Special thanks to the following individuals for their contribution to the 2015 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, Mr. David N. Vickery, and Mr. Michael R. Potter - Universal Consulting Services, Inc  
Ms. Evelyn K. Stephenson, US Army Medical Research Acquisition Activity  
Abstract Reviewers and Session Chairs

### MHSRS Planning

Marriott Harbor Beach Resort and Spa, Ft Lauderdale, FL.

### Supplement Reviewers

<table>
<thead>
<tr>
<th>Ben Antebi, PhD</th>
<th>MAJ Marcus Colyer, USA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Richard Arnold, PhD</td>
<td>Randy Crossland, PhD</td>
</tr>
<tr>
<td>Petr Baranov, PhD, MD</td>
<td>Andrea Crunkhorn, DPT</td>
</tr>
<tr>
<td>LTC Sheryl Bedno, USA</td>
<td>Tianhong Dai, PhD</td>
</tr>
<tr>
<td>MAJ Avi Benov, USA</td>
<td>Tarun Dam, PhD</td>
</tr>
<tr>
<td>LCDR Johathan Berger, USN</td>
<td>Lt Col Susan Dukes, USAF</td>
</tr>
<tr>
<td>Jose Betancourt, PhD</td>
<td>Brian Eastridge, MD, FACS</td>
</tr>
<tr>
<td>Stephane Bibian, PhD</td>
<td>Lt Col Melinda Eaton, USAF</td>
</tr>
<tr>
<td>COL Roman Bilynisky, USA</td>
<td>Paulo Fontes, MD, FACS</td>
</tr>
<tr>
<td>Thomas Blakeman, MSc</td>
<td>Jenny Freeman, MD</td>
</tr>
<tr>
<td>MAJ John Boden, USA</td>
<td>Kathleen Gallaher, MPH</td>
</tr>
<tr>
<td>Michael Boss, PhD</td>
<td>Daniel Gallego-Perez, PhD</td>
</tr>
<tr>
<td>LCDR Matthew Bradley, USN</td>
<td>Dennis Goodes, HSc, MSHIA, MBM, CCC-A</td>
</tr>
<tr>
<td>Richard Branson, MSc</td>
<td>Patrick Grabowski, MPT, PhD</td>
</tr>
<tr>
<td>Surgeon Captain Stephen Bree, Royal Navy</td>
<td>Matthew Hackett, MSE</td>
</tr>
<tr>
<td>Liz Bridges, PhD</td>
<td>LT Jay Haran, USN</td>
</tr>
<tr>
<td>Pam Brown Bear, PhD</td>
<td>CAPT Brett Hart, USN</td>
</tr>
<tr>
<td>COL Richard Burney, USA</td>
<td>Jed Hartings, PhD</td>
</tr>
<tr>
<td>Lynn Caldwell, PhD</td>
<td>Lt Col Jennifer Hatzfeld, USAF</td>
</tr>
<tr>
<td>CAPT Margaret Calloway, USN</td>
<td>Jerry Holmberg, PhD, MS, MT</td>
</tr>
<tr>
<td>Lt Col Jeremy Cannon, USAF</td>
<td>LCDR Dustin Huber, USN</td>
</tr>
<tr>
<td>Kathleen Carlson, MS, PhD</td>
<td>Saber Hussain, PhD</td>
</tr>
<tr>
<td>Howard Champion, MD</td>
<td>MAJ Kyong Hyatt, USA</td>
</tr>
<tr>
<td>Philemon Chan, PhD</td>
<td>Angela Jockheck-Clark, PhD</td>
</tr>
<tr>
<td>Wayne Chappelle, PshD, ABPP</td>
<td>MAJ Jacob Johnson, USA</td>
</tr>
<tr>
<td>Mary Clouser, PhD</td>
<td>Patricia Jordan, PhD</td>
</tr>
</tbody>
</table>
Proceedings of the 2015 Military Health System Research Symposium

Foreword
Terry M. Rauch

Issue Overview
Patricia A. Reilly, Teresa L. Brininger

The 2015 Military Health System Research Symposium Awards
Patricia A. Reilly

Military and Civilian Collaboration: The Power of Numbers

An Orthopedic Performance Improvement Initiative at a Small Military Treatment Facility
John Christopher McGee, Christopher G. Bise, Joel M. Stevens

Reliability, Responsiveness, and Criterion Validity of the Kio Sensor, a New Tool for Assessment of Muscle Function
Patrick Grabowski, Michelle Narveson, Stephen Siegle

The Afghan Theater: A Review of Military Medical Doctrine From 2008 to 2014
Ian Lane, Zsolt Stockinger, Samuel Sauer, Mark Ervin, Michael Wirt, Stephen Bree, Kirby Gross, Jeffrey Bailey, Timothy Hodgetts, Elizabeth Mann-Salinas

Establishing a Joint Theater Trauma System During Phase Zero Operations
Jennifer J. Walker, Zsolt T. Stockinger, Colin G. Chinn

Prehospital Blood Transfusion During Aeromedical Evacuation of Trauma Patients in Israel: The IDF CSAR Experience
Jacob Chen, Avi Benov, Roy Nadler, Daniel N. Darlington, Andrew P. Cap, Ari M. Lipsky, Elon Glassberg

Slack Reducing Band Improves Combat Application Tourniquet Pressure Profile and Hemorrhage Control Rate
Dean Nachman, Avi Benov, Amiram Shovali, Yavnai Nirit, Roy Nadler, Yitzhak Avraham, Elon Glassberg

A Sutureless Vascular Closure Device for Emergent Bovine Xenograft Implantation

The ABC’s of Pancreatic Trauma: Airway, Breathing, and Computerized Tomography Scan?
Matthew Vasquez, Cassandra Cardarelli, Jacob Glaser, Sarah Murthi, Deborah Stein, Thomas Scalea

Measuring Hemodynamic Changes in the Ophthalmic Artery During Applied Force for Noninvasive Intracranial Pressure Monitoring: Test Results in a Porcine Model
Max Twedt, Chase Pfeifer, William Thorell, Greg Bashford

Anesthesia and Postoperative Respiratory Compromise Following Major Lower Extremity Surgery: Implications for Combat Casualties
Samuel M. Galvagno Jr., Jordan Brayanov, George Williams, Edward E. George

Impact of Oxygenation Status on the Noninvasive Measurement of Hemoglobin
Dina Gomaa, Dario Rodriguez Jr., Michael Petro, USAF, Thomas C. Blakeman, Richard D. Branson

Wearable Pulse Oximetry Measurements on the Torso, Arms, and Legs: A Proof of Concept
Marcus Kramer, Aaron Lobbestael, Emily Barten, John Eian, Gregory Rausch

Clinical and Magnetic Resonance Spectroscopic Imaging Findings in Veterans With Blast Mild Traumatic Brain Injury and Post-Traumatic Stress Disorder
Anthony P. Kontos, Anne C. Van Cott, Jodilyn Roberts, Jullie W. Pan, Monique B. Kelly, Jamie McAllister-Deitrick, Hoby P. Hetherington

Validation of Laboratory Animal and Surrogate Human Models in Primary Blast Injury Studies
Namas Chandra, Aravind Sundaramurthy, Raj K. Gupta

Corneal and Corneoscleral Injury in Combat Ocular Trauma from Operations Iraqi Freedom and Enduring Freedom
Anton Vlasov, Denise S. Ryan, Spencer Ludlow, Andrew Coggin, Eric D. Weichel, Richard D. Stutzman, Kraig S. Bower, Marcus C. Hoyer

Frequency Responses to Visual Tracking Stimuli May Be Affected by Concussion
Jun Maruta, Eugene Jaw, Peter Moder, Umesh Rajashekar, Lisa A. Spielman, Jamshid Ghajar

Streamlining Participant Recruitment for TBI and PTSD Research Studies
Shristi Joshi, Kerri Dunbar, Patricia Taylor, Kelli L. Sullivan, Mariam M. Afsal, Christina Song, Maulik Purohit, Michael J. Roy
Expanding Clinical Assessment for Traumatic Brain Injury and Comorbid Post-Traumatic Stress Disorder:  
A Retrospective Analysis of Virtual Environment Tasks in the Computer-Assisted Rehabilitation Environment...  
Marie M. Onakomaiya, Sarah E. Kruger, Krista B. Highland, Paula N. Kodosky, Marcy M. Pape, Michael J. Roy

Long-term Outcomes and Needs of Military Service Members After Noncombat-Related Traumatic Brain Injury  
Kelly J. Miller, Jan E. Kennedy, Karen A. Schwab

Development of a Portable Tool for Screening Neuromotor Sequelae From Repetitive Low-Level Blast Exposure  
Christopher K. Rhea, Nikita A. Kuznetsov, Scott E. Ross, Benjamin Long, Jason T. Jakiela, Jason M. Bailie,  
Matthew A. Yanagi, F. Jay Haran, W. Geoffrey Wright, Rebecca K. Robins, Paul D. Sargent, Joshua L. Duckworth

A Retrospective Cohort Analysis of Battle Injury Versus Disease, Non-Battle Injury—Two Validating Flight  
Surgeons’ Experience...  
William P. Butler, Lawrence W. Steinkraus, Brittany L. Fouts, Jennifer L. Serres

Navy En Route Care: A 3-Year Review of 428 Navy Air Evacuations  
Benjamin Walrath, Alejandra Mora, Victoria Ganem, Stephen Harper, Elliot Ross, Chetan Khared, Gerard Demers,  
Vikhyat S. Bebarta

Use of the Pain Assessment Screening Tool and Outcomes Registry in an Army Interdisciplinary Pain Management  
Center, Lessons Learned and Future Implications of a 10-Month Beta Test...  
Diane M. Flynn, Karon Cook, Michael Kallen, Chester Buckenmaier, Ricke Weickum, Teresa Collins, Ashley Johnson,  
Dawn Morgan, Kevin Galloway, Kristin Jolles

Comparative Efficacy of Multiple Variables of Mesenchymal Stem Cell Transplantation for the Treatment of  
Neuropathic Pain in Rats...  
LiPing Liu, Zhen Hua, Jun Shen, Yan Yin, Jing Yang, Kathleen Cheng, Aijun Liu, Lina Wang, Jianguo Cheng

Chondrotoxicity of Liposomal Bupivacaine in Articular Chondrocytes: Preliminary Findings...  
K. Aaron Shaw, Peter C. Johnson, Steve Zumbrun, Augustine H. Chuang, Craig D. Cameron

Self-Reported Sleep During U.S. Navy Operations and the Impact of Deployment-Related Factors...  
Elizabeth Harrison, Gena L. Glickman, Shiloh Beckerly, Marcus K. Taylor

Physiological Impact of Platelet Apheresis in Pigs: Oxygen Metabolism and Coagulation...  
Wenjun Z. Martini, Cassandra M. Rodriguez, Jonathan Richardson, James K. Aden, Andrew P. Cap, Michael A. Dubick

The Development of a Probabilistic Dose-Response for a Burn Injury Model...  
Anthony Iyoho, Laurel Ng, Philemon Chan

A Novel, Inexpensive Method to Monitor, Record, and Analyze Breathing Behavior During Normobaric Hypoxia  
Generated by the Reduced Oxygen Breathing Device...  
Leonard A. Temme, Paul St Onge, Mark Adams, David L. Still, Jonathan K. Statz, Steven T. Williams

Ultrasound Versus Landmarks for Great Toe Arthrocentesis  
Jason F. Naylor, Kenneth B. Dekay, Benjamin P. Donham, Brian T. Hall

The Efficacy of Two Trabecular Bypass Stents Compared to One in the Management of Open-Angle Glaucoma...  
Anton Vlasov, Won I. Kim

Blast Wave Dynamics at the Cornea as a Function of Eye Protection Form and Fit...  
Steven T. Williams, Thomas H. Harding, J. Keegan Statz, John S. Martin

Does Combat Hearing Preservation Equipment Affect Balance?...  
Art Ambrosio, Ali N. Hoffer, Michael Hoffer

Human Contrast Acuity Variability  
Morris R. Lattimore

Self-Reported Visual Quality of Life After Combat Ocular Trauma...  
Andrew McLaughlin, Marcus H. Colyer, Denise S. Ryan, Rose K. Sia, Eric D. Weichel, Andrew Spiegelman,  
S. G. Eaddy, Robert A. Mazzoli, Kraig S. Bower

Understanding Post-Deployment Reintegration Concerns Among En Route Care Nurses: A Mixed-Methods Approach...  
Felecia M. Rivers, Susan Dukes, Jennifer Hatzfeld, Linda H. Yoder, Sandra Gordon, Angela Simmons

Risk Factors for Incident Postdeployment Mental Health Conditions Among U.S. Air Force Medical Service Personnel...  
Genny M. Maupin, Anthony P. Tvaryanas, Edward D. White, Heather J. Lysjford

Post-traumatic Stress Symptoms in United States Air Force Aeromedical Evacuation Nurses and Technicians...  
Julie M. Swearingen, Tanya M. Goodman, Wayne L. Chappelle, William T. Thompson

Finding Their Way Back In: Family Reintegration Following Guard Deployment...  
Deborah C. Messecar
A Randomized Controlled Trial of Nurse Coaching vs. Herbal Supplementation for Weight Reduction in Soldiers
Mary S. McCarthy, Evelyn B. Elshaw, Barbara M. Szekely, Curtis Hobbs

Physical Screening Predictors for Success in Completing Air Force Phase II Air Liaison Officer Aptitude Assessment
John Christopher McGee, Eric Wilson, Haley Barela, Sharon Blum

Patient Litter System Response in a Full-Scale CH-46 Crash Test
Charles A. Weisenbach, Tyler Rooks, Troy Bowman, Vince Fralish, B. Joseph McEntire

Modeling of Gender Differences in Thermoregulation
Anthony E. Iyoho, Laurel J. Ng, Lisa MacFadden

Proposing Using Waist-to-Height Ratio as the Initial Metric for Body Fat Assessment Standards in the U.S. Army
Stephen A. Bernstein, Michael Lo, W. Sumner Davis

An Augmented Reality-Based Approach for Surgical Telementoring in Austere Environments
Dan Andersen, Voicu Popescu, Maria Eugenia Cabrera, Aditiya Shanghavi, Brian Mullis, Sherri Marley, Gerardo Gomez, Juan P. Wachs

Telesurgery With Miniature Robots to Leverage Surgical Expertise in Distributed Expeditionary Environments
Mark Reichenbach, Tom Frederick, Lou Cubrith, Walter Bircher, Nathan Bills, Marsha Morien, Shane Farritor, Dmitry Oleynikov

Empirical Study on the Impact of a Tactical Biosurveillance Information Visualization on Users' Situational Awareness
Valeriya V. Kettelhut, Trevor C. Vanschooneveld, James C. McClay, David F. Mercer, Ann Fruhling, Jane L. Meza

Military Emergency Medical Service System Assessment: APLICATION of the National Park Service Needs Assessment and Program Audit to Objectively Evaluate the Military EMS System of Okinawa, Japan
Elliott M. Ross, Stephen A. Harper, Cord Cunningham, Benjamin D. Walrath, Gerard DeMers, Chetan U. Kharod

An Assessment of National EMT Certification Among Enlisted Military Medics
Stephen Harper, Remle Crowe, Melissa Bentley, Chetan Kharod, Benjamin Walrath

Public Health Surveillance via Template Management in Electronic Health Records: Tri-Service Workflow's Rapid Response to an Infectious Disease Crisis
Holly Berkley, Matthew Barnes, David Carnahan, Janet Hayhurst, Archie Bockhorst, James Neville

Variation in Postinjury Antibiotic Prophylaxis Patterns Over Five Years in a Combat Zone

Emergency Department Wounds Managed by Combat Medics: A Case Series
Steven G. Schauer, James A. Pfaff

Pertussis Outbreak Among Soldiers During Basic Training: The Need for Updated Protocols
Basheer Halhal, Yuval Glick, Inbal Galor, Ankory Ran, David J. Bacon, Elon Glassberg

Comparative Susceptibility of Different Mouse Strains to Liver-Stage Infection With Plasmodium berghei Sporozoites Assessed Using In Vivo Imaging
Qigui Li, Lisa Xie, Dina Caridiha, Norma Roncal, Qiang Zeng, Jing Zhang, Ping Zhang, Mark Hickman, Lisa Read

Lateral Flow Rapid Test for Accurate and Early Diagnosis of Scrub Typhus: A Febrile Illness of Historically Military Importance in the Pacific Rim
Chien-Chung Chao, Zhiwen Zhangm, Giulia Weissenberger, Hua-Wei Chen, Wei-Mei Ching

Skin Regeneration Using Dermal Substrates that Contain Autologous Cells and Silver Nanoparticles to Promote Antibacterial Activity: In Vitro Studies
Michael A. J. Zieger, Manuel Ochoa, Rahim Rahimi, Gonzalo Campana, Sunil Tholpay, Babak Ziaie, Rajiv Sood

Regenerative and Antibacterial Properties of Acellular Fish Skin Grafts and Human Amnion/Chorion Membrane: Implications for Tissue Preservation in Combat Casualty Care
Skuli Magnusson, Baldur Tumi Baldursson, Hilmar Kjartansson, Ottar Rolfsón, Gudmundur Fertram Sigurjonsson

AMSUS - The Society of Federal Health Professionals should not be held responsible for statements made in its publication by contributors or advertisers. Therefore, the articles reported in this supplement to Military Medicine do not necessarily reflect the endorsement, official attitude, or position of AMSUS or the Editorial Board.
Foreword

Terry M. Rauch, PhD

Military medicine is at a pivot point as it emerges from a decade-plus of combat medical experience. The urgency is high to provide a balanced medical research and development strategy that is responsive to the future capabilities of the warfighter as determined through the Joint Capabilities Integration and Development System and realized through the Defense Acquisition System. By design, exploiting the medical research and development enterprise will drive the modernization of military medicine.

The content of the 2015 Military Health System Research Symposium (MHSRS) reflected the broad scope of medical research and development in the Military Health System (MHS)—from addressing capability gaps realized through the experiences in Iraq and Afghanistan to anticipating medical capabilities needed in future military operations. The 28 breakout sessions covered the spectrum of MHS investments in strategic partnerships and global health engagement, focused on both combat and noncombat medical missions, and emphasized innovative research and development to support the life cycle of our service members.

Each year, the MHSRS grows in stature and importance. Since its first year in 2012, it has emerged as “the” premier scientific forum to integrate and embed emerging evidenced-based practices and technologies into a learning health care system in which the health care providers, scientists, engineers, and patients participate in the generation of knowledge on trends in health and illness, the testing and identification of best practices, and the assessment of the impact of practice changes. This supplement to Military Medicine, along with that of our sister 2015 MHSRS supplement to the Journal of Trauma, are essential tools to communicate the significance of that body of science. I am both pleased and proud to have been invited to lead off this important supplement and recognize the men and women who perform the research and development mission of the MHS.
The first Military Health System Research Symposium (MHSRS) in 2012 marked the merger of the three Service-sponsored research meetings into 1 large joint research symposium under the auspices of Health Affairs. Anchored by the Army’s trauma-focused Advanced Technology Applications to Combat Casualty Care conference, the all-in participation of the Navy/Marine Corps and Air Force provided this new joint assembly with an expanded platform to discuss Department of Defense-sponsored medical research under one unifying goal—improving Service Member care from point-of-injury through rehabilitation. The growth of the MHSRS since 2012 (see Table 1, The 2015 Military Health System Research Symposium Awards in this issue) demonstrates the commitment of DoD-sponsored researchers in the federal government, academia, and industry to advance Service Member care through partnerships, lessons learned, and the utilization of innovative technologies.

This issue starts by recognizing the 2015 MHSRS plenary presentations. In Military and Civilian Collaboration: The Power of Numbers, Skinner et al cite the benefits of maintaining military-civilian partnerships to address major medical capability gaps using the example of the Major Extremity Trauma Research Consortium experience. McGee et al. (An Orthopedic Performance Improvement Initiative at a Small Military Treatment Facility) demonstrate the value of a “physical therapy first” approach to orthopedic care. As one of our 2015 Young Investigators, Patrick Grabowski documents the reliability of the kiio sensor as a measurement for muscle strength (Reliability, Responsiveness, and Criterion Validity of the Kiio Sensor, a New Tool for Assessment of Muscle Function).

The breadth of articles in this Supplement demonstrates how the MHSRS has evolved to reflect the “life cycle” of Service Member care. Many thanks to the authors who submitted manuscripts and to the manuscript reviewers. Each of the articles published here contributes to an accumulating library of medical research data that will inform and guide future military clinicians, scientists, and policy makers as they set the foundation of military medical care pre- and postfuture conflicts.
The 2015 Military Health System Research Symposium Awards

Patricia A. Reilly, PhD

INTRODUCTION
The 2015 Military Health System Research Symposium (MHSRS) continued to break previous records for abstract submission and attendance (Table I). Twenty-four awards recognizing the year’s most outstanding individual and collective research accomplishments were presented at the 2015 meeting. These awardees are highlighted on the following pages. Congratulations to all of the nominees, and special thanks to the supervisors and coworkers who took the time to submit award packages.

DISTINGUISHED SERVICE AWARD
This award is designed to recognize an individual who, over the years, has contributed significantly to the success of military health system research. The 2015 winner is Jay A. Johannigman, Col (USAFR), MD, Director, Trauma and Critical Care, Department of Surgery, University of Cincinnati, Cincinnati, Ohio (Fig. 1).

Citation: Jay A. Johannigman, Col (USAFR), MD, has dedicated his life and efforts to the advancement of medical care for the wounded warrior through his roles as an Individual Mobilization Augmentee to the Offices of the Assistant Secretary of Defense, Deputy Assistant Secretary of Defense, and Force Health Protection and Readiness, and as the Director, Trauma/Critical Care, University of Cincinnati. Early on, he understood the complexity and difficulties inherent within a military medical continuum in need of change and flexibility.

To that end, Dr. Johannigman partnered with peers and, under USAF Surgeon General LtGen (Ret) P.K. Carlton, established the Critical Care Air Transportation Team concept that is credited with a 99% survival rate in military contingencies. Additionally, as a trauma surgeon, he has amassed numerous deployments, employing those skills as a practitioner, leader, and mentor. Dr. Johannigman’s appreciation for these missions, the needs of the caregiver, and the casualty drove him to lead and execute over 40 research studies valued at over 40 million dollars of Department of Defense funding.

Submitted by CMSgt (Ret.) Dario Rodriquez, USAF

OUTSTANDING RESEARCH ACCOMPLISHMENT/INDIVIDUAL
This award is 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. The 2015 winners in the academic and military categories are listed below (Fig. 1).

Academic

Citation: Dr. Kia Washington is recognized for her pioneering work in having successfully created a reliable orthotopic vascularized whole eye transplant small animal model in which gross morphology, retinal blood flow, aqueous humor dynamics, intraocular pressures, and blood-ocular barriers are maintained. The model will provide insights into optic nerve regeneration strategies, allograft tolerance, and functional return. Her project to restore form and function through whole eye transplantation is revolutionary and has the potential to help millions of people with vision loss.

Submitted by Dr. Maxine Miller

Military
LTC Andre Cap, MC USA, Chief, Coagulation and Blood Research, U.S. Army Institute of Surgical Research, San Antonio, Texas.

Citation: For excellence in optimizing Combat Casualty Care, LTC Cap’s sustained leadership in developing the foundational science required to advance the use, storage, and function of platelets for saving lives will continue to impact both military and civilian trauma patients for years to come. This scientific achievement combined with his advocacy for combat casualties and his relentless dedication has greatly advanced the practice of transfusion medicine.

Submitted by Dr. David Baer

OUTSTANDING RESEARCH ACCOMPLISHMENT/TEAM
This award is 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. The 2015 winners in the academic and military categories are listed below (Fig. 2).
Military Team/First Place

Multidrug-resistant Organism Repository and Surveillance Network – Antimicrobial Resistance Monitoring and Research (MRSN-ARMoR), Team Lead: COL Emil Lesho, MC, USA.

Citation: The MRSN-ARMoR team is recognized for its contribution to combating infectious diseases. The team was specifically named by the White House in the 2015 National Action Plan as a key element in the approach to combating antibiotic resistant bacteria. The program was heralded in high-impact medical journals in 2014 as a model for mitigating health care–associated infectious disease in a managed care network. The MRSN-AMRoR team developed a semi-automated sequencing and bioinformatics pipeline that has impacted infection prevention and control efforts throughout the Department of Defense to include in Iraq, Afghanistan, Kuwait, and Honduras. The MRSN-ARMoR is an integrated program between the Army and the Navy, and 100% Defense Health Program funded. Most importantly, the program’s work has helped decrease certain forms of antimicrobial resistance in the Military Health System.

The MRSN-ARMoR team members are:

–From the MRSN (Walter Reed National Military Medical Center): Emil Lesho, Mary Hinkle, Michael Julius, Robert Clifford, Yoon Kwak, Patrick McGann, Erik Snesrud, Michale Sparks, Lakshmi Apalla, Rosslyn Maybank, Ana Ong, Fatima Onmus-Leone, Ian Preston, Eric Steele, Lindsey Nelson, Amanda Roth, Joshua Martinez, Jacob Padilla, LeighAshley Harden, Gerald Ward
–From the U.S. Navy and U.S. Marine Corps Public Health EpiData Center: Uzo Chukwuma, Charlotte Neumann, Kathryn McAuliffe
–From the Global Emerging Infections Surveillance and Response System: Paige Waterman

Submitted by COL Emil Lesho

Military Team/Honorable Mention

National Center for Telehealth and Technology, Joint Base Lewis-McChord, Washington. Team Lead: David D. Luxton, PhD.

Citation: The National Center for Telehealth and Technology team is recognized for the successful completion of the first randomized controlled trial of home-based tele-behavioral health care in the Department of Defense (DoD) and Veterans Administration (VA) health care systems. The study was conducted in response to an identified need within both systems to establish the feasibility of tele-behavioral health treatments that are provided directly to the homes of beneficiaries. This joint DoD and VA research team distinguished itself by independently setting up a research behavioral clinic, preparing safety and technical protocols, and successfully completing the trial on schedule. This high-impact study provided the necessary data regarding the technical feasibility, safety and clinical efficacy of home-based care inform policy decisions that will expand care options for U.S. Service members, Veterans, and their families.

The National Center for Telehealth and Technology team members are: David D. Luxton, Larry D. Pruitt, Amy Wagner, Derek J. Smolenski, Michael A. Jenkins-Guarnieri, Gregory Gahm, Katyna Boykin, Jennifer Green, Kathleen Houston, Kristine Johnson, Elizabeth Speidel, Katherine Stanfill, Lisa Thomas, and Kathleen Woodside.

Submitted by CDR Dennis Faix and Dr. Dennis Goodes

Academic Team/First Place

Rehabilitation Institute of Chicago/Northwestern University, Chicago, Illinois. Team Lead: Levi J. Hargrove, PhD.

Citation: The interdisciplinary team of engineers and clinicians from the Rehabilitation Institute of Chicago/Northwestern University, Chicago, Illinois, is recognized for designing an advanced and intuitive control system for powered prosthetic legs. Unlike current control systems that have limited movement

Table I. Historical Information on the Military Health System Research Symposium (MHSRS)

<table>
<thead>
<tr>
<th></th>
<th>2012</th>
<th>2013</th>
<th>2014</th>
<th>2015</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstracts Submitted</td>
<td>~600</td>
<td>520</td>
<td>1,135</td>
<td>1,196</td>
</tr>
<tr>
<td>Abstracts Accepted</td>
<td>561</td>
<td>516</td>
<td>624</td>
<td>991</td>
</tr>
<tr>
<td>Review Panels</td>
<td>14</td>
<td>20</td>
<td>23</td>
<td>28</td>
</tr>
<tr>
<td>Breakout Sessions</td>
<td>29</td>
<td>25</td>
<td>28</td>
<td>28</td>
</tr>
<tr>
<td>Posters Presented</td>
<td>374</td>
<td>252</td>
<td>387</td>
<td>733</td>
</tr>
<tr>
<td>Young Investigator Submissions</td>
<td>(Presented as 2 Breakout Sessions)</td>
<td>Not a Category</td>
<td>131</td>
<td>184</td>
</tr>
<tr>
<td>Paid Attendees</td>
<td>1,463</td>
<td>1,175</td>
<td>1,524</td>
<td>2,004</td>
</tr>
<tr>
<td>Exhibitors</td>
<td>90</td>
<td>80</td>
<td>76</td>
<td>71</td>
</tr>
<tr>
<td>International Attendees</td>
<td>No Data</td>
<td>No Data</td>
<td>85</td>
<td>84</td>
</tr>
<tr>
<td>Award Nomination Packages</td>
<td>No Data</td>
<td>No Data</td>
<td>29</td>
<td>58</td>
</tr>
</tbody>
</table>

Originating as the U.S. Army sponsored Advanced Technology Applications for Combat Casualty Care meeting in 1993, the name was changed to MHSRS in 2012, when the Assistant Secretary of Defense for Health Affairs became the sponsor and the individual Service research meetings were rolled under the MHSRS. Unlike other years, in 2013 the number of posters to be presented was capped. Introduced first as 2 breakout sessions in 2012, the Young Investigator abstracts were re-invigorated as a competition in 2014. Previous to 2014, award nominations were made as an e-mail message to the planning committee with a paragraph justification.
and are cumbersome to use, this system automatically identifies how the user intends to move (e.g., walking on level ground, up or down stairs) and assists in making natural transitions between these modes. The system is based on signals that are produced when a muscle contracts and on prior gait strides. It does not require surgery or invasive brain-machine implants. In June 2015, the team published the first-ever randomized trial of a powered transfemoral prosthesis in the prestigious Journal of the American Medical Association. This paper joins 14 other 2014–2015 publications that were the direct result of DoD support and which generated extensive media coverage.

The Rehabilitation Institute of Chicago/Northwestern University team members are Levi J. Hargrove, Ann M. Simon,
Aaron J. Young, Nicholas P. Fey, Suzanne B. Finucane, Elizabeth G. Halsne, Kimberly A. Ingraham, Todd A. Kuiken, and John Spanias.

Submitted by Dr. Todd Kuiken

Academic Team/Honorable Mention
U.S. Army Medical Department Center and School, Health Readiness Center of Excellence, San Antonio, Texas. Team Lead: John D. Childs, PhD.

Citation: This team is recognized for their accomplishments in the field of health services research. In 2015, the authors completed a comprehensive analysis across the Military Health System of 753,450 patients with low back pain (LBP). The goal was to evaluate the impact of early and guideline adherent physical therapy on utilization and cost. Utilization outcomes (e.g., advanced imaging, lumbar injections or surgery, and opioid use) and associated costs were examined based on the timing of referral for physical therapy and for adherence to practice guidelines over a 2-year period. The data indicated that early referral to guideline adherent physical therapy for LBP was associated with significantly lower utilization for advanced imaging, lumbar injections or surgery, and opioid use and resulted in 60% lower total costs. The potential MHS-wide cost savings associated with early guideline adherent physical therapy for LBP are substantial. The results of this study will inform future policy and shape optimal processes of care for patients with LBP across the MHS.

The U.S. Army Medical Department and School, Health Readiness Center of Excellence team members: John D. Childs, Julie, M. Fritz, Samuel S. Wu, Timothy W. Flynn, Robert S. Wainner, Erik K. Robertson, Forest S. Kim, and Steven Z. George.

Submitted by LTC Forest Kim

YOUNG INVESTIGATOR COMPETITION
This category is designed to highlight/promote the research accomplishments of residents, fellows, post-docs within 5 years of graduation from a terminal degree, and Service Academy cadets. In 2015, 184 abstracts were submitted to this category. After a 2-tier review process, 12 abstracts were selected for oral presentation at a special plenary session. The competition winners are listed below and pictured in Figure 3.

First place: Surg Lt CDR Ed Barnard, FCEM, for his presentation: Selective Aortic Arch Perfusion for the Reversal of Hemorrhage-induced Traumatic Cardiac Arrest in a Swine Model of Noncompressible Torso Hemorrhage.

Second place: CPT George E. Black, MC, USA, for his presentation: Suspension of Biologic Time in Severe Hemorrhagic Shock: Pilot Study Results from the Biochronicity Project.

Third place: Srinivas Laxminarayan, PhD, for his presentation: Preventing Heat Injuries by Predicting Individualized Human Core Temperature.

MHSRS POSTER COMPETITION
A total of 733 posters were presented over 3 poster sessions. The awards winners are listed in Table II, and pictured in Figure 4.
<table>
<thead>
<tr>
<th>Award</th>
<th>Title</th>
<th>Authors (* = Presenting Author)</th>
<th>Institutional Affiliation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Best in Show</td>
<td>A Probabilistic Concussion Model Relating Head Impact</td>
<td>Laurel Ng*, Melissa Gibbons, Vladislav Volman, Pi Phohomsiri,</td>
<td>L-3 Communications, San Diego, CA</td>
</tr>
<tr>
<td></td>
<td>Kinematics to Internal Axonal Level Injury and Clinical</td>
<td>Durrell Swenson, Jianxia Cui, James Stuhlmiller</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Outcomes for Interpretation of Wearable Sensor Data</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poster Session 1</td>
<td>First Place</td>
<td>Salvatore Libretto*, Sandra Gordon, Lara Hilton, Courtney Boyd,</td>
<td>Samueli Institute, Alexandria, VA</td>
</tr>
<tr>
<td></td>
<td>Evaluation of an Integrative PTSD Treatment Program for</td>
<td>Weimin Zhang, Kimberly McConnell</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Active Duty Service Members</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Second Place</td>
<td>John Evans*, Carol Jones, Debra Murphy, Thomas DeGraba</td>
<td>Naval Hospital, Camp Lejeune, NC</td>
</tr>
<tr>
<td></td>
<td>Establishing Large Scale Data Repositories in a Military</td>
<td>1Naval Hospital, Camp Lejeune, NC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Health Care Setting</td>
<td>2National Centers of Excellence</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Third Place</td>
<td>Glen Michael Fitzpatrick*, Arthur Bode, Veronica Ortiz, Janelle</td>
<td>Cellphire Inc., Rockville, MD</td>
</tr>
<tr>
<td></td>
<td>Immune-Masking of Thrombosomes: Trehalose Stabilized Platelet-Derived</td>
<td>Ober</td>
<td>George Washington University,</td>
</tr>
<tr>
<td></td>
<td>Hemostatic Agent</td>
<td></td>
<td>Washington, DC</td>
</tr>
<tr>
<td></td>
<td>Honorable Mention</td>
<td>Transient Alterations of Cutaneous Sensory Nerve Function</td>
<td>Harvard Medical School, Boston, MA</td>
</tr>
<tr>
<td></td>
<td>by Non-Invasive Cryolipolysis</td>
<td>by Non-Invasive Cryolipolysis</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pelvic Fracture Is Strongly Associated with Lower Limb Traumatic</td>
<td>Jason Forman*, Ed Barnard, Vitoria Ganem, Alejandra Mora, Vikhyat</td>
<td>US Army Institute of Surgical Research,</td>
</tr>
<tr>
<td></td>
<td>Amputation in Combat Related Blast – A Prospective, Multicenter,</td>
<td>Bebarta</td>
<td>Fort Sam Houston, TX</td>
</tr>
<tr>
<td></td>
<td>Observational Study</td>
<td></td>
<td>1US Army Institute of Surgical Research,</td>
</tr>
<tr>
<td></td>
<td>First Place</td>
<td>Evaluation of Kollidon VA64 in the WRAIR PBBI Model:</td>
<td>2Institute of Naval Medicine, Gosport,</td>
</tr>
<tr>
<td></td>
<td>Studies from the Operation Brain Trauma Therapy</td>
<td>Krista Caudle*, Stefania Mondello, Janice Gilsdorf, Frank Tortella,</td>
<td>United Kingdom</td>
</tr>
<tr>
<td></td>
<td>Second Place</td>
<td>Deborah Shear</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Surface Enhanced Raman Spectroscopy as a Point-of-Care Diagnostic for</td>
<td>Nicole J. Crane*, Shubha Yesupriya, Meron Ghebremedhin</td>
<td>Naval Medical Research Center</td>
</tr>
<tr>
<td></td>
<td>Infection in Wound Effluent</td>
<td></td>
<td>1Walter Reed Army Institute of Research, Silver</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Spring, MD</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>University of Messina, Italy</td>
</tr>
</tbody>
</table>

(continued)
<table>
<thead>
<tr>
<th>Award</th>
<th>Title</th>
<th>Authors (* = Presenting Author)</th>
<th>Institutional Affiliation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Third Place</td>
<td>Assessing Objective Neuromarkers in an Immersive Virtual Reality Environment</td>
<td>Brennan Cox, PhD(^1)*, Katherine Service, BS(^1), Seth Reini, PhD(^2), Weimin Zheng, PhD(^1)</td>
<td>(^1)Naval Health Research Center, San Diego, CA (^2)Navy Experimental Diving Unit, Panama City Beach, FL</td>
</tr>
<tr>
<td>Honorable Mention</td>
<td>Differential Contributions of Platelets and Fibrinogen to Early Hyperfibrinolysis During Hemorrhage and Shock and Its Correction with Small-Volume IV Bolus of 7.5% NaCl Adenosine, Lidocaine and Mg(^2+) (ALM)</td>
<td>Geoffrey P. Dobson(^*), Hayley L. Letson</td>
<td>James Cook University, Townsville, Queensland, Australia</td>
</tr>
<tr>
<td>Honorable Mention</td>
<td>Penetrating Ballistic-Like Brain Injury in Rats</td>
<td>Sindhu Madathil(^*), Lai Yee Leung, Katherine Cardiiff, Xiaofang Yang, Frank Tortella, Deborah Shear, Ying Deng-Bryant</td>
<td>Walter Reed Army Institute of Research, Silver Spring, MD</td>
</tr>
<tr>
<td>Poster Session 3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First Place</td>
<td>Burn Scars Modulation Through Laser Deliver of Stem Cells</td>
<td>Luis Rodriguez-Menocal(^1)*, M. Salgado(^1), R.J. Christy(^2), S. Becerra(^2), J. Gill(^1), A. Candalado(^1), S. Natesan(^1), J. Valdes(^1), M. Solis(^1), C.I. Schulman(^1), J.S. Waibel(^1), S.C. Davis(^1), E.V. Badiavas(^1)</td>
<td>(^1)University of Miami, Miami, FL (^2)US Army Institute of Surgical Research, San Antonio, TX</td>
</tr>
<tr>
<td>Second Place</td>
<td>Skin Regeneration Using Dermal Substrates that Contain Autologous Cells and Silver Nanoparticles to Promote Bactericidal Activity - In vitro Studies</td>
<td>Michael Zieger(^1)*, Rahim Rahimi(^2), Manuel Ochoa(^1), Gonzalo Campa(^1), Sunil Tholpady(^1), Rajiv Sood(^1), Babak Ziaie(^1)</td>
<td>(^1)Indiana University School of Medicine, Indianapolis, IN (^2)Purdue University, West Lafayette, IN</td>
</tr>
<tr>
<td>Third Place</td>
<td>Fleet Disease/Non-Battle Injuries and ICD-9 Diagnosis Code Matching in Theatre Medical Data</td>
<td>Chelsea Saia(^*), Tina Luse</td>
<td>Navy and Marine Corps Public Health Center, Portsmouth, VA</td>
</tr>
<tr>
<td>Honorable Mention</td>
<td>Antibiotic-Loaded Keratin Hydrogels as a First-line Therapy for Battlefield Burns</td>
<td>Daniel Roy(^1,2)*, Christine Kowalczewska(^1,2), Kameel Isaac(^1), Lake Burner(^1), Seth Tobmlyn(^1), Robert Christy(^1)</td>
<td>(^1)US Army Institute of Surgical Research, San Antonio, TX (^2)KeraNetics, LLC, Winston-Salem, NC</td>
</tr>
<tr>
<td>Honorable Mention</td>
<td>Global Health Engagement Missions: Lessons Learned Aboard U.S. Naval Hospital Ships-A Qualitative Analysis</td>
<td>Heather King(^1)*, Patricia Kelley(^2), John Malone(^1), Christine Johnson(^1), Gregg Montak(^1), Monique Bouvier(^1), Michael Coronado(^1)</td>
<td>(^1)Naval Medical Center San Diego, San Diego, CA (^2)US Department of Veterans Affairs Office of Research &amp; Development, Washington, DC</td>
</tr>
</tbody>
</table>
FIGURE 4. The 2015 MHSRS poster award winners shown with RADM Doll. Top left: Best in Show winner Laurel Ng, PhD. Top right: Poster session 1 third place winner, Glen Michael Fitzpatrick, PhD. Bottom left: Poster session 2 first place winner, Krista Caudle, PhD. Bottom right: Poster session 2 third place winner, Brennan Cox, PhD.
Military and Civilian Collaboration: The Power of Numbers

MAJ Daniel J Stinner, MC USA*; Joseph C Wenke, PhD†; COL James R. Ficke, MC USA (Ret.)‡; Lt Col Wade Gordon, USAF MC§; CAPT James Toledano, MC USN∥; Anthony R. Carlini, MS∥∥; Daniel O Scharfstein, ScD**; Ellen J. MacKenzie, PhD¶; Michael J. Bosse, MD††; Joseph R. Hsu, MD††; the Major Extremity Trauma Research Consortium (METRC)††

ABSTRACT The purpose of this study was to compare the number and types of extremity injuries treated at civilian trauma centers (CIV CENs) versus military treatment facilities (MTFs) participating in the Major Extremity Trauma Research Consortium (METRC) and to investigate the potential benefits of a clinical research network that includes both civilian trauma centers and MTFs. Two analyses were performed. First, registry data collected on all surgically treated fractures at four core MTFs and 21 CIV CENs over one year were compared. Second, actual numbers and distribution of patients by type of injury enrolled in three METRC studies were compared. While MTFs demonstrated higher percentages of severe injuries including open fractures, traumatic amputations, vascular injuries, contamination, and injuries with bone, muscle, and skin loss when compared to CIV CENS, the CIV CENs treated a substantially higher number and, more importantly, enrolled patients in almost all categories. Comparison of service members to civilians was challenged by several differences between the two patient populations including mechanism of injury, the medical care environment, and confounding factors such as age, social setting and co-morbidities. Despite these limitations, in times without active military conflict, clinical trials will likely rely on civilian trauma centers for patient enrollment; only when numbers are pooled across a large number of centers can requisite sample sizes be met. These data demonstrate the benefits of maintaining a military–civilian partnership to address the major gaps in research defined by the Military.

INTRODUCTION

The burden of extremity injuries is well described from the recent conflicts in Iraq and Afghanistan with 82% of injured Servicemen and women sustaining extremity injuries.1 These injuries account for two-thirds of the hospital and disability costs associated with all serious combat trauma.2 The challenges in treating extremity war injuries have been well-documented and are re-evaluated on an annual basis.3,4 These challenges include: timing and staging of treatment, severe physiologic stress, complex wound management, contamination, infection, bone loss, articular surface involvement, heterotopic ossification, segmental nerve loss, volumetric muscle loss, and compartment syndrome. Challenges do not end with the acute phase of treatment; most patients require rehabilitation and postacute services to optimize recovery and reintegration back into military service and everyday life. Furthermore, questions and challenges of limb salvage versus amputation still exist. For many patients, secondary adverse health effects (e.g., osteoarthritis, cardiovascular disease, obesity, chronic pain, and depression) can develop in the years following injury, which can have significant implications over the course of their remaining life as this is a predominantly young patient population.

The majority of the available evidence guiding our treatment of these severe combat-related extremity injuries is based on Level IV data generated from small studies that often lack a comparison group.5–9 Even though these complex injuries account for a high percentage of all injuries sustained in military combat operations, their absolute numbers are modest, making it difficult to adequately power prospective studies that rely only on injuries treated at military treatment facilities (MTFs). Enhancing collaborations
between the military and civilian trauma communities and combining the population of injured Service Members with the large numbers of patients treated at civilian Level I trauma centers has been posited as an effective strategy for addressing this challenge. One criticism of this approach, however, has been the concern that civilian injuries are not comparable in complexity to injuries sustained in combat. While extremity injuries account for two-thirds of all injuries treated at trauma centers in the United States, the majority result from blunt mechanisms, whereas blasts and high-velocity gunshot wounds are responsible for over 80% of combat casualties. In one of the largest reviews describing combat extremity injuries, Owens and colleagues found that of 3,573 extremity wounds sustained in 1,281 soldiers, 915 (26%) were fractures, with 758 (82%) being open fractures. In contrast, the majority of fractures treated at civilian trauma centers are closed injuries. As a result of the differences in mechanism and severity of injury, as well as the lack of infrastructure for coordinating studies across multiple centers, collaborative prospective research between military and civilian trauma centers has been limited.

The Major Extremity Trauma Research Consortium (METRC) was funded in 2009 by the U.S. Department of Defense to help build this infrastructure and evaluate whether a large enough number of patients treated at civilian trauma centers could populate studies that address current priorities for combat casualty care. The consortium has grown to include the four MTFs that treat the majority of combat casualties, 22 Core Civilian Centers (CIV CENs), and over 30 Satellite Centers throughout the United States. The civilian core centers are large, level I trauma centers chosen on the basis of their volume of extremity trauma, research experience, and academic qualifications. The satellite clinical centers are level I/II trauma centers with lower patient volumes but are important for ensuring adequate numbers and appropriate mix of patients.

The purpose of this study is to compare the distribution of extremity injuries by type and severity treated at civilian centers versus military centers participating in METRC and to investigate the potential for a military–civilian partnership to address existing prioritized research objectives. We hypothesized that including civilian centers with MTFs would increase the number of injuries of interest to the military that are available for study. We also hypothesized that the actual numbers of patients enrolled in specific studies would be substantially higher if patients from both civilian as well as military centers were included.

### METHODS

The study consisted of two analyses based on two different data sources to investigate both hypotheses. In the first analysis, registry data collected on all surgically treated fractures at the core military and civilian centers over a 1-year period were compared. In the second analysis, the actual numbers of patients enrolled in three of the core METRC studies were compared. The protocol for the registry and the three core METRC studies were approved by the Johns Hopkins Bloomberg School of Public Health IRB, the DoD Human Research Protections Office (HRPO), and the local IRB at each participating center.

### Registry Comparisons

In 2010, METRC established a ‘start-up’ registry at the core military and civilian centers to provide estimates of the number and mix of patients potentially available for study recruitment. The registry has been helpful in determining study feasibility and for monitoring enrollment once a study is active. The registry consists of prospectively collected data for trauma admissions from 21 Core CIV CENs and four MTFs (one core civilian center did not participate in the registry). The registry was not initiated at satellite centers. Patients that were eligible for inclusion in the registry were 18 to 84 years of age who were surgically treated for fractures of the upper or lower extremity, pelvis or acetabulum. Excluded were hip fractures in patients 60 years or older, and fractures to the wrist, hand, ankle, clavicle, patella, and the foot other than calcaneus/talus/crush. Each center was asked to collect a minimum data set for all patients meeting the inclusion criteria over a consecutive 12-month period of time between 2010 and 2011. The distributions of fracture characteristics were compared between military and civilian centers. Fractures were classified by upper versus lower extremity, closed vs. open fracture, and by location and type of fracture using the AO/OTA Fracture Classification. Open fractures were further characterized using the Gustilo–Anderson classification and the OTA Open Fracture Classification. Overall comparisons between military and civilian centers were made using permutation tests that account for clustering of fractures by center, but not fractures within individuals. Comparisons were considered to be statistically significant if \( p < 0.05 \).

### Study Enrollment Comparisons

To compare actual enrollment by military and civilian centers in studies of priority to the Department of Defense and its combat casualty program, numbers of patients enrolled in three METRC studies were summarized by center. These studies are briefly described below:

- The FIXIT Study is a prospective clinical trial that consists of two components (randomized and observational studies) to compare two options for treating severe tibia fractures: internal fixation with a plate or an intramedullary nail versus circular external fixation. Main inclusion criteria were: Gustilo type IIIA and selected severe Gustilo type IIIA diaphyseal or metaphyseal tibia fractures. Study enrollment began in 2011 and is ongoing.
- The OUTLET Study is a prospective observational study comparing functional outcomes in patients undergoing limb salvage or amputation for severe distal tibia,
ankle, and/or foot injuries. Main inclusion criteria were: selected open type III pilon and foot/ankle or severe open or closed crush or blast foot injuries. Study enrollment began in 2012 and ended in 2015.

- The BIOBURDEN Study is a prospective observational study to characterize bacteria present (i.e., the bioburden) at the time of wound closure in severely contaminated extremity wounds and to correlate the initial “bioburden” to subsequent infections. Main inclusion criteria were: open type III tibia fracture (plateau, shaft, and pilon) requiring a second procedure following fixation, or traumatic amputation requiring delayed primary closure, skin grafting and/or flap coverage. Study enrollment began in 2012 and ended in 2015.

For purposes of comparison, baseline fracture characteristics for open tibia and/or ankle fractures (AO/OTA 41, 42, 43, and 44) were included in this analysis. As with the registry data, distribution of fracture characteristics were compared between the MTFs and CIV CENs.

Finally, in an effort to predict realistic enrollment expectations at MTFs and CIV CENs for future collaborative studies, registry data were compared with enrollment data for a subset of injuries of particular interest in meeting established research priorities (Gustilo type III open tibia fractures).

RESULTS

Registry Comparisons

A total of 875 fractures (among 613 individuals) were registered at the four MTFs over one year; 14,362 (among 11,487 individuals) were registered at the 21 CIV CENs. Number of fractures registered per month ranged from 6 to 27 among the four MTFs (average of 13.0 per center) and from 20 to 84 for the 21 CIV CENs (average of 45.8 per center) (Fig. 1). Of the 875 fractures treated at MTFs, 168 (19.2 %) presented as traumatic amputations compared to only 62 (0.4%) at the civilian trauma centers. These traumatic amputations are excluded from subsequent tabulations and comparisons of specific fracture characteristics.

A significantly larger percentage of fractures treated at the MTFs were open (34.6%) compared to CIV CENs (22.2%) \((p < 0.01)\). The larger percentage of open fractures in the MTFs versus CIV CENs is consistent across fracture location and fracture type (Table I) although differences are not always statistically significant after taking into account the wide variation in fracture distribution across centers. Of note, however, is that despite larger percentages of fractures that are open at the MTFs, the average number of open fractures per site at the civilian centers (152 per year per site) is more than two times the average across the MTFs (62 per year per site) (Table II). When cases are combined across all centers, the civilian centers treated 9.6 times more open humerus fractures, 10.0 times more open radius/ulna fractures, 10.7 times more open pelvic/acetabulum fractures, 19.9 times more open femur fractures, 14.0 times more open tibia fractures, and 9.8 times more open foot fractures when compared to those treated at MTFs.

The percentages of Gustilo–Anderson type III fractures (of all fractures) were also higher at MTFs compared to CIV CENs (upper extremity: 25.5% vs. 8.2%; lower extremity: 21.3% vs. 10.9%) (Table II). Although the average per center number of type III open fractures to the upper extremities registered over one year was the same for MTFs and civilian trauma centers (12 per year per center), the average per-center number of type III open fractures to the lower extremity was over two times higher at the CIV CENs versus MTFs (59 versus 28 per year per center) (Table II). The total number of type III open fractures treated across all 21 CIV CENs was substantially higher for both upper and lower extremity fractures: 247 vs. 49 type III upper extremity fractures and 1,235 versus 110 type III lower extremity fractures. Of the severe open fractures in the lower extremity, there were 727 type III open tibia fractures treated at the 21 CIV CENs (IIIA = 462, IIIB = 216, IIIC = 49) and only 67 treated at the four MTFs (IIIA = 36, IIIB = 23, IIIC = 8).

The distribution of open fractures by the OTA Open Fracture Classification is presented in Table III. As above, the MTFs as a group (when compared to the civilian trauma centers) registered a larger percentage of the more severe injuries defined by segmental bone loss >2 cm, extensive muscle necrosis, extensive de-gloving, artery injury with ischemia, and contamination embedded in bone. The average per-center numbers of severe injuries treated at the MTFs are similar to the averages in the civilian centers except for contamination; on average, the individual MTFs are treating, on average, 50% more open fractures with contamination embedded in bone compared to civilian centers. Taken together, the total numbers of the most severe injuries treated at CIV CENs were substantially higher than those treated at MTFs (segmental bone loss >2 cm: CIV CENs = 200, MTFs = 36; extensive muscle necrosis: CIV CENs = 259, MTFs = 44; extensive degloving: CIV CENs = 245, MTFs = 36; artery injury with ischemia: CIV CENs = 121, MTFs = 13; contamination embedded in bone, CIV CENs = 348, MTFs = 95) (Table III).

Study Enrollment Comparisons

A total of 1,199 patients were enrolled across the three METRC studies described above, with the civilian centers contributing 1,131 (94%) of the total. Included in this total are 67 traumatic amputations (12 enrolled by the MTFs and 55 enrolled by the civilian centers). These numbers stand in stark contrast to the annual number of traumatic amputations entered into the registry (168 by MTFs and 64 by CIV CENs).

The distribution of fractures by categories of the OTA-OFC shows similar patterns found with the registry data; while the percentage of more severe injuries are higher in the MTFs (except for degloving), both the average per site
and the overall number of severe extremity injuries enrolled is substantially higher in the CIV CENs compared to the MTFs (total numbers for segmental bone loss >2 cm: CIV CENs = 198, MTFs = 14; extensive muscle necrosis: CIV CENs = 198, MTFs = 15; extensive degloving: CIV CENs = 244, MTFs = 12; artery injury with ischemia: CIV CENs = 64, MTFs = 4; contamination embedded in bone: CIV CENs = 284, MTFs = 21) (Table IV).

The distribution of Gustilo–Anderson type III open tibia fractures enrolled across the three studies is summarized in Table V for the MTFs and CIV CENs. The three MTFs together enrolled 39 Gustilo–Anderson type III tibia fractures (19 type IIIA; 16 type IIIB; and 4 type IIIC) compared to 822 enrolled by the CIV CENs (367 type IIIA; 411 type IIIB; and 44 type IIIC). In addition, substantially more severe tibia fracture patterns AO/OTA “C” type fractures were enrolled at CIV CENs than MTFs (CIV CENs = 444, MTFs = 19) (Table V).

Finally, while MTFs accounted for 8.4% of all 794 type III open tibia fractures included in the registry, they only accounted for 4.5% of the 861 type III open tibia fractures enrolled in three large prospective research studies.

**DISCUSSION**

The Department of Defense (DoD) Extremity Trauma Prioritized Research agenda was announced more than a decade ago. Although significant research has contributed to incremental improvements in the evaluation and care of severe extremity injuries, significant knowledge gaps remain. The data presented in this article demonstrate the benefits of a collaboration with civilian trauma centers to conduct prospective clinical research relevant to these research objectives.

Although the four MTFs participating in METRC treat higher rates of severe injuries critical to the prioritized research objectives to include open fractures, traumatic amputations, vascular injuries, contamination, and injuries with bone, muscle, and skin loss when compared to CIV CENs, the CIV CENs treat a much higher number of these patients, both on average per center and overall. It must be noted that the four MTFs participating in METRC not only treat nearly all combat casualties, they also treat, on average, higher acuity patients than other MTFs, making them ideal for collaboration with CIV CENs. This collaboration is important because CIV CENs can enroll much higher numbers of patients (both per center and overall) than even the higher volume MTFs in almost all injury categories. Although comparisons of total numbers of treated and enrolled are based on five times as many CIV CENs (21 centers) versus MTFs (four centers) it is important to keep in mind that the MTF numbers represent the vast majority of combat casualties (i.e., no other MTF currently treats the more severe injuries). Thus a comparison of total numbers illustrates the real potential of a military–civilian partnership.

The low MTF numbers clearly demonstrate that the capacity to conduct independent prospective clinical research of severe extremity trauma relying solely on patients treated at the MTFs is limited. The absolute number of these severe injuries treated at the CIV CENs; however, is 3.5 to 11 times higher than those treated at MTFs. This highlights the fact that there is strength in numbers and that in order to conduct meaningful extremity war injury research, collaboration with civilian level I trauma centers is critical.
Currently, the majority of the available evidence guiding our treatment of severe combat-related extremity injuries consists of Level IV data with small sample sizes, often without a comparison group. Keeling and colleagues, for example, demonstrated excellent results in the treatment of severe open tibia fractures with the use of a circular external fixator. However, this study only consisted of 35 patients and lacked a comparison group. Their results were one of the factors that influenced the development of the FIXIT trial now being conducted by the METRC Consortium, which is comparing circular external fixation to internal fixation in severe open tibia fractures. FIXIT has already enrolled over 350 patients in an effort to definitively determine the optimal method of treatment for these severe injuries.

The criticism that civilian injuries are not comparable to injuries sustained in combat is not supported by the data. Although civilian trauma patients are not commonly injured by blasts or high-velocity gunshot injuries, their injuries do share many characteristics that make them ideal surrogates for testing treatment protocols.

Currently, the majority of the available evidence guiding our treatment of severe combat-related extremity injuries consists of Level IV data with small sample sizes, often without a comparison group. Keeling and colleagues, for example, demonstrated excellent results in the treatment of severe open tibia fractures with the use of a circular external fixator. However, this study only consisted of 35 patients and lacked a comparison group. Their results were one of the factors that influenced the development of the FIXIT trial now being conducted by the METRC Consortium, which is comparing circular external fixation to internal fixation in severe open tibia fractures. FIXIT has already enrolled over 350 patients in an effort to definitively determine the optimal method of treatment for these severe injuries.

The criticism that civilian injuries are not comparable to injuries sustained in combat is not supported by the data. Although civilian trauma patients are not commonly injured by blasts or high-velocity gunshot injuries, their injuries do share many characteristics that make them ideal surrogates for testing treatment protocols.

### TABLE I

<table>
<thead>
<tr>
<th>Bone Fractured and OTA Code</th>
<th>MTFs (n = 4)</th>
<th>Number Closed</th>
<th>Number Open (Percent Open)</th>
<th>CIV CENs (N = 21)</th>
<th>Number Closed</th>
<th>Number Open (Percent Open)</th>
<th>p Value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Upper Extremities</td>
<td>113</td>
<td>79 (41.1%)</td>
<td></td>
<td>2,244</td>
<td>777 (25.7%)</td>
<td></td>
<td>0.05</td>
</tr>
<tr>
<td>Humerus</td>
<td>52</td>
<td>31 (37.3%)</td>
<td></td>
<td>1,200</td>
<td>298 (19.9%)</td>
<td></td>
<td>0.07</td>
</tr>
<tr>
<td>11A,B,C</td>
<td>25</td>
<td>2</td>
<td></td>
<td>533</td>
<td>36</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12A,B,C</td>
<td>16</td>
<td>6</td>
<td></td>
<td>376</td>
<td>109</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13A,B,C</td>
<td>11</td>
<td>23</td>
<td></td>
<td>291</td>
<td>153</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Radius/Ulna</td>
<td>61</td>
<td>48 (44.4%)</td>
<td></td>
<td>1,043**</td>
<td>479 (31.5%)</td>
<td></td>
<td>0.10</td>
</tr>
<tr>
<td>21A,B,C</td>
<td>25</td>
<td>25</td>
<td></td>
<td>487</td>
<td>177</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22A,B,C</td>
<td>36</td>
<td>23</td>
<td></td>
<td>556</td>
<td>302</td>
<td></td>
<td></td>
</tr>
<tr>
<td>All Lower Extremities</td>
<td>349</td>
<td>166 (32.2%)</td>
<td></td>
<td>8,884</td>
<td>2,395 (21.2%)</td>
<td></td>
<td>0.01</td>
</tr>
<tr>
<td>Pelvis/Acetabulum</td>
<td>51</td>
<td>7 (12.1%)</td>
<td></td>
<td>1,994</td>
<td>75** (3.6%)</td>
<td></td>
<td>0.05</td>
</tr>
<tr>
<td>61A,B,C</td>
<td>29</td>
<td>5</td>
<td></td>
<td>947</td>
<td>19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>62A,B,C</td>
<td>22</td>
<td>2</td>
<td></td>
<td>1,047</td>
<td>56</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Femur</td>
<td>98</td>
<td>31 (24.0%)</td>
<td></td>
<td>2,972</td>
<td>617 (17.2%)</td>
<td></td>
<td>0.25</td>
</tr>
<tr>
<td>31A,B,C</td>
<td>38</td>
<td>4</td>
<td></td>
<td>1,005</td>
<td>37</td>
<td></td>
<td></td>
</tr>
<tr>
<td>32A,B,C</td>
<td>41</td>
<td>16</td>
<td></td>
<td>1,431</td>
<td>350</td>
<td></td>
<td></td>
</tr>
<tr>
<td>33A,B,C</td>
<td>19</td>
<td>11</td>
<td></td>
<td>536</td>
<td>230</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tibia</td>
<td>135</td>
<td>105 (43.8%)</td>
<td></td>
<td>3,152</td>
<td>1,471** (31.8%)</td>
<td></td>
<td>0.05</td>
</tr>
<tr>
<td>41A,B,C</td>
<td>55</td>
<td>15</td>
<td></td>
<td>1,328</td>
<td>182</td>
<td></td>
<td></td>
</tr>
<tr>
<td>42A,B,C</td>
<td>32</td>
<td>71</td>
<td></td>
<td>939</td>
<td>880</td>
<td></td>
<td></td>
</tr>
<tr>
<td>43A,B,C</td>
<td>48</td>
<td>19</td>
<td></td>
<td>885</td>
<td>409</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Foot</td>
<td>65</td>
<td>23 (36.1%)</td>
<td></td>
<td>765**</td>
<td>226** (22.8%)</td>
<td></td>
<td>0.73</td>
</tr>
<tr>
<td>81A,B,C</td>
<td>20</td>
<td>10</td>
<td></td>
<td>238</td>
<td>81</td>
<td></td>
<td></td>
</tr>
<tr>
<td>82A,B,C</td>
<td>44</td>
<td>13</td>
<td></td>
<td>500</td>
<td>108</td>
<td></td>
<td></td>
</tr>
<tr>
<td>89A,B,C</td>
<td>1</td>
<td>0</td>
<td></td>
<td>27</td>
<td>37</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*p values are for comparisons made between percentages of open fractures within the specified category. **In eight cases the exact location of the bone was not recorded (including one radius/ulna fracture; one pelvic fracture; three tibia fractures, and three foot fractures).

### TABLE II

<table>
<thead>
<tr>
<th>Injury Type</th>
<th>Total Number (Average per Site)</th>
<th>Percent Distribution</th>
<th>p Value*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MTFs (n = 4)</td>
<td>CIV CENs (n = 21)</td>
<td></td>
</tr>
<tr>
<td>All Upper and Lower Extremities</td>
<td>707 (176)</td>
<td>14,300 (686)</td>
<td>100.00</td>
</tr>
<tr>
<td>All Closed</td>
<td>462 (116)</td>
<td>11,128 (534)</td>
<td>65.3</td>
</tr>
<tr>
<td>All Open Gustilo I/II</td>
<td>86 (22)</td>
<td>1,690 (81)</td>
<td>12.2</td>
</tr>
<tr>
<td>All Open Gustilo III</td>
<td>159 (40)</td>
<td>1,482 (71)</td>
<td>22.5</td>
</tr>
<tr>
<td>Upper Extremities</td>
<td>192 (48)</td>
<td>3,021 (145)</td>
<td>100.00</td>
</tr>
<tr>
<td>Upper Closed</td>
<td>113 (28)</td>
<td>2,244 (108)</td>
<td>58.9</td>
</tr>
<tr>
<td>Upper Open Gustilo I/II</td>
<td>30 (8)</td>
<td>530 (25)</td>
<td>15.6</td>
</tr>
<tr>
<td>Upper Open Gustilo III</td>
<td>49 (12)</td>
<td>247 (12)</td>
<td>25.5</td>
</tr>
<tr>
<td>Lower Extremities</td>
<td>515 (129)</td>
<td>11,279 (541)</td>
<td>100.00</td>
</tr>
<tr>
<td>Lower Open Gustilo I/II</td>
<td>349 (87)</td>
<td>8,884 (426)</td>
<td>67.8</td>
</tr>
<tr>
<td>Lower Open Gustilo III</td>
<td>56 (14)</td>
<td>1,160 (56)</td>
<td>10.9</td>
</tr>
<tr>
<td>Lower Open Gustilo III</td>
<td>110 (28)</td>
<td>1,235 (59)</td>
<td>21.3</td>
</tr>
</tbody>
</table>

*p values are for comparisons made between distributions of closed, open Gustilo I/II, and open Gustilo III fractures within category.
for studying combat-related trauma. Both military and civilian centers are challenged to address segmental bone defects, vascular injuries with ischemia, severe soft tissue injuries, and wounds/fractures with severe contamination. As shown in the data presented in this study, while these characteristics are treated at MTFs at a higher rate compared to CIV CENs, the absolute number treated at MTFs pales in comparison to those treated at CIV CENs.

As the U.S. transitions into a low-volume casualty flow period, the military will rely more heavily on civilian collaborations to conduct clinically relevant combat casualty care research. While the MTFs accounted for 8.4% of the type III open tibia fractures included in the METRC registry, they only accounted for 4.5% of the type III open tibia fractures enrolled in three large prospective research studies that followed. Even more dramatic is the difference in the contribution of MTFs to the total number of traumatic amputations. MTFs contributed 72% of the 232 traumatic amputations included in the registry but only 18% of the 67 traumatic amputations enrolled in the three subsequent METRC studies. This difference is likely due to the fact that the registry data were collected in 2010 and 2011, a period of active combat in Iraq and Afghanistan. In contrast, over 80% of the patients enrolled in the three METRC studies included in this analysis were enrolled between 2013 and 2015, which was a period when the United States military

### TABLE III. Number of Registry Fractures (and Average per Site) by OTA Open Fracture Classification

<table>
<thead>
<tr>
<th>Category</th>
<th>Item</th>
<th>Total Number (Average Per Site)</th>
<th>Percent Distribution</th>
<th>p Value*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>MTFs</td>
<td>CIV CENs</td>
<td>MTFs</td>
</tr>
<tr>
<td>Bone Loss</td>
<td>None or Insignificant Bone Loss</td>
<td>121 (30)</td>
<td>1,663 (80)</td>
<td>49.4</td>
</tr>
<tr>
<td></td>
<td>Bone Missing But Some Contact</td>
<td>67 (17)</td>
<td>1,117 (54)</td>
<td>27.3</td>
</tr>
<tr>
<td></td>
<td>Segmental Bone Loss ≤2 cm</td>
<td>21 (5)</td>
<td>192 (9)</td>
<td>8.6</td>
</tr>
<tr>
<td></td>
<td>Segmental Bone Loss &gt;2 cm</td>
<td>36 (9)</td>
<td>200 (10)</td>
<td>14.7</td>
</tr>
<tr>
<td>Muscle</td>
<td>None or Minor</td>
<td>119 (30)</td>
<td>1,827 (88)</td>
<td>49.6</td>
</tr>
<tr>
<td></td>
<td>Loss of Significant Tissue</td>
<td>77 (19)</td>
<td>1,062 (51)</td>
<td>32.1</td>
</tr>
<tr>
<td></td>
<td>Extensive Muscle Necrosis</td>
<td>44 (11)</td>
<td>259 (12)</td>
<td>18.3</td>
</tr>
<tr>
<td></td>
<td>Unable to Assess</td>
<td>5 (1)</td>
<td>24 (1)</td>
<td>2.0</td>
</tr>
<tr>
<td>Skin</td>
<td>Can Be Approximated</td>
<td>159 (40)</td>
<td>2,602 (125)</td>
<td>64.9</td>
</tr>
<tr>
<td></td>
<td>Cannot Be Approximated</td>
<td>50 (13)</td>
<td>325 (16)</td>
<td>20.4</td>
</tr>
<tr>
<td></td>
<td>Extensive Degloving</td>
<td>36 (9)</td>
<td>245 (12)</td>
<td>14.7</td>
</tr>
<tr>
<td>Arterial</td>
<td>No Arterial Injury</td>
<td>200 (50)</td>
<td>2,844 (136)</td>
<td>81.6</td>
</tr>
<tr>
<td></td>
<td>Artery Injury w/o Ischemia</td>
<td>32 (8)</td>
<td>207 (10)</td>
<td>13.1</td>
</tr>
<tr>
<td></td>
<td>Artery Injury w/o Ischemia</td>
<td>13 (3)</td>
<td>121 (6)</td>
<td>5.3</td>
</tr>
<tr>
<td>Contamination</td>
<td>No Grossly Visible Contamination</td>
<td>91 (23)</td>
<td>1,830 (88)</td>
<td>37.1</td>
</tr>
<tr>
<td></td>
<td>Surface Contamination Visible</td>
<td>59 (15)</td>
<td>994 (48)</td>
<td>24.1</td>
</tr>
<tr>
<td></td>
<td>Contamination Embedded in Bone</td>
<td>95 (24)</td>
<td>348 (17)</td>
<td>38.8</td>
</tr>
</tbody>
</table>

*p values are for comparisons made between distributions of severity level within category.

### TABLE IV. Number of Fractures Enrolled in 3 METRC Studies (and Average per Site) by OTA Open Fracture Classification

<table>
<thead>
<tr>
<th>Category</th>
<th>Item</th>
<th>Total Number (Average Per Site)</th>
<th>Percent Distribution</th>
<th>p Value*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>MTFs</td>
<td>CIV CENs</td>
<td>MTFs</td>
</tr>
<tr>
<td>Bone Loss</td>
<td>None or Insignificant Bone Loss</td>
<td>13 (3)</td>
<td>321 (15)</td>
<td>23.2</td>
</tr>
<tr>
<td></td>
<td>Bone Missing But Some Contact</td>
<td>24 (6)</td>
<td>472 (21)</td>
<td>42.9</td>
</tr>
<tr>
<td></td>
<td>Segmental Bone Loss ≤2 cm</td>
<td>5 (1)</td>
<td>86 (4)</td>
<td>8.9</td>
</tr>
<tr>
<td></td>
<td>Segmental Bone Loss &gt;2 cm</td>
<td>14 (4)</td>
<td>198 (9)</td>
<td>25.0</td>
</tr>
<tr>
<td>Muscle</td>
<td>None or Minor</td>
<td>8 (2)</td>
<td>333 (15)</td>
<td>14.3</td>
</tr>
<tr>
<td></td>
<td>Loss of Significant Tissue</td>
<td>33 (8)</td>
<td>546 (25)</td>
<td>58.9</td>
</tr>
<tr>
<td></td>
<td>Extensive Muscle Necrosis</td>
<td>15 (4)</td>
<td>198 (9)</td>
<td>26.8</td>
</tr>
<tr>
<td>Skin</td>
<td>Can Be Approximated</td>
<td>24 (6)</td>
<td>475 (22)</td>
<td>42.9</td>
</tr>
<tr>
<td></td>
<td>Cannot Be Approximated</td>
<td>20 (5)</td>
<td>352 (16)</td>
<td>35.7</td>
</tr>
<tr>
<td></td>
<td>Extensive Degloving</td>
<td>12 (3)</td>
<td>244 (11)</td>
<td>21.4</td>
</tr>
<tr>
<td>Arterial</td>
<td>No Arterial Injury</td>
<td>37 (9)</td>
<td>819 (37)</td>
<td>66.1</td>
</tr>
<tr>
<td></td>
<td>Artery Injury w/o Ischemia</td>
<td>15 (4)</td>
<td>194 (9)</td>
<td>26.8</td>
</tr>
<tr>
<td></td>
<td>Artery Injury w/o Ischemia</td>
<td>4 (1)</td>
<td>64 (3)</td>
<td>7.1</td>
</tr>
<tr>
<td>Contamination</td>
<td>No Grossly Visible Contamination</td>
<td>12 (3)</td>
<td>284 (13)</td>
<td>21.4</td>
</tr>
<tr>
<td></td>
<td>Surface Contamination Visible</td>
<td>23 (6)</td>
<td>508 (23)</td>
<td>41.1</td>
</tr>
<tr>
<td></td>
<td>Contamination Embedded in Bone</td>
<td>21 (5)</td>
<td>284 (13)</td>
<td>37.5</td>
</tr>
</tbody>
</table>

*p values are for comparisons made between distributions of severity level within category.
was transitioning from a high-volume casualty flow to a low-volume casualty flow. In other words, the overall number of severe combat-related injuries had dropped dramatically and there were fewer severe extremity injuries presenting to MTFs.

There are several limitations to the current study. The mechanism of injury was not used to compare types of injuries, such as blast injuries which were the hallmark mechanism of injury during the conflicts in Iraq and Afghanistan. The authors chose to use injury characteristics instead of mechanism of injury as these are very specific injury patterns that allow for a more standardized and reliable comparison. In practice, treatment of extremity injuries is based on factors such as the bony injury, extent of soft tissue damage, and the level of contamination, which were all accounted for using the different fracture classification systems. We also recognize that although injured service members are cared for throughout the medical evacuation process back to the United States, there are limitations associated with comparing care of these patients to those cared for solely at a single center. Finally, while most MTFs only treat military healthcare beneficiaries, San Antonio Military Medical Center also serves as the only DoD Level I trauma center and treats a consistent flow of civilian trauma patients with these severe extremity injuries. As such, the MTF data may over represent the true military contribution. Furthermore, confounding variables such as age, social setting, and comorbidities may limit extrapolation of MTFs to CIV CENs and vice versa.

In conclusion, while occurring at a higher relative rate at MTFs, the absolute number of patients with severe extremity injuries treated and enrolled in core prospective METRC studies at CIV CENs is robust and in most cases far exceeds the number available at the four MTFs treating the majority of combat injuries. In times of low military casualty flow, it is likely that clinical trials will rely more on CIV CENs than MTFs for patient enrollment. These data demonstrate the benefits of maintaining a clinical research network that includes both civilian trauma centers and military treatment facilities. Only by combining the population of injured service members and combat expertise of military treatment facilities with the patients and clinical research expertise of civilian trauma centers will we be able to conduct the research necessary to establish best practices for treating both service members and civilians with severe extremity trauma. Equally important is the ability to take the lessons learned in treating service members during periods of active combat and translate them to the civilian sector where they can be further refined so that benefits of new technologies and improved treatment strategies can accrue to those injured in subsequent conflicts.

**ACKNOWLEDGMENTS**

Authors representing participating METRC centers are—Boston Medical Center (Boston, MA): Paul Tometta III, MD; Hope Carlisle, RN BSN; Heather Silva. Department of Orthopaedic Surgery, Carolinas Medical Center (Charlotte, NC): Madhav A. Karunakar, MD; Stephen Sims, MD; Rachel B. Seymour, PhD; Denver Health and Hospital Authority (Denver, CO): David J. Hak, MD; Corey E. Henderson, MS. Florida Orthopaedic Institute (Tampa, FL): Roy W. Sanders, MD; Henry Sagi, MD; Barbara Stevenson, MHA; Hennepin County Medical Center (Minneapolis, MN): Andrew H. Schmidt, MD; Jerald R. Westberg, BA. Indiana University Health Methodist Hospital (Indianapolis, IN): Todd O. McKinley, MD; Greg E. Gaski, MD; Laurence B. Kempton, MD. MetroHealth Medical Center (Cleveland, OH): Heather A. Valleri, MD; Mary A. Breslin, BA. Naval Medical Center Portsmouth (Portsmouth, VA): Christopher S. Smith, MD; Lorie M. Gower, CCRC. Center for Orthopaedic Research and Education (Phoenix, AZ): Clifford B. Jones, MD FACS; Debra L. Sietsema, PhD; RN. Penn State University M.S. Hershey Medical Center (Hershey, PA): J. Spence Reid, MD; Andrea H. Horne, CCRC. Department of Orthopaedic Surgery, San Antonio Military Medical Center (Fort Sam Houston, TX): Lt Col Patrick M. Osborn, USAF MC; Michelle Norton. St. Louis University Hospital, (St. Louis, MO): Lisa K. Cannada, MD. University of California at San Francisco (San Francisco, CA): Theodore Miclau, MD; Saam Morshed, MD; PhD; Tgist Belaye, MS; Jonathan K. Kwong. University of Maryland R Adams Cowley Shock Trauma Center (Baltimore, MD): Andrew N. Pollak, MD; Robert V. O’Toole, MD; Theodore Manson, MD; Marcus F. Sciandri, MD; Merryjessica Fuerst, BA. University of Miami Ryder Trauma Center (Miami, FL): Gregory A. Zych, DO. University of Mississippi Medical Center Jackson, MS): Patrick F. Bergin, MD; Mathew Graves, MD; Clay A. Spiteri, MD; Josie M. Hydrick, BS. University of Oklahoma Medical Center (Oklahoma City, OK): William J.J. Erli, MD; Zachary V. Roberts, MD; David C. Teague, MD; Kathy Carl, BA, CCRC; Janet Wells, RN, BS, CCRC. University of Washington, Harborview Medical Center (Seattle, WA): Reza Firoozabadi, MD,MA; Julie Agel, MA. UT Health:

---

**TABLE V.** Number of Gustilo–Anderson Type III Fractures in Registry and Enrolled in METRC Studies—Types III Open Tibia Fractures Only

<table>
<thead>
<tr>
<th>Injury Type</th>
<th>Number (and Percent) Entered in Registry</th>
<th>Number (and Percent) Enrolled in 3 METRC Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All Centers</td>
<td>MTFs (%)</td>
</tr>
<tr>
<td>All Type III Open Tibia Fractures</td>
<td>794</td>
<td>67 (8.4)</td>
</tr>
<tr>
<td>Gustilo Type IIIA</td>
<td>498</td>
<td>36 (7.2)</td>
</tr>
<tr>
<td>Gustilo Type IIIB</td>
<td>239</td>
<td>23 (9.6)</td>
</tr>
<tr>
<td>Gustilo Type IICC</td>
<td>57</td>
<td>8 (14.0)</td>
</tr>
<tr>
<td>AO/OTA Type A</td>
<td>184</td>
<td>17 (9.2)</td>
</tr>
<tr>
<td>AO/OTA Type B</td>
<td>177</td>
<td>13 (7.3)</td>
</tr>
<tr>
<td>AO/OTA Type C</td>
<td>433</td>
<td>37 (8.5)</td>
</tr>
</tbody>
</table>

1Authors representing participating METRC centers are—Boston Medical Center (Boston, MA): Paul Tometta III, MD; Hope Carlisle, RN BSN; Heather Silva. Department of Orthopaedic Surgery, Carolinas Medical Center (Charlotte, NC): Madhav A. Karunakar, MD; Stephen Sims, MD; Rachel B. Seymour, PhD; Denver Health and Hospital Authority (Denver, CO): David J. Hak, MD; Corey E. Henderson, MS. Florida Orthopaedic Institute (Tampa, FL): Roy W. Sanders, MD; Henry Sagi, MD; Barbara Stevenson, MHA; Hennepin County Medical Center (Minneapolis, MN): Andrew H. Schmidt, MD; Jerald R. Westberg, BA. Indiana University Health Methodist Hospital (Indianapolis, IN): Todd O. McKinley, MD; Greg E. Gaski, MD; Laurence B. Kempton, MD. MetroHealth Medical Center (Cleveland, OH): Heather A. Valleri, MD; Mary A. Breslin, BA. Naval Medical Center Portsmouth (Portsmouth, VA): Christopher S. Smith, MD; Lorie M. Gower, CCRC. Center for Orthopaedic Research and Education (Phoenix, AZ): Clifford B. Jones, MD FACS; Debra L. Sietsema, PhD; RN. Penn State University M.S. Hershey Medical Center (Hershey, PA): J. Spence Reid, MD; Andrea H. Horne, CCRC. Department of Orthopaedic Surgery, San Antonio Military Medical Center (Fort Sam Houston, TX): Lt Col Patrick M. Osborn, USAF MC; Michelle Norton. St. Louis University Hospital, (St. Louis, MO): Lisa K. Cannada, MD. University of California at San Francisco (San Francisco, CA): Theodore Miclau, MD; Saam Morshed, MD; PhD; Tgist Belaye, MS; Jonathan K. Kwong. University of Maryland R Adams Cowley Shock Trauma Center (Baltimore, MD): Andrew N. Pollak, MD; Robert V. O’Toole, MD; Theodore Manson, MD; Marcus F. Sciandri, MD; Merryjessica Fuerst, BA. University of Miami Ryder Trauma Center (Miami, FL): Gregory A. Zych, DO. University of Mississippi Medical Center Jackson, MS): Patrick F. Bergin, MD; Mathew Graves, MD; Clay A. Spiteri, MD; Josie M. Hydrick, BS. University of Oklahoma Medical Center (Oklahoma City, OK): William J.J. Erli, MD; Zachary V. Roberts, MD; David C. Teague, MD; Kathy Carl, BA, CCRC; Janet Wells, RN, BS, CCRC. University of Washington, Harborview Medical Center (Seattle, WA): Reza Firoozabadi, MD,MA; Julie Agel, MA. UT Health:
The University of Texas Health Sciences Center at Houston (Houston, TX): Joshua L. Gary, MD; KD Franco, BSN, CCRC; Matthew Galpin, CCRC.
Vanderbilt University Medical Center (Nashville, TN): William T. Obremskey, MD; Cesar S. Molina, MD. Wake Forest University Baptist Medical Center (Winston-Salem, NC): Eben A. Carroll, MD; J. Brett Goodman, MBA; Martha B. Holden, AAS, AA. Department of Orthopaedic Surgery, Walter Reed National Military Medical Center (Bethesda, MD): Xochitl Ceniceros, PhDc, MS, RN. Department of Health Policy and Management, Bloomberg School of Public Health Baltimore MD): Lauren E. Allen, MA; Renan C. Castillo, PhD; Katherine P. Frey, RN, MPH; Lisa Reider, PhD.

REFERENCES

An Orthopedic Performance Improvement Initiative at a Small Military Treatment Facility

Lt Col John Christopher McGee, USAF BSC†; Christopher G. Bise, PT, MS, DPT, OCS†; Joel M. Stevens, DC†

ABSTRACT Objectives: The scientific literature demonstrates a cost-benefit associated with early access to physical therapy. The purpose of this case study is to report the results of an orthopedic performance improvement initiative (OPII) at a small military treatment facility (approximately 4.5K active duty beneficiaries). The OPII was introduced to (1) limit variation by ensuring that all active duty orthopedic consults were audited ensuring proper timing of appropriate services and (2) establish contractual agreement for shared resources with the U.S. Naval Jacksonville Orthopedic Department. Methods: OPII was accomplished through consensus development and strong leadership support. All orthopedic referrals \( n = 260 \) were audited for 6 months. Appropriate recommendations were provided to either continue with orthopedic care or to redirect to a physical therapy first approach. Results: Orthopedic referrals decreased 27% with concomitant 32% increase in physical therapy referrals producing overall savings of $462K (33%). Meanwhile, surgical throughput improved 45%. Seventy percent of the savings were attributed to improved utilization translating to a relative value unit savings per beneficiary of $17.64 (23.7%). Conclusion: Our results support the value of a conservative physical therapy first approach for musculoskeletal conditions and shared mil-to-mil resourcing agreements. Success requires an iterative audit/feedback process balanced with stakeholder consensus.

BACKGROUND
As the pace of research has accelerated, there has been a proliferation of evidence and evidence-based practice strategies. Despite increased availability of evidence-based resources, clinicians and health care organizations have generally failed to assimilate evidence from these resources into routine patient care. This failure has perpetuated “status quo” clinical practice patterns and contributed to unwarranted clinical variation, substandard quality of care, and unsustainable health care costs.1

It is well understood that two of the primary drivers of clinical variation and the associated increase in health care utilization for management of musculoskeletal conditions are increased imaging rates and increased access to specialty care such as orthopedics.2–4 In response, physical therapists have been challenged from within their profession to provide an alternative to the current “status quo” by demonstrating their value through evidence-based, cost-effective management strategies.5 Current literature has begun to demonstrate a cost-benefit associated with conservative care, specifically early access to physical therapy (PT). This “proactive” approach has been proven superior in terms of cost and outcomes when compared with the current “reactive” referral model of care.6 Acknowledging this evidence, the Air Force Medical Operations Agency recently mandated through policy in March 2015 the implementation of direct access PT within the patient-centered medical home environment.7 Overall, there is a push within the national health care system to address the barriers that prevent the translation of high-quality care approaches. This has sparked a research agenda aimed at overcoming such obstacles and improving implementation of best practices such as early access to PT.8,9 Implementation is described as the bridge between the decision to adopt an evidence-based strategy and the sustained use of that strategy in everyday clinical practice.10 In that regard, implementation is the key to creating sustained change of evidence-based, cost-effective approaches in health care delivery. The Consolidated Framework for Implementation Research (CFIR).11 provides a useful tool as a metatheoretical model suited for multilevel systems interventions and identifies five domains as potential targets for successful and sustainable implementation strategies.8

(1) Outer setting (e.g., patient needs, peer pressure, external influences such as network providers and their reimbursement concerns);
(2) Inner setting (e.g., leadership, culture, networks and communications, goals and feedback);
(3) Characteristics of individuals (e.g., knowledge and beliefs, personal attributes);
(4) Characteristics of the clinical innovation (e.g., supporting evidence, complexity); and
(5) Implementation processes (e.g., coordination, planning, opinion leaders, and champions).8

PURPOSE
The purpose of this article is to highlight a performance improvement initiative in a small military treatment facility

---

†Department of Biology, U.S. Air Force Academy, 2355 Faculty Drive, Suite 2P389, USAF Academy, CO 80840.
†Department of Physical Therapy, University of Pittsburgh, 100 Technology Drive, Suite 210, Pittsburgh, PA 15219-3130.
This study was presented as an oral main plenary session at the 2015 Military Health System Research Symposium.
doi: 10.7205/MILMED-D-16-00141
modeled within the CFIR. Our team identified issues with the current referral model and designed a “Physical Therapy First” (PTF) solution to improve appropriate referrals and decrease inappropriate resource utilization.

SETTING
Moody Air Force Base, home of the 23rd Fighter Wing and the 93rd Air Ground Operations Wing, is located remotely in southern Georgia about 12 miles north of the Florida state border. Team Moody organizes, trains, and employs combat-ready A-10, HC-130, HH-60, pararescuemen and force protection assets. Total personnel consist of approximately 6,300 military service members and civilians including geographically separated units in Nevada, Florida, and Arizona. Combined, the two Wings are responsible for executing worldwide close air support, force protection, and personnel recovery operations in support of humanitarian interests, U.S. national security objectives, and overseas contingency operations.12 Located at Moody, the 23rd Medical Group is a small outpatient Medical Treatment Facility which houses Primary Care, Pediatrics, Women’s Health and other services such as Physical Therapy and Mental Health. During the course of this initiative, there were approximately 12 credentialed Primary Care providers which included physicians, physician assistants, nurse practitioners, and two physical therapists on staff that could refer a patient to specialty care such as orthopedics. Overall, this team was responsible for the care of approximately 11,000 beneficiaries to include active duty members, dependents, retirees, and others empaneled to that facility. This initiative however focused specifically on the orthopedic referral patterns for the 4,500 active duty members assigned to Moody and did not include any of the other aforementioned beneficiary categories.

Before March 2013, orthopedics was not a service provided within the 23rd Medical Group. To provide the needed orthopedic services, the 23rd Medical Group referred over 60 active duty orthopedic specialty consults a month to a private managed care network. From July to December 2012, this service was identified as the most common private network referral product (Fig. 1). This resulted in orthopedics being the single largest private network–purchased product in terms of cost with an associated annual expense over $3M for active duty members alone. Furthermore, purchased orthopedic care was responsible for a 250% linear increase in utilization and 74% cost increase in just 4 years. To combat these rising costs, a performance improvement initiative was approved by the 23rd Medical Group Chief of Staff (Chief Medical Director) in March 2013. This initiative used the facility PT Director as the gatekeeper for all musculoskeletal consults requested by the Primary Care staff.

ORTHOPEDIC PERFORMANCE IMPROVEMENT INITIATIVE
The goal of the initiative was to promote cost-effective care through appropriate referral patterns. This was achieved by ensuring that PT was the first line of care and conservative treatment was exhausted before orthopedic specialty referral. Improved treatment outcomes, lower recidivism rates, and greater overall cost savings have all been identified as benefits of early access to PT care.6,8,13–18 Consults were further screened to allow for specialty care referral to the U.S. Navy Jacksonville Orthopedic Department instead of the private managed care network. This local initiative encouraged collaboration between programs within a specific region, thus recapturing or maintaining the health care dollars spent within the Department of Defense.

![Figure 1](https://example.com/figure1.png)

**FIGURE 1.** 23rd Medical Group specialty referral utilization/all beneficiaries (July 2012 to December 2012).
Throughout the initiative, all active duty orthopedic consults were consolidated daily by the Resource Management Officer who was the primary link for coordinating patient care between the military treatment facility and the civilian network. This individual printed and delivered all orthopedic consult requests to the PT Director to initiate chart review and render disposition. Volume ranged from 5 to 13 consults per day and the alternate therapist provider was always available for patient care, to review charts, and to provide opinion on disposition as needed. The PT Director or alternate therapist reviewed the electronic medical record (EMR) for each referral to screen for acuity, red-flags, appropriateness of referral, and to determine whether PT care had ever been provided for the chief complaint. Consult dispositions were rendered on the same business day as received. At no time were patient needs in terms of safety compromised. Cases requiring urgent or emergent orthopedic care such as fractures with associated positive images, distinct motor function loss, tumor effect, and postsurgical complications were immediately returned to the Resource Management Officer for orthopedic specialty referral processing and scheduling as warranted. Cases not meeting these criteria and not having evidence of an associated PT treatment regimen were re-engaged with the consulting provider via EMR correspondence suggesting PTF. At that time, the Primary Care provider then had the option of rerouting the consult for PT as suggested or declining the recommendation for conservative care and continuing with the normal orthopedic referral process. Overall, conflict was prevented as Primary Care providers were educated to inform patients before placing a referral that they were consulting with their musculoskeletal team and a final disposition would be made for the most appropriate course of action for that patient’s care. As a result, patients had prior knowledge that their consult might be rerouted to PTF instead. Furthermore, when a disposition was made to refer a patient to U.S. Navy Jacksonville, the patient ultimately could request to be seen by a civilian network provider instead, particularly if the medical temporary duty posed a hardship or if they simply preferred to be seen locally for whatever reason such as time factors, previous relationship with local network staff, or personal choice. Therefore, at any point in the process, the patient retained the right to request an orthopedic consult, and be treated locally or seek care at U.S. Navy Jacksonville.

Overall, the combined Orthopedic Performance Improvement Initiative (OPII) was a systems-based process tailored to deal with a multifactorial set of clinical barriers and limitations. Furthermore, the OPII was designed within the context of the CFIR model as previously indicated.11

In terms of the outer setting, a major factor driving the implementation of the OPII was the need to reduce the associated costs and referral rates to the private managed care network. Network providers had no discretion in the referral process, so their main concerns were centered on expected volume reduction from referral loss. These concerns were eased by dialogue with the 23rd Medical Group Chief of Staff who communicated that despite a reduction in visits, referrals would likely be better distributed and more appropriate for surgical intervention. The Chief of Staff further communicated that an associated increase in higher charge interventions would more than likely offset the loss of revenue from routine consults and facilitate improved access to their offices for the remainder of their nonmilitary clientele.

In terms of the inner setting, the initiative was endorsed and facilitated by leadership to include the Medical Group Commander, the Chief of Staff, and the full executive staff. Once leadership support was obtained, we engaged the professional staff (physicians, physician assistants, nurses, and therapists) in a forum to help develop consensus on the proposed changes to the protocol. The professional staff was initially briefed with background and supporting evidence to promote buy-in during a routine monthly staff meeting. For continued buy-in, clinical interventions and pathways were reviewed every quarter and the professional staff received feedback on clinical metrics and issues as they arose. We also sought to capitalize on the already healthy relationship between the Primary Care and PT teams before the initiative launch. The pre-existing culture of trust and mutual sharing solidified the unit mission and better enabled Primary Care to work collaboratively within the context of the interprofessional team. However, the ability to continue autonomous practice and maintain final decision-making authority showed respect for their knowledge and beliefs.

In addition to consensus building, we used audit and feedback reporting to the Professional Staff to reinforce that their referral behaviors were being monitored. This perhaps motivated change in behavior given that routine noncompliance with recommended referral patterns would have to be reconciled with the Chief of Staff. This was particularly true if a trend persisted whereby physical therapist dispositions were overruled and the preponderance of those patients were ultimately referred back to therapy by the network orthopedic surgeons. This in fact happened on eight occasions (100%) earlier in the initiative when the physical therapist disposition was overruled. We did not see this undesired trend during the latter half of initiative indicating positive clinical behavior change might have occurred.

With any change in workflow, there is always a fear that the proposed changes will result in an increase in the burden
experienced by the provider. This was one of the significant barriers identified in our early discussions with the professional staff. We addressed this fear by simplifying the proposed design of the OPII. The result was the addition of a single step to the existing procedure, minimizing the disruption in workflow. The subsequent deployment of the OPII was also seamless for patients and allowed for prompt referral without an additional delay. PT time was protected by optimizing appointments on the providers schedule to allow between 30 minutes to an hour daily for the added chart review. Furthermore, as the demand for therapy was expected to increase as a result of a redistribution of care, some care was authorized to be shifted to the local private PT managed care network. This was deemed acceptable by the stakeholders given the lower costs associated with purchased PT care network. This was deemed acceptable by the stakeholders given the lower costs associated with purchased PT over orthopedic care. The results of these small changes not only improved workflow but also promoted additional discussion centered on the individual patient and improving their continuity of care.

The health and success of any new initiative requires an implementation champion to guide the effort.\textsuperscript{11} The implementation of the OPII was no different. In this case specifically, the elements of success included a Physical Therapist with 18 years of clinical experience, a Board Certification in Orthopedic Sports Physical Therapy, an MBA, and a PhD in rehabilitation science with a concentration in health care economics. His membership on the executive staff also gave him access to visionary senior staff that recognized the need for progressive, reasonable, evidence-supported projects which would benefit the mission. Finally, the timing was right for collaboration with the Orthopedic Department at U.S. Navy Jacksonville who had a mutual interest to extend their services as mentioned above. Moreover, their leadership perceived this agreement of services as the beginning of a process that could serve as a template for other medical specialties. The PT Director and the Medical Group Chief of Staff were the critical players who endorsed and tirelessly championed the initiative from birth to fruition. Though the presence of these individuals could be seen as serendipitous, many times the implementation champion(s) already exist in an organization and simply needs to be identified. Given the success of the OPII, the protocol reached by group consensus would be a natural starting point for those workgroups looking to reduce costs and streamline orthopedic care. Adoption of part or all of an existing protocol could lead to rapid translation and implementation in other environments.

RESULTS
All data were captured for a period of 6 months from March 2013 to September 2013. During that period, a total of 260 orthopedic consults were placed by the Primary Care team members and were subsequently audited by the PT Director. As a result of the performance improvement implementation, a reduction in monthly orthopedic referrals was observed from an average of 60 per month to 43 (27% less), thus demonstrating a positive behavioral change. Of the 260 total referrals, 109 (42%) were either deferred to PTF (34, 13%) or recommended for recaptured care (75, 29%) at U.S. Navy Jacksonville (Fig. 2). As previously indicated, all eight patients screened for PTF and denied that disposition by their primary care provider ended up being immediately referred back by their network orthopedic provider for the recommended PT treatment.

Of the 75 orthopedic patients recommended for recapture at U.S. Naval Jacksonville, 18 deferred care and opted to be seen by a local network civilian orthopedic provider. No trend was identified for those who deferred care, but approximately 22% (4) specified that the distance to Jacksonville was too far without associated travel pay; a trend that was eventually reversed when the travel radius was shortened by the 23rd Wing Commander which enabled travel pay for outpatient single-day medical appointments to Jacksonville. Ultimately, this policy change would only reduce the overall cost savings of the initiative by 10% when adjusting for mileage and per diem. Other reasons for patients denying care at U.S. Navy Jacksonville can be observed in Figure 3.

For the 57 remaining cases that were recaptured by U.S. Navy Jacksonville orthopedics, 77% (44) of those ended up with a surgical or procedural disposition. Despite 12 patients (21%) opting to defer a procedure, the combined surgical throughput during this 6-month period was still >55%. This was well above the normative 10% to 15% surgical throughput range that U. S. Navy Jacksonville was producing with the rest of their caseload before their relationship with the 23rd Medical Group. This supports our earlier hypothesis that these cases were more appropriate referrals. By comparison, a follow-up telephone conversation with an orthopedic team member from the San Antonio Military Medical Center verified that their typical surgical throughput averaged just over 10% as the largest combined orthopedic department within the DoD. Furthermore, this joint collaboration added another intangible benefit of ensuring timely EMR entry for improved continuity of care and more rapid medical and military operational decision-making.

Next, as care was redistributed to PTF as a result of the audit process and change in primary care referral patterns,
PT referrals increased from 114 to 150 (32% increase) per month which in turn culminated in a higher degree of network leakage (referral to nonmilitary treatment facility PT provider) to accommodate this difference. Generally, a network referral for PT is considered a much higher cost burden to the government than if care were to be maintained at the military treatment facility; however, this proved to be an acceptable cost in this particular case based on overall cost savings associated with the entire performance improvement initiative. In fact, total active duty outpatient musculoskeletal product (combined orthopedic and PT) costs were slightly higher (approximately $100K [$2.5M to $2.51M]) compared to the same time period in 2012; however, the total combined costs to include inpatient care resulted in a $231K savings overall for 2013 with the initiative in place for only 6 months. This partial figure projects to a $462K annual savings per year which equals 47% of the amount spent the previous year on this product line (Fig. 4). Furthermore, the associated savings far outpaced the standard 4% discount adjustment for inflation and does not include any savings


FIGURE 4. 23rd Medical Group combined active duty orthopedic (orthopedics, podiatry, PT) product costs.
that could be attained by expanding the initiative beyond the active duty population. Lastly, it is worthwhile to consider that the $462K in annual savings from the OPII could be redistributed to purchase up to three additional contract PT’s for the facility to recapture care and additional savings; particularly if the redistribution of PT care resulted in a higher percentage of the overall outpatient musculoskeletal product expenditure as expected. This would also include any further savings associated with care maintained at the military treatment facility.

At first, it might appear that the preponderance of savings was primarily related to the direct recapture of orthopedic cases by U.S. Navy Jacksonville. However, further investigation demonstrated that this mechanism was only responsible for approximately $68.4K ($136.8K annually) or roughly 29.6% of the total cost benefit. This means that the remaining $162.6K ($325.2 annually) or 70.4% of the savings was directly attributed to proper utilization and reduced variation through PTF coupled with primary care referral behavior change associated with the initiative’s implementation. Ultimately, this translated to a relative value unit (RVU) savings per beneficiary of $74.55 (in 2012) to $56.90 (in 2013) or $17.64 (23.7%) less for the military treatment facility based upon their fee schedule (Fig. 5).

Finally, despite not obtaining regional or global outcome functional metrics, there were no adverse patient survey reports collected during this period suggesting that the individual treatment outcomes were equivocal to those associated with usual care before implementation of the performance improvement initiative. Therefore, even if the assumed functional benefits were null, the cost-effectiveness of this initiative was clearly demonstrated in this case.

DISCUSSION

Our results show the success of the implementation model in practice in regard to changing orthopedic referral behaviors. Additionally, we were able to demonstrate the effectiveness of a conservative PTF approach from a business operations standpoint and patient-centered perspective. Our findings provide support for a broader translation of this approach throughout the Military Health System. Furthermore, this approach may be generalizable to the U.S. health care system at large.

When tackling processes that are multivariate in nature, positive change can be elusive.20,21 In today’s complex health care environment, clinical behavioral change is paramount, else we risk sacrificing our future by having our gross domestic product consumed by uncontrollable health care expenditures. For several years, we have known that variation in practice and access to specialty care are two significant factors that result in high utilization patterns.2–4,22,23 However, putting steps into practice to effectively combat those issues is often met with insurmountable resistance as early as the design phase. Moreover, one must accept that any project aimed at controlling costs and inefficiencies will surely be met with resistance. All stakeholders have perspectives which they consider “unique.” They are not often mutually shared, and can be perceived as threats to those with opposing viewpoints. The result is that a threat to one group’s perceived bottom line will undoubtedly spark confrontation, thereby disrupting any corrective process from taking root. Clearly, understanding these perspectives upfront is vital to effectively deal with the barriers and obstacles that threaten to derail the improvement process.20 Guided by the CFIR,11 our team not only demonstrated success with implementing a controversial performance improvement project, but also realized a

---

**FIGURE 5.** 23rd Medical Group combined orthopedic product cost/RVU.
tremendous value for all stakeholders involved without ever being driven too far off course by barriers and obstacles. In this regard, we also demonstrated that the outcome of delivering the right care by the right provider at the right time is a valid proxy for eliminating variation and therefore “bridging the evidence to practice gap.”

Before we launched the (OPII), we were well aware of the pitfalls that could disrupt its longevity. For example, we expected discontent from our local network orthopedic specialists and administrators upon learning that this project would threaten to reduce their clinical volume. We observed first-hand during the ’90s how practitioners naturally countered payment cuts through managed care policies by ramping up procedure volume.4,24 However, we also understood the importance of not demonizing or alienating a particular group of partners who by virtue of commitment to their profession should be respectfully viewed as altruistic caregivers that do in fact provide a valuable service to our population. We further understood the importance of respecting the clinical autonomy of our Primary Care teams so they would not feel as though they were simply cogs in controlling the supply chain. Lastly, we took care to avoid making our patients feel like production widgets by first prioritizing their safety and then by protecting their personal choices in the delivery of their health care.

Beyond creating a respectful culture necessary for cooperation, we had to further emphasize and prove the benefits to each stakeholder group.25 For example, improved surgical or procedural throughput was welcomed by both our network and military orthopedic specialists. Transparency and constant reporting enabled our primary care staff to observe the benefits associated with following the evidence-based guidelines of a PTF approach. We did not directly obtain specific functional outcome metrics, Primary Care follow-up requirements per episode of care, nor did we obtain patient satisfaction surveys to measure clinical outcomes effectiveness. However, some insight to patient satisfaction can be inferred from a lack of adverse reporting. Additionally, savings resulting from the initiative could also be explained by the point that perhaps repeat downstream follow-up medical requirements were at least mitigated to some degree as well which could reflect some level of patient satisfaction. Nonetheless, outcomes tracking, patient surveys, and survival curve analysis of the requirement for downstream medical care should be the target of future studies along with examination of access to care to further validate overall cost-effectiveness of the OPII. The latter would potentially translate to increased efficiency allowing providers to rapidly attend to those in need of urgent care. By establishing a respectful culture, working toward stakeholder consensus and emphasizing a team approach during the care continuum, we were able to demonstrate a commitment to our patients throughout the initiative. Lastly, the magnitude of the cost savings associated with the initiative drew significant attention from our resource management and executive teams which made it difficult to contest once the initiative was operational. Now that the OPII has been realized, and cooperation has been achieved, all of the aforementioned suggestions for continued process improvement can be easily implemented and tracked within the context of the CFIR.11 Examination of the OPII over time would further validate the CFIR model by ensuring continuity and clinical behavioral sustainability.

Finally, we believe we demonstrated the value of a conservative PTF approach when treating patients with musculoskeletal conditions and dysfunctions. Not only did our results support the growing body of evidence favoring this approach in the literature,6,8,13–18 it also provides evidence for the physical therapist as the portal of entry into the health care system when the primary complaint is musculoskeletal in nature. A preponderance of states, including the military, already have Direct Access legislation in place that permit therapists to see patients without referral7,26; however, this is not the typical practice mode for various reasons and it is not our intention to discuss here. Our results give further credence not only to unrestricted direct access but to other models of practice that embed therapists in primary care teams to screen, consult, advise, and to engage patients in earlier treatment, further optimizing clinical outcomes.26 Our situation as a small military treatment facility enabled this initiative to work well. However, given our precedence of success, the PTF approach of the Orthopedic initiative could be easily translated anywhere within the context of the CFIR.11 For example, in scenarios where this clinical model is not available, it would appear that a cost-effective alternative such as the PTF be initiated. At least this expedites the pathway for a patient to be seen earlier in PT, thereby ensuring a safety screen, maximizing appropriateness of care, improving outcomes, and minimizing the probability of reoccurrence, while reducing unwanted clinical variation at a considerable yet undeniable cost savings. Lastly, from a DoD or interagency standpoint, our findings also emphasize the benefit from a shared resourcing perspective.

REFERENCES
7. Air Force Surgeon General Public Affairs: Active-duty Airmen to have direct access to physical therapy clinics; May 5, 2015. Available at

24 MILITARY MEDICINE, Vol. 182, March/April Supplement 2017


Reliability, Responsiveness, and Criterion Validity of the Kiio Sensor, a New Tool for Assessment of Muscle Function

Patrick Grabowski, MPT, PhD; Michelle Narveson, DPT; Stephen Siegle, DPT

ABSTRACT Musculoskeletal injuries are a leading cause of disability in both the general population and military, and the ability to effectively quantify musculoskeletal function remains problematic. The aim of the current study is to investigate the reliability, responsiveness, and criterion validity of the kiio Sensor for measurement of muscle strength, a key aspect of comprehensive functional assessment. Forty-four (24 male, 20 female) physically active civilian adults (mean [SD] = 21.2 [1.5] years of age) with no history of upper extremity injuries in the last year and no current complaints of pain, weakness, or functional limitation completed two sessions of maximum shoulder external rotation contractions. The forces were measured with a kiio Sensor and isokinetic dynamometer. The devices showed strong correlation ($r = 0.89$) and no significant mean difference ($0.3 \pm 3.3$ lb, $p = 0.47$). Intrasession reliability for both devices was analyzed using intraclass correlation coefficients (ICC$_{3,1} = 0.96$ [95% confidence interval = 0.93–0.98]), with the kiio Sensor having slightly less standard error and trial-to-trial variability. The kiio Sensor 7-day reliability was ICC$_{3,3} = 0.97$ (95% confidence interval = 0.94–0.98). The kiio Sensor demonstrates excellent reliability, responsiveness, and validity compared with a gold standard isokinetic dynamometer. Several key attributes contributing to this technology’s military relevance are discussed.

INTRODUCTION Musculoskeletal injuries are a leading cause of disability in the general population and U.S. military, resulting in enormous direct and indirect costs.$^{1-4}$ Prerequisite to any program of injury management, from prevention to rehabilitation, is the ability to effectively quantify the many facets of musculoskeletal function. The World Health Organization defines function at three different levels: the body part function and structure, the whole person, and the whole person in a social context.$^5$ Disability results from dysfunction at one or more of these levels in the form of impairments, activity limitations, and participation restrictions, respectively. To effectively assess any individual, all three levels must be tested.$^6$ Although tests of the activity and participation levels are important for estimating daily function, they lack sensitivity to underlying impairments of muscle function allowing potential compensatory movements, which may predispose to future injury. For example, strength impairment of the ankle plantar flexors is a risk factor for development of Achilles tendinopathy.$^7,8$ However, plantar flexor strength has very low correlation with tests of activity and participation.$^9$ Thus, assessment at the level of body part must also be considered. Additionally, activity and participation tests provide only limited information to guide the progress of individuals who fail to meet required standards. For instance, an individual who fails to achieve the required proficiency in a push-up test would benefit from knowing if this was related to weakness in the pectoralis musculature, the dynamic stabilizing muscles of the glenohumeral joint, or the scapulothoracic stabilizers.$^{10}$ This is especially critical after injury, as medical and rehabilitative treatment primarily aim to resolve impairments in body part function and structure.$^7,11$ Therefore, to comprehensively guide decisions for musculoskeletal training either to prevent or rehabilitate injury, specific objective measures at the level of key individual body parts is a necessity, and should be used in conjunction with other functional tests. Objective performance standards serve a number of important purposes, from identifying and reducing injury risk factors, to easing clinical decision-making with legally defensible criteria, to improving efficiency of fitness training and rehabilitation. To develop and administer specific, comprehensive, and validated performance standards, effective methods for measurement of muscle function must be identified.

Despite recent technology advancement,$^{12}$ measurement of regional musculoskeletal function remains a complex entity rooted in long-standing paradigms. The most popular methods, which have been in use by clinicians and fitness professionals for decades, all have positive aspects but significant limitations. Manual muscle testing (MMT) is the most widely used method. An examiner applies hand pressure to a body segment as the individual exerts countereffort. The examiner then subjectively rates the maximal exertion of the tested muscle on a 0 to 5 Likert-type scale. This technique is efficient and cost-effective, but unreliable and negatively impacted by human factors such as size and strength of the examiner.$^{13-15}$ Additionally, this only provides a single metric of maximal isometric strength. Recent studies show that more complex metrics requiring a force/time curve, such as rate of force development, provide useful information for return to activity decisions.$^{16}$ For these reasons, MMT fails to provide adequate data needed to develop performance standards.$^{17}$ On the other end of the spectrum are isokinetic dynamometers...
(ID), which provide highly complex and dynamic measures, such as indices of power and work, with excellent reliability and are generally considered the gold standard for muscle assessment.\textsuperscript{18,19} The major limitations for IDs include high cost (up to $50,000 per instrument), nonportability with a large footprint (up to 64 ft\textsuperscript{2}), and complex setup and operation.\textsuperscript{20,21} The ID instruments are computerized, recording a number of useful metrics based on a torque/time curve, but their high cost and labor-intensive methodology have limited their applicability on a large scale. The most commonly used compromise between MMT and ID devices is handheld dynamometry (HHD). The HHD adds a quantitative force sensor, thus adding to the equipment necessary to conduct tests.\textsuperscript{27} A solution to this is to provide external stabilization to the device, which can be cumbersome, time consuming, and adds to the equipment necessary to conduct tests.\textsuperscript{28} Some have wireless technology to expedite data collection, and some have the ability to generate a force/time curve for more complex performance metrics. The major limitation for these devices is that numerous authors find them subject to the same human factors as MMT, such that if the muscle to be tested is stronger than the examiner, the data are no longer reliable.\textsuperscript{23–26} A solution to this is to provide external stabilization to the device, which can be cumbersome, time consuming, and adds to the equipment necessary to conduct tests.\textsuperscript{27} Recently, tension myometers have become more appealing as strength testing devices (sometimes referred to as pull-type HHD).\textsuperscript{28,29} These strain gauges eliminate the examiner from the testing procedure, resulting in reliable strength data. Unfortunately, very little data exist on the measurement properties of myometers. A new device, the kiio Sensor (Kiio, Madison, Wisconsin), is a tension myometer that has been validated and shown to be reliable in mechanical laboratory testing, but has not been studied with human subjects (personal communication with Kiio). Therefore, the purpose of this study is to investigate the reliability, responsiveness, and criterion validity of the kiio Sensor in comparison to a gold standard ID.

**METHODS**

**Participants**

Forty-four (24 male, 20 female) civilian adults (age, mean [SD] = 21.2 [1.5] years) with no history of upper extremity injury in the last year and no current complaints of pain, weakness, or functional limitation participated. Males averaged 179 cm tall (range 168–199 cm), with mean weight of 84.1 kg (range 64.4–161.0 kg). For females, mean height was 166 cm (range 159–176 cm), mean weight was 68.5 kg (range 53.0–88.2 kg). Thirty-nine were right-hand dominant based on self-report of hand used for writing. Individuals were recruited from the local campus physical therapy and exercise science departments. All participants reported moderate-to-high levels of average weekly physical activity (≥5 on the Tegner Activity Level Scale\textsuperscript{30}). Individuals were excluded if they had any history of cardiovascular disease, neurologic or vascular impairment of the upper extremity, or cervical injury in the previous year. All participants provided informed consent before the study, which was approved by the University of Wisconsin-La Crosse Institutional Review Board.

**Procedure**

Individuals participated in two separate sessions 1 week apart, at the same time of day to maximize consistency of recent activity levels. Both sessions began with a 3-minute warm-up on an upper extremity ergometer, followed by submaximal isometric external rotation (ER) contractions at approximately 25, 50, and 75% of full effort. During the first session, subjects performed a familiarization trial with each device, after which they performed three maximal isometric contractions of standing shoulder ER using both the kiio Sensor and a Cybex Norm (CSMI, Stoughton, Massachusetts) ID, for a total of six repetitions, all performed with the dominant arm. The instrument sequence was randomized and balanced among participants to eliminate order effects. Participants received standardized instructions to perform at maximal effort, but no verbal encouragement during exertion. They were instructed to inform the investigator of any pain experienced during testing so the procedure could be terminated appropriately. All trials with a single instrument were separated by 1 minute, and participants rested 5 minutes between instruments to minimize the effects of fatigue.\textsuperscript{31} Individuals were asked to maintain their current level of physical activity, with no new activities or exercise protocols commencing between test sessions. When participants returned for session two, they performed three maximal isometric contractions of standing shoulder ER with the kiio Sensor only. For both sessions, all contractions were performed with the shoulder in a position of 30° abduction in the scapular plane (30° anterior to the
frontal plane), and 0° of rotation (Fig. 1A). The standing position previously provided reliable shoulder measures.32,33 This position was maintained during use of the kiio Sensor by keeping the participant standing at the dynamometer station, using the elbow rest as a light tactile guide only, the elbow was not strapped in place as it normally would be with use of the dynamometer (Fig. 1B). The ID was calibrated on a weekly basis. The calibration of the kiio Sensor was checked weekly using the same set of weights as the ID, and it maintained accuracy within ±0.1 lb throughout the data collection period without needing recalibration.

**Statistical Analysis**

All data from the ID were collected with Humac2009 software (CSMI) at a sampling rate of 100 Hz, and were entered manually into a spreadsheet. All data from the kiio Sensor were sampled at 80 Hz, transmitted automatically to kiio Flex software (Kiio) running on a Windows tablet (Dell Latitude 10; Dell Inc., Round Rock, Texas), and then exported to the spreadsheet for further analysis. For ID data, torques recorded by the Cybex system were converted to force by measuring the mechanical lever arm from the dynamometer axis of rotation to the handle, and dividing the torque value by this distance. Reliability coefficients (intraclass correlation coefficient [ICC3,1] with 95% confidence intervals) and the standard error of measure were calculated for both the kiio Sensor and ID for the three trials within session one.34 The method error coefficient of variation (CV_ME) was also calculated for each device, providing a measure of the percentage of variation relative to the mean from trial to trial.35 Reliability and CV_ME were also analyzed for the kiio Sensor over the 7-day period between tests, using the pooled data (average of 3 trials) from sessions one and two. Additionally, the minimum detectable difference (MDD95) was calculated based on the 7-day data.35 Criterion validity of the kiio Sensor relative to the ID was analyzed using a Pearson correlation coefficient to evaluate the strength of the linear relationship of the data and a matched pairs t test to evaluate for systematic bias.35 Bland–Altman plots with 95% limits of agreement were constructed to visually examine for systematic discrepancies across the range of forces measured.36 All statistical analyses were completed with SPSS v22 (IBM, Armonk, New York).

**RESULTS**

Mean (SD) maximal ER force for session 1 was 20.4 lb (6.5). Reliability outcomes are presented in Table I. The Pearson correlation coefficient was r = 0.89, and the matched-pairs t test showed no difference between the two devices (mean difference = 0.3 ± 3.3 lb, p = 0.47). Limits of agreement were ±6.6 lb (Fig. 2). Mean (SD) maximal ER force for session 2 was 21.6 lb (6.0), with reliability and responsiveness outcomes in Table I.

**DISCUSSION**

This is the first study to investigate the reliability, responsiveness, and criterion validity of the kiio Sensor for strength measurement in humans. The maximum forces measured, and the high reliability of the gold standard ID in our protocol is in agreement with previous research on shoulder ER isometric strength.32,37,38 Further analyses indicate that the kiio Sensor provides clinically acceptable data with excellent reliability and responsiveness for human strength assessment that is as good as or better than the ID.35 This is consistent with previous research with another tension myometer (Mecmesin, West Sussex, United Kingdom), which was used to assess isometric shoulder strength as part of the Constant score, and demonstrated good repeatability with isometric ankle dorsi/plantar flexion strength measures as well.28,39,40 Bland–Altman analysis of kiio Sensor and ID data shows no evidence of systematic bias over the range of forces measured (Fig. 2). However, the limits of agreement of 6.6 lb are fairly high relative to the maximal forces measured for shoulder ER, and therefore repeated shoulder ER measures on any individual should be performed using the same type of device, rather than using them.

**TABLE I.** Reliability and Responsiveness Outcomes

<table>
<thead>
<tr>
<th></th>
<th>Within Session</th>
<th>7-Day Reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ICC3,1 (95% CI)</td>
<td>SEM (lb)</td>
</tr>
<tr>
<td>ID</td>
<td>0.96 (0.93–0.98)</td>
<td>1.3</td>
</tr>
<tr>
<td>kiio Sensor</td>
<td>0.96 (0.93–0.98)</td>
<td>1.2</td>
</tr>
</tbody>
</table>
Interchangeably over time with a single subject. Given the excellent reliability and small percentage of variation from trial-to-trial, low MDD, and strong criterion validity of the kiio Sensor, this device appears to provide a number of potential advantages over currently used methods, warranting further discussion.

The kiio Sensor provides efficient, objective quantification of muscle force/time, eliminating the problems of subjectivity, rater bias, and single data points encountered with the MMT and most HHD methods. The ID and other tension myometers also solve some of these problems, but the kiio Sensor does so at a fraction of the cost. Current list prices of approximately $1,000 are roughly 2% of an ID, and less than 50% of the Mecmesin myometer. The kiio Sensor is also very portable, being smaller and lighter than other tension myometers on the market, while the ID fails to be portable in any way. Thus, widespread use may be more easily achievable in various settings, from rehabilitation clinics, to fitness training facilities and beyond. The attachment of the kiio Sensor is adaptable to various implements (e.g., static straps, steel cable, elastic bands), and anchor points, allowing for wide-ranging use in both static and potentially dynamic test applications. Although the ID is the only device capable of standardizing joint rotation velocities, the ability of the kiio Sensor to provide kinetic measures throughout a range of motion with force/time data enables the possibility of collecting more complex metrics of body part function. This is an area of study with ample room for innovation, as the clinical use of rate of force production and joint acceleration and deceleration are only beginning to be investigated. Additionally, there may be ways to incorporate the device into the other levels of functional assessment, gaining more quantitative data from activity and participation tests. Perhaps the most important attribute of the kiio Sensor is its wireless connectivity. This allows for data to be quickly and seamlessly transmitted to a tablet, which worked reliably throughout this study. According to the manufacturer, this device can communicate with both computers and mobile devices, and integrates with a secure cloud database. This capability enables collection of mass quantities of information in a cost-effective manner, enabling large-scale analytics to identify trends having significant implications for prediction and prevention of injury or reinjury. In the military environment, this type of device represents a means to gather muscle performance data on large numbers of active duty members efficiently with limited expense. Data can be collected in various settings, without the need for extensive protocol training, specialized equipment, or dedicated testing locations. Consider the implications of attempting this with isokinetic technology, including the time, equipment, and space requirements, and the need for new assessment tools at the level of body part structure and function is clear. With tension myometers, it is feasible to assess multiple muscles on an individual in a matter of minutes, with automated generation of databases contributing to musculoskeletal injury prevention and rehabilitation initiatives.

Despite these advantages, the kiio Sensor has two limitations. The first is sampling rate that is limited to 80 Hz. Although this is sufficient for most clinical analysis, it is significantly slower than other devices. New Biodex IDs (Biodex Medical Systems, Inc., Shirley, New York), with appropriate auxiliary outputs, are capable of sampling rates up to 2,000 Hz, and the Mecmesin device can achieve 5,000 Hz. The significance of this may be minor, however, as some studies have shown that a sampling rate as low as 50 Hz results in negligible differences in force data when compared to frequencies up to 1,000 Hz. The second limitation is maximum load. The kiio Sensor is rated up to 250 lb, whereas IDs, such as the Cybex device in this study, are commonly tested for accuracy up to 500 lb of force. The Mecmesin myometer is capable of measuring up to 225 lb, whereas other tension myometers have various ranges. This range may limit the kiio Sensor’s ability to assess some muscles on some individuals, but it is possible to connect multiple sensors in parallel, increasing the capacity to at least 500 lb. Therefore, this limitation is likely to be rarely encountered. Considering its overall capabilities, the addition of some key tests of body part function using technologically advanced tools such as the kiio Sensor could provide a significant increase in usable data with which to manage musculoskeletal injury prevention and rehabilitation. Because of the efficiency and cost-effectiveness of such instruments, the relative effort of incorporating these measures is small in comparison to the scope and financial impact of musculoskeletal injuries.

LIMITATIONS

A methodological limitation likely added some small element of error to the data. In order to convert torques recorded with the ID to forces for comparison with the kiio Sensor, the moment arm of the ID needed to be measured manually. The investigators made every effort to keep this measurement as consistent as possible, thereby minimizing impact on the data. Of greater importance, this study was restricted to healthy individuals for shoulder ER only, which limits generalizability to patient populations and other body regions. Additionally, the individuals were civilians, which raises the question of whether the results apply to military populations. Although the population studied may have characteristics in common with some segments of active duty members, this study only represents a first step in the process. Now that reliability and criterion validity of this device in asymptomatic individuals are established, the effort and cost of future study are warranted. This includes the investigation of various body regions, various populations (i.e., military, healthy, injured, wider age ranges), and the usefulness of more complex data available from the kiio Sensor (e.g., rate of force production, area under the force/time curve) for both static and dynamic tests. Once accomplished, meaningful objective performance standards can then be established and prospectively validated.
CONCLUSIONS
Current methods of muscle assessment are limited in their ability to objectively quantify impairments of localized muscle function, rendering them impractical for widespread augmentation of comprehensive functional testing. The Kiio Sensor demonstrates excellent reliability, responsiveness, and validity compared with a gold standard ID in a group of healthy participants. Because of several key attributes, this technology may be an excellent tool for muscle assessment in widespread settings, and with additional study could assist in the establishment, validation, and administration of objective performance standards. Although this study is limited to the measurement of isometric shoulder ER, the device appears readily applicable to the measurement of more complex metrics for many muscle groups, efficiently increasing the quantity of useful data for the prevention and rehabilitation of musculoskeletal injuries.

ACKNOWLEDGMENTS
Equipment and software were loaned free of charge by Kiio, Madison, Wisconsin.

REFERENCES
The Afghan Theater: A Review of Military Medical Doctrine From 2008 to 2014

Col Ian Lane, L/RADC†; CAPT Zsolt Stockinger, MC USN‡; COL Samual Sauer, MC USA†; COL Mark Ervin, USAF MC‡; COL Michael Wirt, MC USA§; Surg Capt Stephen Bree, RN∥; COL Kirby Gross, MC USA¶; Col Jeffrey Bailey, USAF MC**; Brig Timothy Hodgetts, L/RAMC††; COL Elizabeth Mann-Salinas, AN USA§

ABSTRACT  This article forms part of a series that will explore the effect that Role 2 (R2) medical treatment facilities (MTFs) had on casualty care during the military campaign in Afghanistan and how we should interpret this to inform the capabilities in, and training for future R2 MTFs. Key aspects of doctrine which influence the effectiveness of R2 MTFs include timelines to care, patient movement capabilities, and MTF capabilities. The focus of this analysis was to review allied doctrine from the United States, United Kingdom, and the North Atlantic Treaty Organization to identify similarities and differences regarding employment of R2 related medical assets in the Afghan Theater, specifically for trauma care. Several discrepancies in medical doctrine persist among allied forces. Timelines to definitive care vary among nations. Allied nations should have clear taxonomy that clearly defines MTF capabilities within the combat casualty care system. The R2 surgical capability discrepancy between United States and North Atlantic Treaty Organization doctrine should be reconciled. Medical evacuation capabilities on the battlefield would be improved with a taxonomy that reflected the level of capability. Such changes may improve interoperability in a dynamic military landscape.

INTRODUCTION
This article forms part of a series that will explore the effect that Role 2 (R2) Medical Treatment Facilities (MTFs) had on casualty care during the military campaign in Afghanistan and how we should interpret this to inform the capabilities in, and training for future R2 MTFs. Roles of care refer to the increasing medical capabilities available for the combat injured. Generally, at the point of injury, combat life savers (soldiers trained to perform basic first aid) and trained combat medics apply Tactical Combat Casualty interventions to stabilize casualties and prepare for evacuation. Role 1 (R1) represents unit-level care at a field medical station, where a Licensed Independent Provider can provide advanced airway management and possibly initiate fresh whole blood transfusion in preparation for evacuation to surgical support. R2 provides more robust medical resources than R1, and is the first level of care where damage control surgery (DCS) and advanced resuscitation may be provided, but offers limited patient holding ability. R2 is the first MTF in the chain of evacuation, also referred to as Deployed Hospital Care. Role 3 (R3) is a deployed field hospital offering expanded surgical and imaging capabilities; patient holding duration is technically unlimited. Role 4 (R4) is a fixed facility that is in the home country of the deployed force or that of ally and offers all medical and surgical specialties from acute care to long-term rehabilitation.

The focus of this analysis is to review doctrine to identify similarities and differences regarding employment of R2 related medical assets in the Afghan Theater specifically for trauma care. The scope is limited to North Atlantic Treaty Organization (NATO) and United States (US)/United Kingdom (UK) joint and single Service doctrine, as the overwhelming majority of MTFs and evacuation assets in Afghanistan were from these two nations. An R2 Registry was implemented by the US Joint Trauma System in 2008 so doctrine from 2008 to 2014 was reviewed.

What Is Military Doctrine and How Is It Organized?
Military doctrine is the expression of how the military operates, linking theory, history, experimentation, and practice. Its objective is to describe how to think, not what to think. Yet, despite its centrality to military thinking, doctrine has been described as ill-defined, confusing, and poorly understood. NATO’s definition of doctrine, used unaltered
by many member nations including the US Department of Defense, is:

“Fundamental principles by which military forces guide their actions in support of objectives. It is authoritative, but requires judgment in application.”

It goes on to say that

“policy, as agreed by the highest National Authorities, normally leads and directs doctrine, and that applied doctrine is necessary for effective coalition building.”

The UK follows this line stating that “Except where there is a specific need for national doctrine, the UK will adopt and employ NATO doctrine.”

Military doctrine has been variously categorized but contemporary taxonomies tend not to align doctrine to a particular level of conflict or environment. The UK advocates four broad levels: philosophy, principles, practices, and procedures. Describing the relationships between these levels as:

“Philosophy is conceptual, enduring, pervasive and largely descriptive. It provides understanding. Principles, which are more specific, build upon the philosophical foundations to summarize that understanding. Both are likely to provide clearer context than faster-moving doctrine can, provided they are malleable. Practices describe the ways in which activity is conducted. Procedures link practices together. Both are intended to be prescriptive. Lower-level doctrine could change relatively rapidly and pragmatically, often from a bottom-up direction. However, practices and procedures should always be consistent with the higher-level philosophy and principles, which change only as a result of measured consideration, which is usually a top-down process.”

Using this template, Table I shows how the doctrine examined within this article fits into the organizational hierarchy.

Key aspects of operational factors which influence the effectiveness of R2 MTFs:

- Timelines to care
- Patient movement capabilities
- MTF capabilities

**TIMELINES**

Once injured, the principal factor that determines mortality, morbidity, and residual disability is time required to provide medical care. This is true of all medical emergencies, but is of overriding significance when dealing with surgical emergencies—particularly surgical control of hemorrhage in the combat setting. Hence, evacuation time is the major clinical driver dictating the type and location of medical assets in operations and conflicts, and timeliness in providing appropriate intervention to the wounded or ill is crucial. The provision of high-quality early intervention has been shown to improve outcomes, while any delay either before care is initiated or between subsequent levels of care, will be deleterious to patient outcomes.

**Allied Doctrine, Allied Joint Publication (AJP) 4-10 (A)** was the extant NATO doctrine for the whole of the

<table>
<thead>
<tr>
<th>TABLE I. Levels of Doctrine</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Level</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Philosophy</td>
</tr>
<tr>
<td>Principles</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Practices (Includes Joint and Allied Environmental, Functional and Thematic Doctrine)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

ADP, Army Doctrine Publication; AFM, Army Field Manual; AJP, Allied Joint Publication; AMedP, Allied Medical Publication; AMS, Army Medical Services; ANNEX, United States Air Force Doctrine; ATP, Army Technical Publications; BATLS, Battlefield Advanced Trauma Life Support; BDD, British Defence Doctrine; CGO, Clinical Guidelines for Operations; CPG, Clinical Practice Guidelines; FM, Field Manual; JDP, Joint Doctrine Publication; JP, Joint Publication; JTS, Joint Trauma System; JWP, Joint Warfare Publication; MCRP, Marine Corps Reference Publication; MIMMS, Major Incident Medical Management Support; NWP, Navy Warfare Publication; NTTP, Navy Tactics, Techniques, and Procedures.
period. Its principal medical planning timeline was the 1-2-4 hour principle:

“Primary (definitive) Surgery” for critically injured patients within one hour of wounding. If this is not achievable then Damage Control Surgery (DCS) should be available within two hours followed by primary surgery within four hours.”

Subordinate NATO publications have not provided further guidance. Allied joint medical planning doctrine (AJMedP-1) recommends the planning of Medical support (MEDEVAC) and did not specify any timelines by which timelines when discussing forward medical evacuation (Priority 3 within 72 hours) but only referred to clinical priorities for patient evacuation (Priority 1 requiring immediate transfer, Priority 2 within 24 hours, and Priority 3 within 72 hours) but only referred to clinical timelines when discussing forward medical evacuation (MEDEVAC) and did not specify any timelines by which casualties should reach a level of care. Standardization agreement (STANAG) contradicts AJMedP-2, prescribing a 2-hour evacuation time limit for urgent cases and 4 hours for priority cases.

UK Joint Doctrine initially (up to March 2011) was based on Joint Warfare Publication 4-03 which prescribed the 1:2:4 hour principle, albeit subtly different to the NATO description:

“rapid access to first aid and BATLS (battlefield advanced trauma life support)/BARTS (battlefield advanced resuscitation techniques and skills) resuscitation within one hour of wounding; access to surgical resuscitation (e.g., DCS) for those who require it within two hours of wounding; and primary surgery within four hours of wounding.”

It also recognized that when required by the unique operational environment the principle could be adapted accordingly. This was superseded by Joint Doctrine Publication (JDP) 4-03 which advocated a new clinical paradigm: 10(min)-1-2 (according to the Allied Command Operations Directive on Medical Support to Operations as cited by JDP) where:

“bleeding and airway control for the most severe casualties should be achieved as soon as possible—ideally within 10 minutes of wounding. MEDEVAC assets should reach the seriously wounding with skilled medical aid within one hour of wounding at the latest. Casualties that require surgery or further resuscitation should, where possible, be in an MTF equipped for this within two hours of wounding.”

UK Army Doctrine advocated the 10-1-2 timeline to guide decision making regarding the configuration and location of the MEDEVAC and MTFs, while recognizing the enduring utility of the 1-2-4 hour principle that focuses on the timeliness for casualty movement between DCS and Primary surgery. Both US Joint (JP4-10 2006) and US Army (FM 4-02.2 May 2007) doctrine described patient precedence for evacuation as:

“within two hours for Urgent cases, within four hours for Priority cases and within 24 hours for Routine cases.”

The US position changed (as cited in Rasmussen, 2016), however, following Congressional Interest in late 2008 and 2009 and resulted in changes for prehospital evacuation of:

“One hour for urgent and urgent surgical missions to appropriate medical care.”

This was incorporated into the Army FM 4-02.19 in July 2009 and remained extant in the October 2011 version (ATTP 4-02), the August 2013 version of FM 4-02, and the August 2014 version of ATP 4-02. In these later publications, the guidance was that Urgent cases should be evacuated as soon as possible and within 1 hour, yet the “Urgent-Surgical” category does not specify time to surgical intervention. Priority cases remained as within 4 hours and Routine as within 24 hours. These changes were not made to the Joint Doctrine until the current version was published in 2012. Subordinate doctrine publications such as FM 4-02.25 (Employment of Forward Surgical Teams [March 2003]) offered no further guidance regarding timelines to care. US Army doctrine is generally stated as implementing or being in consonance with the North NATO STANAG. The US Joint Medical Doctrine referenced ABCA (America, Britain, Canada, Australia/New Zealand) publications but no Allied documents.

In summary, there were differing doctrinal timelines in use over the January 2008 to October 2014 time period in the Afghan theater of operations as described in Table II.

<table>
<thead>
<tr>
<th>Organization</th>
<th>Evacuation Time Planning Policy</th>
<th>Time Period</th>
</tr>
</thead>
<tbody>
<tr>
<td>NATO</td>
<td>1-2-4 hours</td>
<td>(January 8–October 14)</td>
</tr>
<tr>
<td>UK</td>
<td>1-2-4 hours</td>
<td>(January 8–March 11)</td>
</tr>
<tr>
<td>UK</td>
<td>10 minutes-1 hour-2 hours</td>
<td>(March 11–October 14)</td>
</tr>
<tr>
<td>US</td>
<td>2 hours for Urgent Cases and 4 hours for Priority Cases</td>
<td>(January 8–July 9)</td>
</tr>
<tr>
<td>US</td>
<td>1 hour for Urgent cases and 4 hours for Priority Cases</td>
<td>(July 9–October 14)</td>
</tr>
</tbody>
</table>

NATO, North Atlantic Treaty Organization; UK, United Kingdom; US, United States.
In terms of what level of care should be reached within these timelines, NATO Doctrine explicitly states that it should be to definitive surgery, ideally within 1 hour but if not DCS within 2 hours. The UK advocates skilled medical aid within 1 hour and surgery within 2 hours, whereas the US started the campaign with a 2-hour guideline for evacuation of urgent cases without explicitly stating to what level of care. This was changed to 1 hour in July 2009 with the addition of the statement to “appropriate medical care.”

**PATIENT MOVEMENT CAPABILITIES**

Evacuation of casualties is a crucial part of the deployed Health Service Support system and requires specific medical personnel and assets. Time to care creates interdependency between evacuation, treatment, and the theater-holding policy, with each directly impacting the other if the standard of patient care is to be maintained. Thus patient movement is not simply a transportation task but is part of the continuum of care and a medical responsibility.

NATO doctrine advocates that a medical evacuation system should have the following capabilities:

(a) The ability to evacuate casualties to a MTF 24/7, in all weather, over all terrain and in any operational circumstances

(b) The provision of the necessary clinical care throughout the journey

(c) The ability to regulate the flow and types of patients.

Unlike the Roles used to describe MTFs, NATO doctrine describes MEDEVAC, be it ground or air (Aero-medical Evacuation [AE]), in terms of where along the chain of evacuation it operates giving three main categories:

(a) **Forward MEDEVAC/AE**—Point of wounding to the initial MTF. This is required by operational circumstances to meet clinical timelines.

(b) **Tactical MEDEVAC/AE**—between MTFs within the Joint Operational Area.

(c) **Strategic MEDEVAC/AE**—from the Joint Operational Area, to the home nation or other country/safe area.

Although NATO doctrine states the priorities and dependency of patients requiring evacuation (see above), it provides no guidance as to the levels of medical capability required; the focus is on the transport assets and the process to control them. Where specific skills are mentioned, guidance remains broad. AJP 4-10(A) merely states the range of potential capabilities when discussing prehospital ground evacuation transportation:

“There is variation in terms of capabilities and patient capacity. Most will be equipped for basic life support only, but at the top of the scale are advanced support units, staffed with emergency care medical specialists and/or trained specialist paramedic personnel who can provide extended resuscitative care, administer drugs, and begin administration of intravenous fluids in addition to providing basic first aid.”

It takes a similar line with Tactical AE of pre- and post-operative patients, recognizing only the requirement for specialist clinical staff and equipment. AJMedP-227 in its discussion on Incident Response Teams suggest that medical capability could range from paramedical staff to primary health care professionals with advanced resuscitation training, to specialist secondary care teams.

US Joint Doctrine focuses on transport assets, priorities and process although in the Appendices of both publications reference is made to the Critical Care Air Transport Team requiring specialty or critical care capability. US Army doctrine focuses on the priorities for evacuation and not medical capabilities. UK Joint Doctrine does not contain a specific section of medical transfer/evacuation and like the Allied and US doctrine it focuses on priorities and responsibilities rather than capabilities. Guidance in the later UK Joint doctrine refers only to appropriately trained medical staff except when describing the Medical Emergency Response Team (MERT) capability:

“It is based on para-medics or emergency medicine nurses but may be augmented by medical officers experienced in skills such as advanced airway management, rapid sequence induction and the maintenance of anesthesia.”

UK Army doctrine also acknowledges that the MERT requires crew augmentation for prehospital emergency care interventions but does not specify further.

**MEDICAL TREATMENT FACILITIES**

NATO MTF Role Terminology should provide a common language that enables planners to determine the theater laydown and facilitates interoperability. In practice, however, national caveats and mission specific nuances have blurred the boundaries over the last 10 to 15 years.

AJP 4-10(A) categorizes MTFs into four tiers or Roles on a progressive basis (Table III). Each Role of care is defined by a minimum clinical capability and not by its capacity or maneuverability. In principle, each MTF contains the minimum capabilities of the Roles below it, while an MTF cannot be reduced below the minimum capabilities of its given numeric descriptor. UK Joint Doctrine initially used NATO terminology, but in the later publication, more caveats are introduced. UK Army Doctrine uses NATO terminology without exemptions but does note that boundaries can be blurred. The earlier versions of US Joint Doctrine did not use the term Roles, instead describing healthcare capabilities from prevention through to definitive care, and only referred to the NATO definitions in a later chapter. This changed in July 2012 when the NATO definitions were included in the main text. US Army Doctrine initially uses the term “Levels” rather than Roles.
but, in broad terms, they describe the same medial capability. This changed in the later doctrine \textsuperscript{21,29} with Roles replacing Levels in line with the NATO terminology. Its R2 description remained consistent throughout, stating that they have the capability to provide packed red blood cells (liquid), limited x-ray, and clinical laboratory support but not surgery. A note emphasizing this appears in the October 2011 publication highlighting the differences with the Allied publications (See below). There are minimal differences in the definitions of R1 and R3 MTFs used in Allied, \textsuperscript{8,15} UK, \textsuperscript{13,15,30} and US \textsuperscript{28,29,34} doctrine; Joint or single Service. The significant differences are in the descriptions of what constitutes a R2 capability.

R2 MTF

NATO defines R2 as:\textsuperscript{35,36}

“providing an intermediate capability for the reception and triage of casualties, as well as being able to perform resuscitation and treatment of shock to a higher level than Role 1. It will routinely include DCS and may include a limited holding facility for the short term holding of casualties until they can be returned to duty or evacuated. It may be enhanced to provide basic secondary care including primary surgery, intensive treatment unit and nursed beds.”

NATO doctrine\textsuperscript{36} also introduced a delineation in R2 capabilities; those able to support R2 light maneuver (R2[LM]) and the more clinically capable variant R2 enhanced (R2[E]). The R2(LM) MTFs are described\textsuperscript{36} as able to conduct triage and advanced resuscitation procedures up to DCS. They will usually evacuate postsurgical cases to R3 (or R2[E]). In addition to R1 capabilities, R2(LM) will include

(a) Specialist medical officer led resuscitation with required support elements
(b) Routine DCS with postoperative care
(c) Field Laboratory capability
(d) Basic imaging capability
(e) Reception, regulation and evacuation of patients
(f) Limited holding capacity.

The same doctrine\textsuperscript{38} describes R2(E) MTFs as:

“small field hospital providing basic secondary health care, built around primary surgery, ICU and nursed beds. It is able to stabilize post-surgical cases for evacuation to Role 4 without needing to put them through a Role 3 MTF first.”

In addition to R2(LM), R2(E) will include

(a) Primary (definitive) surgery
(b) Surgical and medical intensive care capability
(c) Nursed beds
(d) Enhanced field laboratory including blood provision.

Initially, UK Joint Doctrine\textsuperscript{13} did not recognize the NATO subdivision but merely stated that R2 MTFs “may, in certain circumstances, include DCS when it will be known as R2+.” This is rectified in a later publication\textsuperscript{30} describing R2(LM) as providing “advanced resuscitation up to DCS” and R2(E) MTFs able to provide Primary surgery and evacuate directly to R4. The later doctrine also includes blood availability but only at R2(E) MTFs. UK Army doctrine\textsuperscript{13} from March 2012

\begin{table}[h]
\centering
\caption{Levels (Roles) of Trauma Injury Care}
\begin{tabular}{|l|l|}
\hline
Levels (Roles) of Trauma and Injury Care & Function \\
\hline
Role I Initial level of care/immediate lifesaving measures. \hspace{1cm} Battlefield Care to Battalion Aid Station \hspace{1cm} Similar to civilian first responders. \hspace{1cm} Also includes: Battlefield Care (Self-Aid/Buddy Aid, Combat Lifesaver and Combat Medic), \hspace{1cm} Battalion Aid Station (far forward aid station with at least one physician available). \hline
Role II \hspace{1cm} Small, highly mobile, austere surgical team. \hspace{1cm} Forward Surgical Team \hspace{1cm} Provides life-and-limb saving surgical care and typically the first level of surgery available. \hline
Role III \hspace{1cm} Limited capabilities, some laboratory, X-ray, mental health and dental services may be available. \hspace{1cm} Combat Surgical Hospital \hspace{1cm} Provides full range of surgical, medical, laboratory, and radiology capability. \hspace{1cm} Air Force Theater Hospital \hspace{1cm} Care also includes dental, physical therapy, mental health, obstetrics/gynecology, and primary care services. \hline
Role IV OCONUS \hspace{1cm} Definitive medical and surgical care. \hspace{1cm} Example: Landstuhl Regional Medical Center \hspace{1cm} Outside of area of military operations or combat, but not within CONUS. \hline
Role IV CONUS \hspace{1cm} Stabilization point before evacuation to CONUS. \hspace{1cm} Walter Reed National Military Medical Center \hspace{1cm} Definitive medical and surgical care OCONUS. \hline
Medical Center, Brook Army Medical Center \hline
\hline
\end{tabular}
\end{table}

Adapted from Horne et al, 2014; Silva 2014; Cubana et al, 2013.\textsuperscript{32–35}
is coherent with the joint doctrine in its definitions of both R2(LM) and R2(E).

US Joint Doctrine initially acknowledged Allied terms only adopting them in the July 2012 publication. US Army Doctrine, however, retained its definition and added a note to this effect in October 2011:38

“Note. The R2 definition used by NATO forces (Allied Joint Publication-4.10[A]) includes [the following] terms and descriptions not used by US Army. US Army doctrine subscribe to the basic definition of a R2 MTF providing greater resuscitative capability than is available at Role 1. It does not subscribe to the interpretation that a surgical capability is mandatory at this Role per the NATO doctrine. The NATO descriptions are

– A medical company with a collocated forward surgical team may be referred to as a light maneuver R2 facility.
– An enhanced R2 MTF may be used in stability operations scenarios and consists of the medical company, forward surgical team, and other specialty augmentation as deemed appropriate by the situation.”

It should be noted that one of the key capabilities of Forward Resuscitative Surgical Teams is its ability to function effectively when independent of a R2 MTF. US Army Forward Surgical Teams, US Air Force Mobile Field Surgical Teams and US Navy Forward Resuscitative Surgical Squadrions are all able to integrate with traditional R2 MTFs but are also designed to rely on evacuation assets to rapidly clear stabilized patients, sometimes immediately after surgical procedures are completed. One damage control surgical capability, the US Air Force Tactical Critical Care Evacuation Team—Enhanced, took the next logical step of integrating forward resuscitative surgical care directly into the evacuation platform—allowing evacuation and surgical stabilization to occur in concert. During the Afghanistan conflict, tactical evacuation capabilities routinely served this role and compensated for the increased patient acuity by providing advanced clinical providers (Emergency Medical Technician-Paramedics, Critical Care Nurses, and Emergency Medicine Physicians) when needed.

DISCUSSION
How have operations in Afghanistan impacted on medical doctrine? For the most part this article has focused on the higher levels of doctrine which we noted change only as a result of measured consideration, usually a top-down process. The one significant example of change at this level (US time to care, the “Golden Hour Initiative”)18 only occurred after direction from the Executive authority (highest National Authorities), but even this failed to make it into the joint doctrine until 2012. Otherwise it can be argued that higher level doctrine did not change during the period to reflect reality on the ground; a reality that saw the medical approach to trauma develop significantly. The changes that did occur were captured in the lower tactical levels as standard operating procedures and tactics, techniques, and procedures which were able to react to these changes through bottom up demand. In his thesis “A Revolutionary Approach to Improving Combat Casualty Care,” Hodgetts makes the case that over this period we have seen a revolution in military medical affairs. A summary of the doctrinal changes Hodgetts states in his thesis is at Table IV.

These changes are prescriptive and are more about “what” to do rather than “how” to think. That said, the effect these changes have had on the outcomes for trauma casualties on the battlefield cannot be disputed. The lessons from this campaign will influence higher level doctrine but before changes can be incorporated it is necessary to be clear about what is enduring and applicable universally rather than adaptations specific to that theater of operations or campaign. The question now is how should our experiences in Afghanistan shape our higher level doctrine for the future?

Some change has already occurred; the latest edition of the Allied Joint Doctrine for Medical Support (AJP 4-10[B]) now includes the 10-1-2 guidelines and the level of care to be reached within each time frame as well as the concept of Damage Control Resuscitation (DCR). Yet, despite a stated willingness to adopt Allied doctrine there are still many national caveats regarding what constitutes a R2 MTF and the time lines to a particular Level of Care. R2 MTFs can, under current guidance, be anything from a higher capability than a R1 MTF to a small hospital. This span is probably too great and hinders both planners and interoperability among allies when different capabilities are mandated (e.g., lack of required surgical capability for US R2 elements). As the R2(E) is accepted as being a “deployed hospital” then perhaps it would be simpler if “R3” identified any “deployed hospital” with a suffix denoting its level of capability (level I, II, III). The Afghan campaign highlighted the nonlinear nature of medical support where patients can move from point of injury to surgical care without any

<table>
<thead>
<tr>
<th>Significant Doctrinal Changes During Period of Operation Enduring Freedom (c. 2002 to 2014)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABC to &lt;C&gt;ABC</td>
</tr>
<tr>
<td>Tourniquet Use</td>
</tr>
<tr>
<td>4 Stages of Combat Resuscitation</td>
</tr>
<tr>
<td>Rapid Primary Survey</td>
</tr>
<tr>
<td>MIST Report</td>
</tr>
<tr>
<td>Clinical Guidelines for Operations</td>
</tr>
<tr>
<td>Damage Control Resuscitation</td>
</tr>
<tr>
<td>Hemostatic Resuscitation</td>
</tr>
<tr>
<td>Immediate Surgical Intervention Upon MTF Arrival</td>
</tr>
</tbody>
</table>

ABC, airway, breathing, circulation; “<C>,” catastrophic hemorrhage; MIST, mechanism, injuries, symptoms, treatment; MTF, medical treatment facility.
intermediate steps. Thus there is no need to categorize deployed hospitals as a R2 simply to show that it is further forward. Ultimately, this would then allow R2 MTFs to focus on the intermediary nonsurgical capability on the way to the deployed hospital, and subsequently, allow the use of the term for MTFs without surgery. This is something the US has kept within its doctrine as it envisages such facilities being the norm in any larger scale near-peer conflict.

The UK Joint doctrine has now been archived and replaced by AJP 4-10(B), but still retains a caveat stating the UK uses the 10-1-2+2 timeline. The primary difference for the US is that R2 does not have to contain surgical capability, therefore timelines are to an appropriate MTF within an hour and not to surgery. This less demanding position is necessary in the higher doctrine as it needs to be relevant to all future campaigns and not just what happened during the most recent operation. Doctrine must also reflect the realities of a large-scale conflict with a near peer opponent.

Conversely, another potential solution would be to disassociate forward surgical capabilities from the entanglements of the Role definitions. The tactical advantage of small surgical teams integrated into a joint trauma system can melt away if encumbered by doctrinal attachment to MTFs designed primarily to support trauma care delivery. These small teams can functionally bring lifesaving capabilities to operational areas that would otherwise be supported at the R1 (R1) level, fully integrate with R2 MTFs and augment R3 advance surgical capabilities to increase surgical throughput. If forward surgical care doctrine is to be most effective in future operational environment against near peer adversaries, it must recognize the tactical advantage of small size and unencumbered mobility of forward surgical capabilities operating independent of MTFs and reliant on tactical evacuation for relief of limited holding bed capability. This concept is consistent with most existing NATO and other doctrine that focuses on time to appropriate surgical intervention/DCR as opposed to defined Role of care.

One of the significant changes seen in Afghanistan was the increase in the range and the medical capabilities carried on patient movement assets. Typically there were three levels of capability available to the Patient Evacuation Coordination Cell to task described in Table V. There may be an advantage in having levels of MEDEVAC assets in the same way we have levels of MTFs, each with an agreed level of medical capability (probably not far off those above). In the same way MTF capabilities assists planners in configuring the theater laydown, so will agreed MEDEVAC capabilities. It will equally help develop the mutual understanding required if the higher levels of interoperability are to be achieved. This will then drive changes that support the intelligent tasking of the various MEDEVAC capabilities; a requirement for quicker more timely medical information and the availability of a Medical Common Operating Picture.

Limitations in this review include the limited scope of the analysis from 2008 to 2014; future evaluation will focus on allied doctrine from 2014 and beyond. The primary goal of this review was trauma-related combat casualty care, yet primary care and disease nonbattle injury comprises much of deployed medical care and is the driving force behind much of the current R2 doctrine, particularly for the US Army. These doctrinal differences were not addressed in this article.

A larger and more difficult question may be the definition of an alliance or Joint force. Currently there is no mandate to reach consensus on doctrine, yet failure to do so inherently hinders interoperability. Perhaps the solution is more integrated war-gaming and sharing of resources and greater inter-agency and inter-national compromise to reach consensus.

CONCLUSION

Several discrepancies in medical doctrine persist among allied forces. Timelines to definitive care vary among nations. We as allied nations should have clear taxonomy that clearly defines MTF capabilities within the combat casualty care system. The R2 surgical capability discrepancy between US and NATO doctrine should be reconciled. Medical evacuation capabilities on the battlefield would be improved with a

<table>
<thead>
<tr>
<th>Casualty Evacuation Type</th>
<th>Highest Level of Medical Provider</th>
<th>“Level” of Medical Capability</th>
<th>Offensive Arms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lift of Opportunity</td>
<td>Combat Life Saver</td>
<td>&lt;1</td>
<td>Y</td>
</tr>
<tr>
<td>US Army Air Ambulance (“DUSTOFF”)</td>
<td>Flight Paramedic</td>
<td>1</td>
<td>N</td>
</tr>
<tr>
<td>US Army Air Ambulance (Augmented)</td>
<td>Critical Care Registered Nurse</td>
<td>2</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>and Flight Paramedic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>US Air Force Rescue Squadron (“PEDRO”)</td>
<td>Para-medic</td>
<td>2</td>
<td>Y</td>
</tr>
<tr>
<td>US Air Force Tactical Combat</td>
<td>Emergency Medicine Physician, Nurse Anesthetist,</td>
<td>3</td>
<td>N</td>
</tr>
<tr>
<td>Casualty Evacuation Team (TCCET)</td>
<td>Emergency or Critical Care Nurse</td>
<td></td>
<td></td>
</tr>
<tr>
<td>US Air Force Tactical Critical Care</td>
<td>Surgeon, Emergency Physician, Nurse Anesthetist,</td>
<td>3+ (Surgical capability)</td>
<td>N</td>
</tr>
<tr>
<td>Evacuation Team–Enhanced (TCCET–E)</td>
<td>Emergency Nurse, Operating Room Technician</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medical Emergency Response Team (MERT), United Kingdom)</td>
<td>2 Physicians, Emergency Nurse, 4 Para-medics</td>
<td>3</td>
<td>Y</td>
</tr>
</tbody>
</table>
taxonomy that reflected the level of capability. Such changes may improve interoperability in a dynamic military landscape.

REFERENCES


Establishing a Joint Theater Trauma System During Phase Zero Operations

COL Jennifer J. Walker, MC USA*; CAPT Zsolt T. Stockinger, MC USN†; RADM Colin G. Chinn, MC USN‡

ABSTRACT Objectives: Military personnel risk injury due to accidents, disasters, and military threats during Phase Zero “shaping” operations. Medical facilities must be poised to respond. Methods: The U.S. Pacific Command (PACOM) Area of Responsibility (AOR) covers more than 50% of the earth’s surface; relevant Clinical Practice Guidelines must include the maritime setting and extended evacuation periods. Military hospitals in the region are not connected by a defined Trauma System. There is variable adherence to trauma training requirements before assignment in this AOR. Demand for trauma care at any 1 location is low and trauma teams have little opportunity to maintain competency for high-risk/low-volume interventions. There is no documentation of total demand for trauma care in the AOR. Trauma care in PACOM is often deferred to civilian facilities. Results: Core elements of a Joint Theater Trauma System (JTTS) as established during combat operations in U.S. Central Command are applicable during Phase Zero. A PACOM JTTS was established to address the region’s readiness to respond to Phase Zero trauma as well as escalation of regional threats. Information technology coordination was a critical hurdle to overcome. Conclusion: PACOM lessons learned are applicable to other Geographic Combatant Commands developing a JTTS during Phase Zero operations.

INTRODUCTION

From a Joint doctrine planning perspective, “Phase Zero” operations include steady state, peacetime, shaping operations like theater security cooperation missions and exercises and engagements with partner nations (Fig. 1). Service members in the U.S. Pacific Command (PACOM) are frequently engaged in these operations throughout the region. While major trauma during Phase Zero is fortunately rare for Service members, medical providers, and military treatment facilities (MTFs) must be poised to respond to accidents, natural, and man-made disasters, and escalation of military threats.

The value of military trauma systems has been well documented and analyzed elsewhere. Elements of the Joint Theater Trauma System (JTTS) developed for combat operations in the U.S. Central Command (CENTCOM) are applicable during Phase Zero operations. Establishing the core aspects of a JTTS during peacetime allows a region’s medical teams to capitalize on experiences gained during conflict in Afghanistan and Iraq, and not be caught off guard if the security situation were to abruptly change. Escalation from Phase Zero operations will require prompt response from medical teams.

Development of a JTTS in the PACOM Area of Responsibility (AOR) during Phase Zero operations comes with specific challenges, and with the assistance of the San Antonio-based Department of Defense (DoD) Joint Trauma System, major differences between CENTCOM and PACOM were identified.

Differences including extremely varied terrain, immense span of control, and lack of dedicated personnel or other resources to support implementation must be overcome. Medical coordination between the Services can be a challenge. Nonetheless, lessons learned in PACOM are applicable to other Geographical Combatant Commands wishing to develop a JTTS during Phase Zero operations.

The goal is to improve trauma delivery and patient outcomes across the continuum of care and military operations, using existing medical infrastructure with improved communication networks along with evidence-based interventions and attention to process improvement.

The PACOM AOR

The PACOM AOR covers 36 countries and over 50% of the Earth’s surface—spanning “from Hollywood to Bollywood”—much of it in a maritime environment (Fig. 2). The region is home to several of the world’s largest armies, introducing the possibility of rapid escalation from Phase Zero operations. Other regional threats including radicalization and growing tensions between states may escalate with little notice. Border disputes, nuclear capability, and transnational terrorism threaten security in the AOR. There are many overall threats to health in the Asia–Pacific region, including economic disparity, food, and water security issues, and man-made and natural disasters, which may also impact overall security within the region.

The PACOM Command Structure includes Army, Navy, Air Force, and Marine Corps subordinate component commands. However, the Command Surgeon at HQ USPACOM does not uniformly have direct command and control or tasking authority over the medical support assigned to the nine bedded hospitals and numerous Role 1 and 2 facilities within these commands. The Army Role 3 facility in Seoul and the Role 4 Tripler Army Medical Center on Oahu fall

---

*Graduate Medical Education, Womack Army Medical Center, 2817 Reilly Road, Fort Bragg, NC 28310.
†Joint Trauma System, 3698 Chambers Pass, Suite B, U.S. Army Institute of Surgical Research, Fort Sam Houston, TX 78234.
‡Research, Development and Acquisition, Defense Health Agency, 7700 Arlington Boulevard, Suite 5101, Falls Church, VA 22042.

The work was presented at the Military Health System Research Symposium, Fort Lauderdale, FL, August 17-20, 2015.

doi: 10.7205/MILMED-D-16-00167
under Regional Health Command—Pacific, which is subordinate to U.S. Army Medical Command, not U.S. Army Pacific (USARPAC). Likewise, Navy Role 3 facilities at Okinawa, Yokosuka, and Guam fall under Navy Medicine-West in San Diego, not the Commander, U.S. Pacific Fleet, who indirectly controls Navy shipboard Role 2 assets. Air Force Role 3 facilities in Alaska, Osan, Misawa, and Yokota take direction from the Pacific Air Force (PACAF) Surgeon. There is no defined Trauma System connecting these facilities.

While there are deployable medical assets available to respond to an escalation of threat in the region, these assets are not all immediately available. Surgical assets assigned to Marine Forces Pacific may take up to 72 hours to be in place. USARPAC can provide various levels of support to medical operations including a Forward Surgical Team that requires a 72 to 96 hour movement window. Permanent PACFLT assets available for response include the USNS Mercy, a 1000-bed Role 3 facility that may take a month to embark and transit from San Diego and be operational in theater; one aircraft carrier with one operating room, three intensive care beds, and 51 bed capacity; one large deck amphibious assault ship with one of four operating rooms staffed, and 64 bed capacity; and a recently added Expeditionary Resuscitative Surgical System, making a small Role 2 surgical capability deployable on nearly any platform and fully operational within 96 hours.5 PACAF’s assets include an Expeditionary Medical Support Health Response Team, a mobile tent hospital facility with surgical capability, ready to load within 24 hours after personnel arrive at Anderson Air Force Base in Guam from bases throughout the AOR. This highlights the need to rely on our existing MTFs to provide care in the region during any escalation from Phase Zero.

It should be noted that much if not most of the PACOM AOR is not covered by MTFs, and Phase Zero trauma care is often deferred to civilian facilities through Tricare Area Operations-Pacific (TAO-P). However, local standard of care at civilian facilities is variable and medical record keeping may be incomplete or come through the military trauma system without translation.

**Initiating a JTTS During Phase Zero Operations**

Some of the region’s medical facility leadership warmed slowly to initial JTTS development efforts, perhaps due to anticipation for additional resourcing requirements; developing a JTTS connecting the region’s subordinate commands and bedded facilities mandated identification of a “common denominator” in their chains of command. The USPACOM Chief of Staff signed a JTTS Charter tasking the PACOM Surgeon to Chair an Advisory Board with participation from the Component Surgeons at USARPAC, PACFLT, PACAF, Marine Forces Pacific, and Special Operations Command-Pacific. The purpose of the PACOM JTTS, as outlined in
the Charter, was to recommend medical trauma policy guidance, processes, equipment standardization, training, and trauma registration, with objective to optimize peacetime and combat joint intratheater casualty care from point of injury through final disposition. The region’s medical leadership adopted the JTTS, once it became clear that the concept was to take what already existed and connect it into a better-integrated system.

Development of JTTS Working Groups

There is no dedicated staff or “JTTS team” assigned in PACOM AOR, as there was in CENTCOM from 2005 to 2014, which proved to be one of the biggest challenges to implementation of a JTTS during Phase Zero. The PACOM JTTS Charter tasked the PACOM Component Surgeons to designate staff members to support the PACOM JTTS Clinical Operations and Medical Informatics Working Groups. The Commander of the U.S. Army’s 18th Medical Command (Direct Support) became an enthusiastic supporter of JTTS development and the unit’s Oahu-based staff offered to chair the video teleconferenced Working Groups. It became obvious that JTTS needed to have representation from the MTFs, and not just from the Component Surgeon staffs, in order to get the JTTS message out and really make it a “system”. CENTCOM-based JTTS position descriptions for the Trauma Medical Director, the Trauma Program Coordinator, and the Medical Informatics Technology Coordinator were personalized for the PACOM AOR. Following buy-in from the individual services’ regional medical leadership, MTFs were tasked to designate appropriate representatives to participate in the development and execution of JTTS. These personnel all support the PACOM JTTS as an additional duty during
Phase Zero operations, and participation in these working groups was occasionally trumped by other component- and facility-specific staff duties. The individual MTF Trauma Medical Directors and Program Coordinators participated in the Clinical Operations Working Group alongside the clinical representatives from the Component Surgeons; our experience is that the physicians and nurses were exceptionally eager to participate in the development of this network. Medical information technology (IT) Coordinators from the MTFs contributed to the Medical IT Working Group with representatives from the Component Surgeons; however, we were challenged to make Phase Zero JTTS relevant to these working group members given the defined and time-consuming nature of peacetime medical documentation requirements.

Phase Zero tasks and milestones for the PACOM JTTS working groups had to be developed. Initial Clinical Operations Working Group tasks focused on identifying trauma subject matter experts in the region and elsewhere, availability of trauma training in the region, and application of CENTCOM clinical practice guidelines (CPGs) that are relevant to PACOM. The Medical IT Working Group tasks focused on establishing a central repository for trauma-related documents, training materials, and contact information for trauma experts, with data accessible to all Services. Both groups collaborated on initial entry of data into the San Antonio-based DoD Trauma Registry (DoDTR) and conduct of a PACOM Virtual Trauma Care Conference similar to that already existing in CENTCOM.

Identifying and Linking Key Trauma Care Personnel in AOR

Through personal contact with clinical leadership at each bedded facility, a contact roster for trauma care personnel at Role 2, 3, and 4 facilities was developed and maintained on Joint Knowledge Online. We discovered that MTFs in the region are not routinely staffed for the potential trauma demands associated with intensified operations. Fellowship-trained trauma surgeons are not assigned to all MTFs in Phase Zero; general surgeons assigned to many of the MTFs in the AOR tend to be junior, often straight out of residency training, with limited experience in complex trauma surgery including major abdominal, thoracic, and vascular cases. Virtually linking these junior clinical teams with more experienced providers elsewhere in the region and at the Joint Trauma System in San Antonio added to their readiness.

Compounding this issue, unlike deployment in CENTCOM, there is no standardized “predeployment training” before assignment in this region. We also discovered that Components do not uniformly adhere to Component-specific trauma training guidelines. The theme seemed to be that losing units did not want to send a provider to training before reassignment to PACOM, and once the provider arrives at the gaining unit there is not enough clinical overlap or funding to allow for attendance at trauma courses outside the region. Why is this relevant in Phase Zero? Consider a Soldier who suffers a severe traumatic brain injury while assigned at the Army base at Yongsan Republic of Korea. There is no neurosurgeon assigned at the Army hospital. The general surgeons assigned there did not have military-specific trauma training before assignment and have not recently refreshed on lifesaving neurosurgical techniques. The host nation providers may or may not follow the Joint Trauma System’s CPGs for management of severe head trauma. Before assignment in the region, attendance at a trauma care course that addresses skills needed for trauma management in a resource-constrained environment serves as a refresher for infrequently practiced lifesaving skills.

The operational tempo in Phase Zero operations does not support maintenance of perishable trauma care skills. Most of the Role 3 or 4 facilities rarely care for severely traumatized patients. Demand for trauma care at any one MTF is routinely low, and trauma teams have little opportunity to maintain competency for high-risk/low-volume interventions. It is not just an issue for the surgeons and other physicians—the emergency room and intensive care unit nurses, medics, and corpsmen do not have the opportunity to keep their skills refreshed either. Given the geography, consistent adherence to “Golden Hour” evacuation standards is impossible, so our Role 1 medical teams need to be prepared and rehearsed as well.

To mitigate this potential gap in readiness, addressing the apparent deficiency in the region’s ability to medically support the Commander’s priority to “be ready to fight tonight” the Clinical Operations Working Group identified and advertised trauma training resources and opportunities from within and outside of the AOR. Trainers in ATLS and other recognized courses were identified, and courses were opened to all Services and all facilities in the AOR. Novel and inexpensive training platforms were devised, like a static C-130 evacuation training model for medical staff assigned to Kadena Air Force Base in Okinawa, Japan.

Clinical Practice Guidelines

Initially, the Clinical Operations Working Group asked clinical experts in PACOM to review the CENTCOM-centric CPGs and adapt them for PACOM. We quickly realized that the CENTCOM-based CPGs are largely applicable as is, once specific references to facilities and capabilities available in CENTCOM are edited out. Trauma Care CPGs have been shared with host nation providers—through personal contact and through TAO-P. Time will tell if the host nation providers and facilities adopt or modify the U.S. CPGs.

The Clinical Operations Working Group identified several CPGs that should be developed for the PACOM AOR, including:

- Catastrophic attack on a ship in blue water
- Near drowning
- Care of patient during prolonged stays at Role 1 and Role 2 facilities
- Care of trauma patient during prolonged evacuation times

Unfortunately, the AOR does not have the personnel or resources to develop these CPGs from scratch. The Joint Trauma System has taken the initiative to rewrite all existing
Documenting the total demand for Phase Zero trauma care throughout the AOR is problematic. While there is unfortunately no DoD policy that requires DoDTR data entry during Phase Zero, one of the tasks of the PACOM JTTS Charter was to establish the ability to enter data into the DoDTR. Tripler Army Medical Center is a contributor to the DoDTR, but only for trauma patients that have been evacuated from Landstuhl—they are not currently initiating DoDTR data entry for trauma that occurs on Oahu or transfers from elsewhere in PACOM AOR.

DoDTR data entry is relevant during Phase Zero operations: how can we confidently assure our commanders that we are ready to “fight tonight” if we have no true idea how well-practiced our trauma teams are? While a single facility can fairly easily document the trauma care delivered at that individual facility, there is no other process in place to document the total demand for trauma care collectively throughout the region during Phase Zero operations. Incidentally, the DoDTR meets the American College of Surgeon’s requirement for Trauma Registry participation to achieve official American College of Surgeon trauma center designation, for deployed or combat-related trauma. State Trauma Registry documentation requirements in Alaska and Hawaii might overlap with DoDTR data entry.

With the help of the DoDTR team at the Joint Trauma System in San Antonio, guidelines were developed for MTFs to enter baseline trauma care data into the DoDTR. Initial DoDTR data entry Inclusion focused on active duty military receiving care in an MTF or through Tricare at a local facility, for 1 or more of a set of trauma International Classification of Diseases, 9th Revision, Clinical Modification codes. Data Entry Fields were identified to document baseline demand for trauma care in Phase Zero, not to develop or verify adherence to CPGs (as they are used in CENTCOM).

Aside from one-deep offices at MTFs in Alaska and Hawaii, there are no additional personnel to enter data into the DoDTR. Attempting to duplicate the successes of the CENTCOM JTTS, the PACOM JTTS Medical IT Working Group analyzed the methods of CENTCOM DoDTR remote and direct data acquisition. During Phase Zero operations, most of the PACOM trauma treatment is scanned in record in the Armed Forces Health Longitudinal Technology Application, whereas medical treatment received in a combat zone is recorded in a searchable, web-based application known as Theater Medical Data Storage. Theater Medical Data Storage allows for tracking injured patients as they move through theater levels of care as well as remote acquisition of trauma data. Given the nonstandardized documentation methods throughout the AOR during Phase Zero—including scanned-in trauma records in foreign languages and various facility-specific trauma care reporting tools—it is difficult to rely on remote acquisition of data into the DoDTR. The direct method of DoDTR data acquisition and entry using Trauma Program Coordinators is more accurate, but costly. Joint Trauma System personnel host an initial 10 + hours of web- and teleconference-enabled training to familiarize with DoDTR processes. Some overseas facilities may have annual personnel changeover, necessitating annual renewal of training. Regardless of remote or direct acquisition of trauma data, it is difficult—if not impossible—to extract and record trauma data from TAO-P records into the DoDTR, given language barriers and multiple standards for documentation which may or may not be adhered to.

The Virtual Trauma Care Conference
A monthly “PACOM JTTS Virtual Trauma Care Conference” brought together geographically separated teams along with trauma care experts from throughout the DoD, highlighting issues related to garrison-based trauma cases from the region. This Virtual Trauma Care Conference, certainly one of the “best practices” implemented in PACOM during Phase Zero, discusses all steps along the care continuum—including prehospital care through patient movement back to continental United States. Cases are reviewed with attention to adherence to Joint Trauma System CPGs. Following the case discussion, a trauma subject matter expert discusses a timely topic that is relevant to military medical care.

Recent cases discussed included:

— Service member with catastrophic head trauma during rappelling training in Nepal; provided good opportunity to discuss imaging and neurosurgical capability in the AOR.

— Service member blown off a roof by helicopter rotor wash during a training accident in Thailand, suffering a severe pelvic fracture; good opportunity to review use of pelvic binders.

— Service member injured in a land mine blast during a training exercise near the Cambodian border; good opportunity to review prehospital care and use of blood products at local national facilities.

Finally, all facilities are offered the opportunity to share their recent wins or challenges with the group. Conveying this event over Defense Collaboration Services (DCS) is one way to standardize communication networks. Unfortunately, not all facilities have access to voice on DCS. To mitigate this issue, slides are displayed via DCS and phone line teleconferencing is used for voice.

Blood Support During Phase Zero Operations
The Joint Blood Program Officer at PACOM implemented advances in blood support throughout the AOR. While these efforts were conducted outside the PACOM JTTS continuum and fit the suggested model of planning for future
contingencies, nevertheless they add to trauma support during Phase Zero operations. U.S. Naval Hospital Okinawa was selected as the site for liquid/never frozen plasma production, which is Food and Drug Administration approved with a shelf life of 26 days. This local production mitigates the lack of plasma thawing equipment at Role 2 facilities used during exercises and missions. Prepositioning of U.S. Food and Drug Administration-approved blood products in deployed MTFs or host nation hospitals assures a safe supply for U.S. Service members participating in exercises and engagements. The Joint Trauma System’s Fresh Whole Blood CPG standardized collection procedures for the Expeditionary Resuscitative Surgical System in support of maritime small unit missions.

In addition to these clinical advances that directly affect U.S. Service members, the Joint Blood Program Officer at PACOM sponsored Blood Safety Workshops in several countries in the region. Through partnerships with Department of State, U.S. Centers for Disease Control, U.S. Agency for International Development, the World Health Organization, Red Cross, local Ministries of Defense, and others, these workshops include participation from host nation military and civilian providers and have been lauded as models of interagency/international/civil–military coordination. With a focus on building banking, transfusion, and testing capability, these workshops enhance local national preparedness for disasters.

CONCLUSION
Lessons Learned from the development of the PACOM Trauma System are applicable to other Geographic Combatant Commands wishing to implement JTTS during Phase Zero operations:

— Establishing the skeleton for a regional JTTS during Phase Zero operations identifies potential gaps and/or weaknesses in medical, logistical, and transportation capabilities.

— Identifying key personnel (Trauma Medical Directors, Trauma Program Coordinators, and Medical IT Coordinators) facilitates communication networks in the event of escalation.

— Sharing lessons learned from garrison-based trauma throughout the AOR helps to maintain a baseline trauma “readiness.” This is best facilitated through periodic real-time teleconferencing to discuss specific cases and issues involved in their care, to highlight strengths of the system and identify real (as opposed to theoretical) weaknesses that must be addressed.

— Data capture is key to identifying baseline trauma requirements, areas for performance improvement, and further development of clinical best practices.

— Each COCOM should collaborate with the DoD Joint Trauma System, as the DoD’s subject matter experts on trauma care and trauma system development, to design and implement a JTTS appropriate to that COCOM’s needs and capabilities.

ACKNOWLEDGMENTS
Thank you to COL Kirby Gross and Col Jeffrey Bailey for exceptional peer mentoring.

REFERENCES


Prehospital Blood Transfusion During Aeromedical Evacuation of Trauma Patients in Israel: The IDF CSAR Experience

Jacob Chen, MD†‡; Avi Benov, MD†‡§; Roy Nadler, MD†; Daniel N. Darlington, PhD†; LTC Andrew P. Cap, MC USA†; Ari M. Lipsky, MD∥∥; Elon Glassberg, MD*

ABSTRACT Background: Data regarding the effect of prehospital blood administration to trauma patients during short-to-moderate time evacuations is scarce. The Israel Air Force Airborne Combat Search and Rescue is the only organization that deals with aeromedical evacuation for both military and civilian casualties in Israel and the only one with the ability to give blood in the prehospital setting. Methods: Data on packed red blood cells (PRBCs) administration in the evacuation missions from January 2003 to June 2010 were analyzed and actual transfusion practice was compared to clinical practice guidelines (CPGs). Results: Over the studied 101 months, a total of 1,721 patients were evacuated by Combat Search and Rescue. Of these, 87 (5.1%) trauma patients were transfused with PRBC. Demographics included 83% male and 17% female with a median age of 23 years. Main mechanisms of injury included gunshot wounds (36%), motor vehicle accidents (28%), and blast injuries (24%) with an average of 2.6 injured regions per casualty. The most commonly injured body regions included lower extremities (52%), chest (45%), and abdomen (38%). Overall, 10 (11%) casualties died. Lifesaving intervention included tourniquets (27%), endotracheal intubation (24%), tube thoracostomy (24%), and needle thoracostomy (21%) times. For 98% of the patients, clinical judgment led to administration of red blood cells before indicated by the CPG. The heart rate tended to decrease during the evacuation, whereas there was no clear trend in systolic or diastolic blood pressure or shock index. Conclusions: In our aeromedical experience, transfusion of PRBCs for trauma patients was safe, feasible, and most likely beneficial. PRBCs were administered according to the flight surgeons’ clinical judgment and not in complete adherence to CPGs in most cases. Data collected from this and similar studies worldwide have led to change in CPGs with the shift from hypertensive resuscitation to hypotensive-hematostatic Remote Damage Control Resuscitation.

INTRODUCTION Trauma is the leading cause of death and disability in young patients between the ages of 1 and 44. Blood loss associated with trauma is the major contributor to preventable mortality in the first 24 hours and represents a targetable parameter for therapies to reduce death due to trauma. Control of hemorrhage constitutes the basis for most therapeutic approaches to traumatic injuries and early volume resuscitation and transfusion of blood components serves as a crucial part of the medical care provided to hemorrhaging patients.2-6

Resuscitative treatment for hemorrhagic shock has a long history. The first record of successful treatment of hemorrhagic shock with fluid resuscitation is from as early as the 17th century.7 The treatment of combat-related injuries has included transfusion of blood products since as early as the World War I,8 though the safety of universal donor blood transfusion was not fully established until the Korean War.9,10 As a result of military experience, type O red blood cell transfusion has become an accepted therapy for resuscitation of trauma victims with hemorrhagic shock in the civilian setting,11 in both hospital and prehospital environments.11,12 Though little research has been conducted in the area of air ambulance blood transfusions, in-flight blood transfusion has been described for more than three decades.12-17

The Airborne Combat Search and Rescue unit (CSAR) is the Israel Defense Force (IDF) heliborne combat medevac extraction unit, subordinate to the Israeli Air Force. During peacetime, the unit provides on-scene support for military medical providers and civilian Emergency Medical Services, as well as inter-hospital transfers of patients; it is also capable of a wide range of technical rescues (e.g., sea, desert, and high altitude). During war time, the unit also provides combat battlefield extraction and medical treatment when necessary to pilots and combat units. CSAR operates on UH-60 and CH-53 helicopters and en route care is delivered by a team of 2 Advanced Life Support providers (physician or paramedic), at least one of whom is a flight surgeon (a reserve physician who is a general surgeon, anesthesiologist, or intensivist) and several combat flight medics.

In this study, we set out to describe the Israeli experience of prehospital blood transfusion during aeromedical evacuation, and to evaluate adherence to the IDF Medical Corps
Methods

**IDF-MC CPG: Transfusion of Red Blood Cells**

At the time of the study period, CPGs regarding blood transfusion were written so that they were consistent with then-current ATLS guidelines. They included:

- Only data regarding patients administered with PRBCs was collected in our study.
- Two units of O-positive PRBCs were to be taken on every aeromedical evacuation. When patients were transfused during inter-hospital transfers, PRBCs were taken from hospitals after PRBCs group and type was established.
- Blood was kept in the flight hangar in a designated refrigerator at 4°C. Refrigerators’ temperature was monitored daily by a medic from CSAR unit according to IDF-Surgeon General (IDF-SG) directive no 962.750.001 mandating twice a day inspection of refrigerators with a designated thermometer (TFA 30.1042 refrigerator thermometer, TFA Dostmann GmbH + Co. KG, Reicholzheim, Germany) to insure that the PRBCs are kept according to regulation.
- Blood was taken on board the helicopter in a simple cooler with cool packs to maintain temperature until January 2009. Starting in January 2009, PRBCs were stored in a thermal transport container (Original Golden Hour Container; Minnesota Thermal Science, Plymouth, Minnesota), a reusable iceless thermal container. The thermal container can be used for up to 72 hours. Maintenance and inspection of stored blood products were performed on a daily basis by the on-call air crew according to IDF-SG directive no 962.750.001.
- Patients appropriate for transfusion included those with obvious or suspected acute blood loss in whom hemodynamic shock persisted after 2 or more liters of crystalloid infusion. Hemodynamic shock was defined as a systolic blood pressure (SBP) <90 mm Hg and/or clinical signs of shock such as altered mental status (with no evidence of traumatic brain injury), heart rate (HR) >120 bpm, pallor, respiratory rate above 30 breaths per minute, or delayed capillary refill.
- Pediatric patients qualified for transfusion if they demonstrated evidence of persistent hemorrhagic shock after the administration of 40 mL/kg of crystalloid bolus, administered in 20 mL/kg increments. For pediatric patients, blood was transfused in 10 mL/kg increments, to a maximum of 40 mL/kg.
- When massive uncontrolled hemorrhage was deemed likely by flight surgeon clinical judgment, blood transfusion could be initiated even before crystalloid administration.
- Blood transfusion took place either by intravenous (IV) or intraosseous route, through a dedicated IV Y-tubing line with a filter. PRBCs were diluted with 200 mL of saline using the Y-tubing line. No blood warming system was used during the research period.

Data Acquisition and Statistical Analysis

The study was conducted as a case series of all documented cases known to be treated with PRBCs between January 1, 2003 and July 31, 2010. The CSAR unit was the only aeromedical critical care service in Israel during this study period. Only trauma patients who were recorded as having received blood products were included in the analysis.

Patient data were recorded in-flight on standardized data collection sheets, and subsequently transferred to a computerized database after mission completion. Data retrieved included gender, age, injured body regions, mechanism of injury (MOI), interventions (including lifesaving interventions [LSI], defined below), blood products received, and patient assessments. Vital signs were recorded at least four times during each mission: on the ground on receiving the patient, after liftoff, before landing, and before transferring the patient to the emergency department or trauma bay. Vital signs included HR, systolic and diastolic blood pressure, mean arterial pressure, and shock index defined as HR divided by SBP.

Injured body regions were grouped into 1) head and neck, 2) trunk, divided into chest and abdomen, and 3) extremities, divided into upper and lower. MOI was grouped into 3 main categories: 1) blast (including improvised explosive device, landmine, mortar, shrapnel, bomb, and grenade), 2) gunshot wound (including shrapnel originating from gunshots), and 3) other.

LSIs were defined according to IDF-MC CPGs to include advanced airway management (endotracheal intubation, cricothyroidotomy, or laryngeal mask airway insertion), needle thoracostomy, tube thoracostomy, arterial tourniquets, and intraosseous access or central IV line placement as primary vascular access. Mission characteristics were defined according to the types of evacuation: combat or noncombat medical evacuation (evacuation from point of injury to hospital [combat or noncombat]) and inter-hospital transfer (transfer of patient between hospitals in Israel). Medical evacuations can also involve extraction form point of injury (where there is a need to extract a casualty [e.g., drowning, cliff, cable extraction]).

Safety

Safety evaluation included assessment of transfusion reactions, other transfusion complications, and exposure to blood by medical teams due to the complexity of handling blood while flying.

Statistical Analysis

Data were recorded in a Microsoft Excel spreadsheet (Microsoft, Redmond, Washington) and analyzed using JMP 12 statistical
software (SAS, Cary, North Carolina). Results for continuous data are presented as medians and interquartile ranges (IQRs), categorical data are presented as absolute numbers and percentiles.

This study was approved by the IDF-MC Institutional Review Board (study number 1049-2011), and the requirement for written informed consent was waived.

RESULTS

Patients and Injury Characteristics

During the study period, 1,721 aeromedical evacuation patients were recorded. Medical evacuations accounted for 76% (1,306) of patients, whereas extractions and inter-hospital transfers accounted for 17.5% (303) and 6.5% (112), respectively. A total of 89 (5.2%) patients received in-flight blood transfusion, of whom 87 (98%) were trauma casualties and two (2%) were nontrauma patients. The two latter patients both required inter-hospital transfer, one suffering from respiratory disease and the other from obstetric/gynecologic uncontrolled bleeding. The 87 trauma patients constituted the study population and the remainder of the results relate to these patients.

Seventy (80%) patients were evacuated as part of a designated medical evacuation mission of whom 60 (69% of 87) were part of combat evacuation missions and 10 (31%) were noncombat missions. All of the nonmedical evacuation missions (n = 17, 20%) were inter-hospital transfers. No casualty that was extracted received blood transfusion. Of all medical evacuation missions, 5.4% received blood transfusion; of all inter-hospital transfers, 15.5% received blood transfusion. Median age was 23 years (20–32). Seventy-two (82.8%) patients were male and 17 (17.2%) were female. Thirty-nine (44.8%) casualties were IDF soldiers and 48 (55.2%) were civilians. Main mechanisms of injury were gunshot wounds (n = 32; 36%), motor vehicle accidents (n = 25; 28%), and injuries inflicted by explosions (n = 21; 24%). Other mechanisms of injury included stab wounds (n = 4; 4%), plane crashes (n = 2; 2%), and falls from height (n = 2; 2%). For one (1%) casualty, the MOI was not documented.

Total of 223 anatomical regions were injured yielding an average of 2.6 injured regions per casualty. The most commonly injured body regions included lower extremities (n = 45; 52%), chest (n = 39; 45%), and abdomen (n = 33; 38%) (Fig. 1).

Overall, 10 (11%) patients died: 7 (8%) patients were pronounced dead at the time of hospital arrival, 2 (2%) died in the emergency department after resuscitation attempts failed, and 1 (1%) died in the intensive care unit.

Mission Characteristics

The median number of evacuees per flight was 1 (IQR = 1–2). According to standing protocol, all patients were transferred to civilian trauma centers. The median mission time (from call to hospital), airborne medical crew time on ground, and flight time from scene to hospital were 59 minutes (IQR = 42–86), 8 minutes (IQR = 3–21), and 15 minutes (IQR = 10–19), respectively. Fifty-three (59%) missions were performed during daytime and 37 (41%) during nighttime.

Vital Signs

SBP and diastolic blood pressures from point of care (pre-flight to the hospital), as well as HR and shock index, are shown in Table I. The HR tended to decrease during the evacuation (from on-site to hospital), whereas there was no clear trend in SBP or diastolic blood pressure or shock index.

Lifesaving Interventions

A total of 108 LSIs were performed on 61 patients (70.1%) receiving blood transfusion. Tourniquets were the most

FIGURE 1. Anatomical wound distribution.
common LSI applied 24 (27%) times; endotracheal intubation and tube thoracostomy were used 21 (24%) times each; and needle thoracostomy was used 19 (21%) times. Table II shows LSI performed on casualties treated with PRBCs.

**Fluids and PRBC Transfusion**

For 98% (n = 85) of the patients, clinical judgment made by the on-site physician led to early administration of PRBCs. IV crystalloid fluid was given to 48 (55%) patients before PRBC transfusion (median 500 mL, IQR = 300–1,000), and only 2 (2%) patients received 2,000 mL of crystalloid before transfusion of PRBCs was started in accord with the CPG. The median volume of PRBCs transfused to each patient was 500 mL (median = 500 mL, IQR = 300–475). Thirty-seven (43%) patients were transfused with less than one unit of PRBCs (median = 250 mL, IQR = 200–300) and 20 (23%) patients received more than one unit (median = 500 mL, IQR = 500–900) of PRBCs.

**Safety**

There were no documented cases of transfusion reactions, other immediate transfusion-related complications, or any technical problems while administering blood, during flight or emergency department treatment.

**DISCUSSION**

Aeromedical emergency care services have grown significantly during the past 20 years. However, critical evaluations of prehospital blood transfusions during flight missions are limited.\textsuperscript{13–15,17} We sought to contribute to the medical literature by describing our experience with in-flight transfusions.

**TABLE II.** Lifesaving Interventions

<table>
<thead>
<tr>
<th>Intervention</th>
<th>N (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Endotracheal Intubation</td>
<td>21 (24)</td>
</tr>
<tr>
<td>Cricothyroidotomy</td>
<td>1 (1)</td>
</tr>
<tr>
<td>Laryngeal Mask Airway</td>
<td>1 (1)</td>
</tr>
<tr>
<td>Needle Thoracostomy</td>
<td>19 (21)</td>
</tr>
<tr>
<td>Tube Thoracostomy</td>
<td>21 (24)</td>
</tr>
<tr>
<td>Arterial Tourniquet</td>
<td>24 (27)</td>
</tr>
<tr>
<td>Intravascular Access</td>
<td>7 (7)</td>
</tr>
<tr>
<td>Central Intravenous Line</td>
<td>14 (16)</td>
</tr>
<tr>
<td>Total</td>
<td>108</td>
</tr>
</tbody>
</table>

In our study of patients with trauma-related injuries, the severity of illness was considerable, with an overall mortality of 11%. The high number of patients in need of LSIs and the average 2.6 injured body regions per casualty further emphasize their severity of illness.

Over these past two decades, and during the time of our presented experience, there has been a shift from the administration of large volumes of crystalloids to the early administration of blood products for treatment of massive hemorrhage. Accordingly, the casualties in our study were treated when large-volume crystalloid infusion was still common practice. An important finding of our study is that in the vast majority of cases, flight surgeons acted on clinical judgment which led to early administration of PRBCs even before the requisite two liters of crystalloid had been given. The overall effect of this deviation from the CPG cannot be determined from our data, but given recent developments, was most likely beneficial.

Over the 7.5-year window during which we included casualties in this study, we found that an average of only 12 patients per year received PRBCs. Consequently, hundreds of units of PRBCs were likely wasted every year in order to maintain the ability to transfuse PRBCs. This raises the question of whether this type of resource utilization is appropriate because it is utilized so rarely, it is the authors’ opinion that although used so rarely, due to the nature of their missions CSAR units, in Israel and worldwide must maintain the ability to deliver blood products. We also believe that adopting a policy of returning unused PRBCs to hospital blood banks for use before expiration can minimize the waste of PRBCs.\textsuperscript{21}

We have also shown in our study that transfusions were more frequently associated with inter-facility transfer missions rather than on-scene missions, with 15.2% and 5.4% of patients in each group receiving blood transfusions. This may be due to the severity of the injuries sustained by these casualties, which also prompted their transfer to facilities that could offer higher levels of care.

Previous publications by Berns et al\textsuperscript{13} and by Higgins et al\textsuperscript{17} reported smaller percentages of trauma cases, higher percentages (62–91%) of inter-facility usage of blood transfusion, and mortality between 45% and 84%. These differences are likely attributable to the different populations served by our respective aeromedical services, as well as the protocols for blood transfusion to which each adheres.
Berns et al\textsuperscript{13} reported that the hemoglobin level was raised in 75% of the cases and was increased by 1.3 g on arrival to receiving hospital. Higgins et al\textsuperscript{17} confirmed that the hematocrit levels increase and suggested hemodynamic parameter improvement following blood transfusion.

In the current study, short evacuation times did not allow the transfusion of more than two units of red blood cells, and in several instances casualty evacuation was completed before blood transfusion was complete. It seems reasonable to assume that in many evacuations a choice between transfusions of either reconstituted freeze-dried plasma (FDP) or PRBCs will be necessary. Early plasma transfusion is associated with improved survival in hemorrhaging trauma patients,\textsuperscript{24,25} and though it is not currently considered standard initial therapy, some data suggest that plasma should be transfused before the transfusion of red blood cells.\textsuperscript{26} Accordingly, the IDF-MC CPGs were altered to include initial transfusion of FDP, followed by transfusion of PRBCs when appropriate.

Adverse effects of blood transfusion are well known,\textsuperscript{27–29} though complications associated with uncrossmatched O blood administered in the field have not been well documented. Dalton et al\textsuperscript{15} reported a 6-year experience of aeromedical service blood transfusion and concluded that the dangers of blood transfusion have been reduced with screening procedures and modern storage techniques. In his study, 87 trauma patients underwent blood transfusion and the only directly related adverse reaction was a transient episode of self-resolving shortness of breath in a single patient. Recently, Powell-Dunford et al\textsuperscript{30} demonstrated the safe use of PRBCs administered by medevac teams in Afghanistan with 61 cases of blood transfusion without any known instance of adverse reaction or local blood product wastage. Our study further validates the safe transfusion of PRBCs with no recorded short-term adverse reactions among our trauma cohort. Although no safety issues were documented during transfusion, we cannot exclude the possibility that a female casualty with Rh-negative blood that received O-positive PRBCs can suffer from adverse consequences with future pregnancies.

Looking ahead, the IDF-MC has recently started implementing Remote Damage Control Resuscitation principles in the prehospital setting. These include control of compressible hemorrhage, hypotensive-hemostatic resuscitation, transfusion of blood products, and use of tranexamic acid tranexamic acid and FDP for the care of hemorrhaging patients,\textsuperscript{20,22,23} not only during aeromedical evacuations but also at point of injury for ground medical teams.

This study has important limitations. The main limitation concerns the lack of an appropriate control group, making an accurate assessment of the relative effects of PRBC transfusion difficult. Other limitations are its small size, heterogeneity of medical care providers, and that data were collected retrospectively, resulting in missing data. Several endpoints were subjective, potentially resulting in reporting biases. Barton et al\textsuperscript{20} reported that medical personnel may have documented successful responses to justify the use of their treatment; however, such instances are difficult to verify. Finally, this study was performed in a mixed military/civilian environment in Israel, and may not be representative of other trauma systems.

**CONCLUSION**
In our aeromedical experience, transfusions of PRBCs for trauma patients were safe, feasible, and most likely beneficial. PRBCs were administered according to the flight surgeons’ clinical judgment and not in complete adherence to CPGs in most cases. Data collected from this and similar studies worldwide have led to changes in CPGs, shifting from hypertensive resuscitation through hypotensive resuscitation to Remote Damage Control Resuscitation (hypotensive-hemostatic resuscitation).

**REFERENCES**

Slack Reducing Band Improves Combat Application Tourniquet Pressure Profile and Hemorrhage Control Rate

CAPT Dean Nachman, MC IDF; MAJ Avi Benov, MC IDF; MAJ Amiram Shovali, MC IDF; MAJ Yavnai Nirit, MC IDF; CAPT Roy Nadler, MC IDF; LT Yitzhak Avraham, MC IDF; COL Elon Glassberg, MC IDF

ABSTRACT Background: The Combat Application Tourniquet (CAT) is the tourniquet of choice in the Israeli defense forces. Applying the device loosely before windlass twisting is a main pitfall in CAT application. This study objective is to assess the effectiveness of a novel design modification of the CAT, aiming to prevent loose applications, by minimizing the slack. Methods: Using the HapMed leg tourniquet trainer, an above the knee traumatic amputation was simulated. Active duty combatants and Special Forces basic medics were randomly assigned to apply the modified (n = 67) or conventional CAT (n = 65) once. Applied pressure, hemorrhage control status, time to stop the bleeding, and estimated blood volume loss were measured. Results: Using the modified CAT, the mean (±SD) pressure applied was significantly higher compared to the conventional one (231.49 ± 37.84 mm Hg vs. 213.31 ± 45.51 mm Hg, p < 0.05). Hemorrhage control rate was 86.6% in the modified CAT group versus 67.7% in the conventional CAT group (p < 0.05). Analyzing only the applications that succeeded in hemorrhage control, blood loss (171.12 ± 72.43 mL vs. 187.75 ± 91.72 mL, p > 0.05) and time to stop bleeding (27.27 ± 13.15 seconds vs. 27.5 ± 11.25 seconds, p > 0.05) were similar. Conclusions: The modified CAT demonstrated an upgraded pressure profile and hemorrhage control rate, potentially indicating its improved efficacy.

INTRODUCTION
Deaths resulting by injuries that are potentially survivable in the battlefield are mainly the result of hemorrhage, with trauma to the extremities being a major cause.1 Tourniquets have been used since the days of ancient Greeks, some two millennia ago,2 but their usage has been subjected to various controversies concerning their safety and effectiveness ever since.3 In recent years vast research led to profound improvements in tourniquet design and application methods, leading to the introduction of modern tourniquets to the battlefield.4 Substantial experience with tourniquet application have been gained by military personnel in recent military conflicts, demonstrating improvement in mortality rates of soldiers suffering major limb injury, with relatively minimal morbidity.5,6 It has been estimated that tourniquets saved the lives of 1,000 to 2,000 U.S. military personnel during “Operation Enduring Freedom” and “Operation Iraqi Freedom.”9

Recent increase of violence acts in a civilian setting, such as firearms shooting and explosions, highlights the need for implementing insights from treating mass trauma casualties in the battlefield to the civilian sector.10 The Hartford consensus recommendations, issued after several active shooting events, stress the importance of early hemorrhage control in mass casualty events, even by unexperienced laypersons.11 Thus, in addition to the obvious need in battlefield trauma care, it is essential to provide laypersons and novice first responders an effective and intuitive means for hemorrhage control.

One of the most commonly used tourniquets is the 38-mm-wide, windlass and band designed Combat Application Tourniquet (CAT; Composite Resources, Rock Hill, SC). The CAT is the only advanced tourniquet used by the Israel Defense Forces (IDF),12 and is widely issued to deployed U.S. soldiers.13 Despite the importance of hemorrhage control, there is a lack of actual data on tourniquet effectiveness arising from the battlefield. A prospective survey of casualties from the Iraqi campaign demonstrated a success rate of 79% to the CAT, which achieved the highest rate among all prehospital tourniquets,5 whereas a more recent study from the campaign in Afghanistan found that in 83% of the CAT applications distal pulses were present, indicating ineffectiveness.14 As a result of several after-action reports of failed CAT applications in the IDF, we conducted a study (data not published) aiming to identify the specific flawed phase in CAT application, leading to device ineffectiveness. It was found that the leading cause to failure was applying the device loosely, i.e., not pulling the band tight enough before windlass twisting. This finding matches reports by others, who have also demonstrated the importance of slack reduction for an effective application.15,16 Insufficient slack reduction leads to an ineffective windlass twist, not compressing the limb despite further twisting. Furthermore, excess windlass twists might cause the CAT to roll and even brake.4,13

The purpose of the present study was to assess the effectiveness of a novel modification in the design of the CAT, aiming to minimize slack and improve applied pressure.
METHODS

Study Design

This cross-sectional randomized study was granted a waiver by the Institutional Review Board of the IDF Medical Corps and was conducted in accordance with the guidelines of good clinical practice.

Modified CAT

To address the tightening phase insufficiency, we have developed a novel modification of the CAT, by adding a band designed to reduce the slack (Figs. 1A and 1B). The ad on band is attached to the buckle of the CAT, allowing slack to be minimized, by pulling it to the counter direction of the original band after the last pass through the buckle. The rest of the application of the modified CAT was identical to the application of the conventional one. It is of importance to highlight that the modified CAT is an investigational modification, and that it has not been reviewed by any regulatory administration.

HapMed Manikin

The HapMed Leg Tourniquet Trainer (CHI Systems, Fort Washington, PA) is a manikin simulating a bleeding amputation, proximal to the knee of the right thigh. The trainer has the ability to collect performance statistics such as time of application, time to stop bleeding, applied pressure, blood loss volume, and hemorrhage control status. The HapMed manikin was used in the last few years in several studies, examining tourniquet efficiency. This study used injury scenario one, i.e., a casualty with a small build.

Study Population and Setting

Tourniquet testing was performed by a total of 132 participants; all were active duty males, 18–21 years old. Sixty-one participants were combatants during basic military training (combatants group). Seventy-one participants were basic medics from the Special Forces of the IDF. Among them, 39 were 1 week after the completion of combat medic training course (Emergency Medical Technician Basic [EMT-B] beginners group), and 32 had additional 17 weeks of rigorous “in-unit” medical training (EMT-B advanced group). After completion of a short questionnaire regarding previous experience with tourniquet application and estimation of self-efficiency in a scale of 1 (the lowest) to 10 (the highest), the participants were block randomized to either modified (N = 67) or conventional (N = 65) CAT application, according to their previous medical training. Each group received an identical predefined explanation regarding the HapMed manikin and correct application of the modified or conventional CAT, respectively. Following the short explanation, all assigned participants applied the modified or conventional CAT once, as described in section “Study Protocol.”

Study Protocol

The participants stood 10 ft (3 m) away from the HapMed manikin, with a CAT tourniquet held in their hand, so that the buckle of the CAT is free. The manikin was fixed on a table and the participants were given a start notification to begin CAT application. After applying the tourniquet, the participants announced the examiner when to stop the clock. Performance statistics was collected at the end of every application from the electronic controller of the HapMed. Participants were blinded to the specific investigated outcome variables, and were instructed to apply the tourniquet to the best of their ability to stop the bleeding. In addition, all the participants assigned to the modified CAT group, who also had previous experience with a conventional CAT, answered a preference questionnaire. Participants from the conventional CAT group were not requested to
answer the preference questionnaire, since they were not experienced with the modified CAT.

**Sample Size Calculation**

The study was designed to identify a 20 mm Hg increase of applied pressure. A sample size of 126 participants was calculated using WINPEPI (version 11.53, 2015; Brixton Health, Stone Mountain, Georgia), with a power of 80%, significance level of 5%, and SD of 40 mm Hg that was estimated from previous studies we have conducted on similar populations.

**Statistical Analysis**

Descriptive statistics are presented as means and SD. Continuous data were compared using two-tailed Student’s *t* test for independent groups. Pearson’s *χ*² was used to compare categorical data. Correlation was assessed using Pearson’s correlation coefficient. The probabilities of observing chance effects on the dependent variables of interest are presented as *p* values, with significant level set at 0.05.

**RESULTS**

The CAT was applied by all 132 participants. The mean number of previous applications of the CAT (±SD) was 11.88 (±15.14) for all participants, with no statistically significant difference between the conventional and the modified CAT groups (12.62 ± 15.46 vs. 11.09 ± 14.88, respectively, *p* > 0.05). The EMT-B advanced group had significantly more experience with CAT application relative to the combatants and EMT-B beginners group (28.27 ± 5.11, 4.48 ± 2.97, 4.33 ± 0.95, respectively, *p* < 0.0001) (Table I).

Estimation of self-efficiency was evaluated using a scale of 1–10: 1–3 was considered low, 4–7 medium, and 8–10 high. Since only three participants had a self-efficiency estimation in the low group, we united the low and medium groups. A total of 69.7% of all participants had a high self-efficacy estimation, with no statistically significant difference between the conventional and modified CAT groups (72.3% vs. 67.2%, *p* > 0.05) (Table I). Participants with low self-efficacy estimation achieved a pressure of 225.11 ± 39.9 mm Hg, whereas participants with high self-efficacy estimation achieved a pressure of 221.32 ± 44.7 mm Hg, that is to say that no relation was found between self-efficacy estimation and mean applied pressure.

The combined mean applied pressure of all the subgroups using the modified CAT was significantly higher compared to the conventional CAT (231.49 ± 37.84 mm Hg vs. 213.31 ± 45.51 mm Hg, *p* < 0.05) (Fig. 2, Table II). In a subgroup analysis, the mean pressure applied by the conventional CAT was significantly higher in the EMT-B advanced group than the combatants group or the EMT-B beginners group (239.8 ± 46.68, 207.81 ± 41.1, 201.19 ± 44.28 mm Hg, respectively, *p* < 0.05). Furthermore, the mean applied pressure using the modified CAT was significantly higher compared to the conventional CAT in the combatants group (237.17 ± 32.57 vs. 207.81 ± 41.1, *p* < 0.01), and a similar trend was identified in the EMT-B beginners group (223.61 ± 35.2 vs. 201.19 ± 44.28, *p* = 0.09). In contrast to the less trained subgroups, the EMT-B advanced group reached similar pressure values with both the conventional and the modified CAT (239.8 ± 46.68 mm Hg vs. 229.6 ± 49.68 mmHg, respectively). Of importance, when using the modified CAT, participants from both the combatants and the EMT-B Beginners subgroups reached high pressure values, similar to the EMT-B advanced group. That is to say, that insufficient previous training had no ameliorating effect once the modified CAT was used. Seven participants, four from the modified CAT group and three from the conventional CAT, were excluded from the statistical analysis of the applied pressure because of a basic misunderstanding regarding the application of the CAT, such as skipping the windlass twisting phase, leading to 0 mm Hg applied pressure. Previous application number was positively correlated to applied pressure (*p* < 0.01) using the conventional CAT, but there was no such correlation using the modified CAT.

Tourniquet hemorrhage control rate for the modified CAT was significantly higher than with the conventional one (86.6% vs. 67.7%, *p* < 0.05) (Fig. 3, Table II). A similar trend was demonstrated in a subgroup analysis of the

<table>
<thead>
<tr>
<th>Group</th>
<th>No. of Participants</th>
<th>Participants applying the Modified CAT</th>
<th>Previous CAT Applications (Mean ± SD)</th>
<th>High Self-Efficacy Estimation</th>
<th>N (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All participants</td>
<td>132</td>
<td>67 (50.75)</td>
<td>11.88 ± 15.14</td>
<td>92 (69.7)</td>
<td></td>
</tr>
<tr>
<td>Conventional CAT</td>
<td>65</td>
<td>0 (0)</td>
<td>12.62 ± 15.46</td>
<td>47 (72.3)</td>
<td></td>
</tr>
<tr>
<td>Modified CAT</td>
<td>67</td>
<td>67 (100)</td>
<td>11.09 ± 14.88</td>
<td>45 (67.2)</td>
<td></td>
</tr>
<tr>
<td>Combatants</td>
<td>61</td>
<td>32 (52.45)</td>
<td>4.48 ± 2.97</td>
<td>34 (55.7)</td>
<td></td>
</tr>
<tr>
<td>EMT-B Beginners Group</td>
<td>39</td>
<td>18 (46.15)</td>
<td>4.33 ± 0.95</td>
<td>36 (92.3)</td>
<td></td>
</tr>
<tr>
<td>EMT-B Advanced Group</td>
<td>32</td>
<td>17 (53.12)</td>
<td>28.27 ± 5.11</td>
<td>22 (68.7)</td>
<td></td>
</tr>
</tbody>
</table>

CAT, Combat Application Tourniquet; EMT-B B, Emergency Medical Technician-Basic Beginners group; EMT-B A, Emergency Medical Technician-Basic Advanced group, N, Number. *p* < 0.0001 relative to combatants and EMT-B B (Student’s *t* test). *p* < 0.05 relative to combatants and EMT-B A (Pearson’s *χ*² test).
combatants and EMT-B beginners groups, but not in the EMT-B advanced group.

Analyzing only the applications that achieved hemorrhage control, there was no statistically significant difference found between the modified and conventional CAT groups in time to stop bleeding (27.5 ± 11.25 seconds vs. 27.27 ± 13.15 seconds, respectively) and volume of blood loss (187.75 ± 91.72 mL vs. 171.12 ± 72.43 mL, respectively).

All the participants applying the modified CAT (n = 67) completed an additional survey after the application of the modified CAT, regarding their preferred tourniquet, conventional or modified. There was a clear preference of the modified device, with 97% (n = 65) of the participants preferring it over the conventional CAT.

DISCUSSION

Tourniquet application is an essential, lifesaving skill that is taught by many different organizations worldwide. The CAT tourniquet has a track record as a lifesaver in battle trauma; however, its effectiveness is greatly influenced by the previous training and experience of the applicator.18 In the IDF, the CAT is issued to all front line medical personnel and combatants, a heterogenic population with variable training in bleeding control and CAT application.12 In this study, the combatants and EMT-B represent the point-of-injury first responders to battlefield trauma. This less trained population was the one to apply the majority of the tourniquets during recent conflicts, and therefore improving their CAT application effectiveness is a major concern.19,20 This concern will probably be accentuated in future battlefields of urban environments, in which evacuation and advanced medical treatments might be delayed, leaving fellow combatants to provide point-of-injury bleeding control.

In the present study, we tested the effect of a modification of the CAT, which addresses slack reduction, a major drawback in CAT application. Applied pressure was selected as a main measure, since the pressure applied by a tourniquet is a major determinant for effective arterial occlusion, alongside tourniquet width that is fixed and limb girth that varies greatly between soldiers, leaving applied pressure as the only user-dependent variable.21 The modified CAT, in comparison with the conventional CAT, allowed an 18.18 mm Hg improvement in mean applied pressure, almost similar to our predefined goal of an increase of 20 mm Hg. This improvement in applied pressure translates to an 18.9% increase in hemorrhage control rate, in the specific HapMed trainer scenario. A subgroup analysis revealed that the

<table>
<thead>
<tr>
<th>TABLE II. Applied Pressure and Hemorrhage Control Rate by the Conventional and Modified CAT</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Group</strong></td>
</tr>
<tr>
<td>Conventional CAT, Overall</td>
</tr>
<tr>
<td>Modified CAT, Overall</td>
</tr>
<tr>
<td>Combatants Conventional CAT</td>
</tr>
<tr>
<td>Combatants Modified CAT</td>
</tr>
<tr>
<td>EMT-B Beginners Group Conventional CAT</td>
</tr>
<tr>
<td>EMT-B Beginners Group Modified CAT</td>
</tr>
<tr>
<td>EMT-B Advanced Group Conventional CAT</td>
</tr>
<tr>
<td>EMT-B Advanced Group Modified CAT</td>
</tr>
</tbody>
</table>

CAT, Combat Application Tourniquet; EMT-B B, Emergency Medical Technician-Basic Beginners group; EMT-B A, Emergency Medical Technician-Basic Advanced group. N, Number. *p < 0.001 relative to conventional CAT group. †p < 0.01 relative to combatants with conventional CAT group. ‡p < 0.05 relative to combatants and EMT-B B with conventional CAT (Student t test for mean applied pressure, Pearson’s χ² for tourniquet effectiveness).
modified CAT had a substantial positive contribution to the ability of less experienced groups (the combatants and EMT-B beginners) to effectively apply the CAT. Such an effect was not demonstrated among the experienced group (the EMT-B advanced). That is to say, the modified CAT upgraded the pressure applied by combatants and EMT-B beginners to the level of the EMT-B advanced group, compensating for the differences in previous training and the complete lack of experience in the application of the modified CAT.

Surprisingly, while the application of the modified CAT was the first experience of the participants with this device, time to stop bleeding and blood loss were similar in both CAT types. The applications that did not stop the bleeding were excluded from this analysis, because time to stop bleeding is not relevant if bleeding had not stopped, and the volume of the blood loss is greatly influenced by the interaction of the user with the manikin. An application that has not reached hemorrhage control could lead to a very low blood volume loss if the participant decides to stop the application after a very short period of time. These results emphasize that the modification had not compromised the straightforward application flow of the CAT. In addition, a clear preference for the modified CAT over the regular one was shown among the participants who applied it. These results emphasize the potential of this intuitive modification, especially when applied by the less trained first responders, on whom we trust to perform lifesaving point-of-injury bleeding control.

A secondary conclusion arising from our findings is that, interestingly, the estimation of self-efficacy was not associated with the effectiveness of the application. This suggests that major personality confounders influence this variable and that the participants are unaware of their real capabilities. Integration of manikins with performance feedback during the training process might reduce this gap in self-awareness, leading to better training and application performance.

The present study has several limitations. Indeed, using the HapMed manikin to simulate above the knee amputation and bleeding allows for reproducibility, objective readings of physical and physiological parameters, and fewer ethical concerns. On the other hand, applying the CAT on a manikin does not simulate the physical and physiological variability of actual thigh amputations, such as blood loss, which is an extrapolated calculation made on the basis of the time of application and applied pressure. Furthermore, the settings of the study did not include a simulation of the stressful battlefield environment. Our study protocol allowed only one application attempt per participant, although in a real trauma scenario extensive manipulation (and reapplications) of a tourniquet may achieve better results. Another limitation of the present study is that the results are collected from the first application of the modified CAT, without previous training, although such training might have allowed for even better results with the modified CAT. We assumed that the modification we designed for the CAT improves application by slack reduction, but slack was not assessed in order to not interfere with the application process. Being a pilot study, with a few dozen participants in every subgroup, it is possible that additional differences between the groups could be better detected with larger study groups.

**CONCLUSION**

In this study, we have demonstrated the major contribution of a simple modification of the CAT, which significantly improves its hemorrhage control rate and pressure profile, without compromising its intuitive application. This could translate into more lives being saved, especially in the hands of less experienced first responders that without further training can achieve a similar hemorrhage control rate to the most experienced medics. Additional studies are needed to better understand the exact mechanism through which the modification achieves this improvement and the outcomes of further training.

**REFERENCES**


A Sutureless Vascular Closure Device for Emergent Bovine Xenograft Implantation

Joseph F. Knapp, MD, FACC†; Matt Maxwell, MD, FACST‡; Cassidy Briggs, BA*; Erin T. Shaw, DVM, DACVS*; Milton V. Marshall, PhD, DABT*; David Cheung, PhD*

ABSTRACT A novel vascular staple (C-staple) was developed that does not enter the vasculature lumen during anastomoses. The objective of this study was to demonstrate C-staple safety when used with a bovine xenograft and compare efficacy of the C-staple procedure with Anastoclip surgical clips or suturing when used with a bovine xenograft. Eight sheep had an acute comparison between suturing and C-staples using both common carotid arteries. Sixteen sheep had xenograft placement in the left carotid artery, eight with C-staples and eight with Anastoclips in a chronic study. Over 6 months, Doppler ultrasound interrogation of the common carotid arteries was performed. After 6 months, arteries were evaluated histopathologically. Cross-clamp and surgical times were longer in the C-staple group than the suture group, and xenograft implantation times were statistically longer with C-staples than with Anastoclips. After 6 months, C-staple biocompatibility was similar to Anastoclips. Patency and hemodynamics of the bovine xenograft were not statistically different between the two groups. C-staples performed as well as the Anastoclips except for implant times, likely due to delivery system differences. Histological findings and clinical outcomes were no different with the two devices. Further refinements of the C-staple delivery system are necessary before proceeding to clinical trials.

INTRODUCTION

Acute limb ischemia is a surgical emergency, and revascularization time is critical to limb salvage and a successful clinical outcome. A readily available and easily reconstructed and delivered vascular conduit would prove ideal for reestablishing blood flow to injured areas of the body. Our work has focused on shortening intraoperative time to surgical revascularization using an ovine model and aortic grafting. Efforts to develop a reliable vascular conduit have included autologous vein grafts, polymeric materials such as ePTFE, and xenogeneic grafts. Unfortunately, autologous veins are not always reliably available after trauma; when available, they must be surgically harvested, extending the time to ischemic limb revascularization. Use of polymeric materials or xenogeneic grafts could theoretically shorten the time needed to anastomose new vasculature material if autologous vein harvest is not required. To date, polymeric grafts of a caliber appropriate for peripheral limb implantation have not demonstrated reliable patency. Use of xenogeneic grafts has also been problematic because of an immune response to the foreign material. To overcome these problems, a freeze-dried xenogeneic (bovine) arterial graft with little or no antigenicity was developed by the International Heart Institute of Montana Foundation (IHIMF).

We have previously demonstrated the suitability of the freeze-dried bovine vascular graft in sheep when implanted in the carotid artery. After 6 months, patency with the small bore bovine xenograft was equivalent to or superior to autologous veins or ePTFE grafts. Large bore xenografts performed better than autologous vein grafts, with endothelialization occurring after implantation. L+D-Hydro-treated bovine xenografts have also been successfully used in the Blalock–Taussig procedure in children and infants, demonstrating the utility of L+D-Hydro-treated bovine xenografts.

To further reduce the time needed to anastomose the bovine xenograft, a surgical staple (C-staple) was developed that does not enter the lumen when used for closure. The C-staple described in this manuscript was designed to be used with the IHIMF bovine xenograft. The hypotheses to be tested are that the C-staple is superior to Anastoclips for xenograft anastomoses in a chronic study and operating time with C-Staples is as fast or faster than suturing in an acute study.

Continuous (running) suture lines are most commonly used for vascular anastomoses. However, continuous sutures can result in intimal damage or create an anastomotic stricture, lead to flow/shear/turbulence issues, and result in thrombosis and/or intimal hyperplasia. Alternatively, the time to complete a clipped anastomosis compared to using continuous sutures is considerably shorter, as much as 50% or more and translates into a direct reduction in tissue ischemic time during vascular surgery. Interrupted sutures take longer to complete an anastomosis but reduce the potential of creating an anastomotic stricture. Vascular clips/staples, like interrupted sutures, allow a vessel to grow, and vasodilate while minimizing compliance disturbances at the anastomoses.

Staples or clips used for vascular anastomoses can potentially lead to thrombus formation if the clip metal protrudes into the lumen. The C-staples provide vascular closure without penetrating the lumen during anastomoses, as does the

*The International Heart Institute of Montana Foundation, 500 West Broadway Suite 350, Missoula, MT 59802.
†International Heart Institute–Providence St. Patrick Hospital, 500 West Broadway, Suite 350, Missoula, MT 59802.
This work was presented as a poster presentation at the Military Health System Research Symposium, Fort Lauderdale FL, August 17–20, 2015 doi: 10.7205/MILMED-D-16-00080
Anastoclip GC (originally called vascular closure system). The objective of this study was to evaluate the IHIMF C-staples with anastomosis of a bovine vascular graft in the carotid artery of sheep. C-staple biocompatibility and performance were compared directly with the commercially available Anastoclip GC in a chronic study and standard surgical suturing in an acute study. End points were a comparison of implant time between suturing and C-staples in an acute study, as well as biocompatibility and graft patency in a chronic study.

MATERIALS AND METHODS

Animals
Twenty-four-castrated male Targhee sheep weighing an average of 65 or 76 kg were used for acute (8) or chronic (16) studies, respectively. The sheep were maintained outdoors in a pasture or a barn except when receiving xenograft implants and when the grafts were explanted, at which time they were kept indoors at the University of Montana research animal facility (Missoula, Montana). All animal protocols were approved by the University of Montana Institutional Animal Care and Use Committee and the U.S. Army Animal Care and Use Review Office.

Vascular Grafts
Bovine vascular grafts approximately 25 cm in length after processing were prepared by the L+D-hydro process described in U.S. Patent 7,008,763. Xenografts were rehydrated in 0.9% saline containing 200U heparin/mL for an average of 28.19 ± 16.73 minutes (range 15–77 minutes) with occasional agitation and submerged to ensure no air bubbles were trapped in the lumen prior to carotid artery implantation.

Implant Procedure
Animals were randomly assigned to treatment groups, and the same type of closure was used at proximal and distal ends of each xenograft end-to-end anastomosis. Both carotid arteries were used in the acute study, and ~0.5 to 1 cm was excised from each carotid artery before bovine xenograft anastomosis using sutures or C-staples in the primary and contralateral carotids. In the long-term study, ~0.5 to 1 cm of the left carotid artery was excised before 4.0 to 4.9 cm xenograft end-to-end anastomosis with Le Maitre Vascular (Burlington, Massachusetts) Anastoclip GC (Anastoclip), large (1.7 mm) or C-staples. Before final suture, Anastoclip, or C-staple application to the distal anastomosis, the graft was back filled with blood to flush air before completing the anastomoses. Ethicon (Somerville, New Jersey) Surgicel hemostatic sponges were used to arrest minor bleeding. In chronic animals, a Bupivicaine line block (up to 2 mg/kg) was injected into the subcutaneous and muscular tissue for postoperative pain control, and muscular and subcuticular layers were closed with 3-0 Vicryl in a continuous suture pattern.

Graft and Carotid Artery Imaging
Right and left carotid arteries were imaged in all animals by Doppler ultrasound (DUS) to measure inner diameter (ID), peak systolic velocity (PSV), and end diastolic velocity (EDV) using a Vivid e 8L-RS linear array transducer probe of a GE Healthcare (Waukesha, Wisconsin) Vivid-i Ultrasound (GE Healthcare, Pittsburgh, Pennsylvania) with the probe oriented in the direction of carotid arterial blood flow. Two-dimensional DUS images and velocity tracings were obtained at the proximal, mid, and distal left carotid artery/graft complex and the contralateral right carotid artery locations. In long-term animals, DUS interrogation was performed preimplant to obtain a baseline reading, immediately post implant, at 1-week post implant and monthly for 6 months.

Explant Procedure
After final DUS interrogation, all sheep were euthanized with KCl under deep anesthesia. At the 6-month time point, the carotid/graft was photographed and flushed with LRS or saline ex vivo by gravity pressure to remove any remaining blood, followed by 10% neutral buffered formalin (NBF) for fixation; explanted grafts/vessels were stored in neutral buffered formalin. A 5 cm section of right carotid was retained for histopathological comparison.

Histopathology
All anastomoses (two per treated vessel) in the chronic study were processed and embedded in blocks of Spurr’s embedding medium. Blocks were sectioned at one level, sectioned serially twice, and stained with hematoxylin and eosin and elastin trichrome. All mid-graft, proximal, and distal reference samples were processed, embedded in paraffin, sectioned at one level and similarly stained. Contralateral arteries were trimmed at one level (mid-graft section) and processed in paraffin, sectioned, and stained with hematoxylin and eosin and elastin trichrome.

RESULTS

Acute Study
The objective of the acute study was a comparison between the most common type of vascular anastomosis, suturing, with C-Staples. No statistically significant (p < 0.05) difference in cross clamp time was found for xenograft implantation using C-staples (39.00 ± 7.76 minutes; range 27–55) or sutures (33.25 ± 5.72 minutes; range 28–48). Total operative time was also not statistically different (p < 0.05) between xenograft implantation using C-staples (72.75 ± 15.08 minutes; range 54–96) or sutures (66.38 ± 11.19 minutes; range 44–86). All xenografts were patent immediately postoperatively except one animal with sutures that suffered graft failure from tearing. No statistically significant differences in PSV, EDV, or IDs were seen between...
C-staple or suture groups at corresponding anatomic sites (data not shown).

**Chronic Study**

No operative or immediate postoperative mortality occurred in the long-term study. Leaks were occasionally seen at arterial xenograft branches closed with hemoclips during processing and were repaired with sutures during xenograft implantation. One sheep with a xenograft implant using Anastoclips died on the seventh postoperative day due to respiratory distress secondary to hemorrhage from a mid-graft rupture. Representative images of a xenograft implanted with C-staples or Anastoclips (Figure 1A and B) are shown after arterial flow resumption; Figure 1C and D illustrate representative xenografts at explant.

A statistically significant ($p < 0.05$) increase in cross clamp time was seen with C-staples (47.88 ± 12.58 minutes; range 30–74 minutes), compared to Anastoclips (30.50 ± 5.92 minutes; range 19–39). Total operative time was significantly ($p < 0.05$) greater for xenograft implantation using C-staples (92.63 ± 15.98 minutes; range 67–127 minutes) than Anastoclips (73.29 ± 8.71 minutes; range 57–85 minutes).

**C-staple Closures**

The objective of the chronic study was to compare C-Staples with the marketed vascular closure system, Anastoclips. Comparisons were made between the two closure systems in operative time, cross-clamp time, and usability by the surgeon. In addition, biocompatibility and arterial hemodynamics were compared between the two closure systems. After 6 months, five xenografts anastomosed with C-staples were patent based on PSV. Although one xenograft was occluded at 1 week and low flow was reestablished by 6 months, histopathologic evaluation indicated the graft was suboccluded, which is consistent with the observed low flow. Two xenograft implants were occluded after 1 week, and they remained occluded after 6 months; occlusion was confirmed histopathologically. One animal had notable neck swelling suggesting a minor, nonlethal leak after 1 week, and the xenograft was occluded at 1 month; it remained occluded at subsequent examinations; after 6 months, occlusion was confirmed histopathologically.

**Anastoclip Closures**

Four xenografts anastomosed with Anastoclips were patent. One xenograft was occluded in the immediate 1-month Doppler examination, with minimal flow reestablished by 6 months; this graft was considered to be occluded by histopathologic evaluation. Two xenografts were occluded at the 1-week Doppler examination, and they remained occluded at 6 months. One xenograft was occluded at the 4-month Doppler examination, and it remained occluded at 6 months.

**Xenograft Characteristics After Implantation**

PSV and internal diameters of carotid arteries and grafts were compared throughout the study. Data from five xenografts implanted with C-staples (animals S1, S6, S7, S8, and S16) and four xenografts implanted with Anastoclips (animals S3, S4, S9, and S12) were available for flow rate evaluation at the 6-month time point. Other animals were excluded due to nonpatency, and the early postoperative death animal was also excluded. PSVs, EDVs, and IDs were compared over 6 months to evaluate blood flow through
The xenografts were used to determine the effect of the closure system on hemodynamics. Proximal and distal measurements in Figure 2 were obtained from the native artery just outside the anastomotic sites. No statistically significant differences were observed between groups when compared at the same anatomic sites over time (Figure 2); PSV was greatest at mid-graft at all time points. EDV trends mimicked those of PSV (data not shown). Internal diameters at the anastomotic sites were unable to be imaged because of the presence of C-staples or Anastoclips.

As expected and given that rehydrated xenografts were somewhat smaller in diameter than the native carotid, in the early postoperative stages a general trend of increased velocity of blood flow through the smaller diameter vessel was seen, especially in the mid-graft region. Accordingly, PSV increased at the sites of the proximal anastomosis, mid graft, and distal anastomosis, correlating with a decrease in internal diameter of the xenograft compared to the native carotid arteries. Both C-staple and Anastoclips groups exhibited this trend although these data do not show a significant difference in alteration of blood flow velocities as a result of xenograft implantation either between or within groups. After 6 months, the xenografts were generally dilated relative to their implant size, corresponding to a mitigation of flow rate variance due to implantation.

Four animals had sites of distinct differences in flow rates after 6 months: animals S7 and S8 in the C-staple group and animals S4 and S9 in the Anastoclips group had considerably higher PSV and EDV at the distal anastomotic site. Animal S4 had multiple adhesions to the exterior of the graft, and S9 was thought to have a small aneurysm that was confirmed histopathologically. As significant flow accelerations correlated with degree of stenosis, these data suggest these

FIGURE 2. Peak systolic velocity at different areas of the xenograft with different closure types.

FIGURE 3. Inner graft diameters at different areas of the xenograft.
four animals may have had an anastomotic occurrence that affected blood flow. The sites of increased blood flow were similar in both C-staple and Anastoclip groups, indicating that the event causing increased blood flow was likely not related to the type of anastomotic closure used for implantation. For the other five animals, neither anastomotic site had significant flow acceleration, suggesting no specific anastomotic effects on blood flow occurred through the xenograft. As a contralateral control, DUS interrogation of the right carotid arteries was performed at three (proximal, mid, and distal) sites (data not shown). A much greater consistency was seen in the right (untreated) carotid ID, PSV, and EDV than the left carotid that had the xenograft implant.

Mid-graft diameters of patent C-staple and Anastoclip groups tended to increase toward the end of the 6-month period of time (Figure 3). Although this increase was most noticeable in the Anastoclip group, no statistically significant differences between groups were seen. Values from the right (untreated) carotid artery are also included for comparison and show good consistency over time. The contralateral native carotid artery diameter was fairly constant at about

FIGURE 4. Microscopic evaluation of graft anastomoses from animals with C-staples and Anastoclips stained with Elastin Trichrome, illustrating anastomosis or graft patency. Animal S6 (C-staple), (A) distal anastomosis and (B) proximal anastomosis; (C) Animal S3 (Anastoclip), distal anastomosis and (D) proximal anastomosis; (E) Animal S3 mid-graft; and animal S16 (C-staple), distal anastomosis with partial occlusion.
7 mm, whereas the xenograft IDs were smaller, around 5 mm. Thus, results were similar with both closure systems and the bovine xenograft.

**Histopathological Evaluation**

At explant, animal S1 had possible aneurysm formation at a curved area of the graft, which was confirmed histopathologically. The graft in S4 was adherent to surrounding tissues on all sides, and it had an S-shaped curve. The graft in S9 had diffuse, moderate adventitial reaction and aneurysm formation was noted approximately 1 cm proximal to the distal anastomosis, approximately 1.8 cm in diameter. However, no aneurysm was observed histopathologically. S16 was partially occluded. Examples of patent grafts at the anastomotic sites can be seen in Figure 4. Panels A-D in Figure 4 show a patent lumen at both distal and proximal anastomotic sites. A section of mid-graft is shown in panel E, in which a well-developed endothelial layer is present. Panel F shows the distal anastomosis of a partially occluded graft. Mean percent area stenosis were similar in each group: 52% in the C-staple group and 45% in the Anastoclip group; however, these differences were not statistically significant because of the small group sizes. Except as noted above, the grafts were encased in host tissue with little/no evidence of an inflammatory reaction either along the xenograft or at the proximal anastomotic sites (Figure 5). Overall, the minimal to mild inflammatory response to the C-staples or Anastoclips was similar.

Microscopic evaluation of carotid arteries from 17 sheep treated with C-Staples, Anastoclips, or contralateral controls demonstrated the following at 6 months:

- Four of eight grafts with C-staple anastomoses, and three of seven grafts with Anastoclip anastomoses were fully patent. All other test and control grafts were occluded or suboccluded at the anastomoses and/or in the graft.
- All patent grafts and anastomoses showed very advanced or complete healing as evidenced by full endothelialization and organized and mature neointima. Animal S1 was a single exception where preaneurysmal dilation of the graft was associated with nonocclusive moderately extensive mural thrombosis.
- The two anastomotic systems showed optimal local biocompatibility in this model as indicated by a low-grade chronic inflammatory response (foreign body response to the C-staple/clip members).

The C-staple and Anastoclip closures produced similar levels of neointima with comparable average neointima thickness and percent stenosis values at the anastomoses and in the graft.

Lumen area was, on average, slightly greater in the C-staple group compared to the Anastoclip group; however, this variation was not considered to be biologically meaningful in consideration of the high intra-group variability.

No significant differences in clinical chemistry and hematology values were observed between animals implanted with C-staples or Anastoclips throughout the study period.

**DISCUSSION**

Few vascular anastomotic devices have successfully reached commercial development. Most device designs fall into three major design categories: metal clip or staple, retainer, and inserts; many devices are hybrids of two or more categories. Metal clips and staples with “C-” or “U-” shaped geometry offer the advantage of providing some degree of vasodilatation compared to other shaped geometries. Frame/retainers are comprised of a metal or plastic round or oval frame containing spikes or prongs and a retainer containing holes or slots to receive the mating spikes or prongs. Biomechanical properties of the resulting anastomotic site with a frame/retainer are not identical to those of the native blood vessel. The lack of blood vessel compliance produces abnormal
stress that causes adjacent blood vessel tissues to react and hypertrophy can occur, resulting in restenosis. Inserts use a metal/polymer tube or stent used with sutures or mechanical clamps. Like the retainers, inserts severely restrict biomechanical properties of the resulting anastomosis. With the exception of the Anastoclip, surgical staples are designed for joining tissue sheets, organ resections, and large diameter vessels like the colon. The Anastoclip GC is designed for end-to-end and side-to-end anastomoses; C-staples were designed for end-to-end anastomoses. C-Staples will not slip or yield on pressurization of the anastomotic site while avoiding entry of the metal surface within the vessel lumen. The resulting anastomotic site is free of restrictions that impedes the natural pulsatile motion of the blood vessel without metal staple exposure in the vessel lumen after anastomosis as reflected in PSV values over time.

One complication of xenograft implantation in the current study was related to xenograft length. A longer length of graft than the amount of carotid excised was needed to have sufficient graft tissue for eversion at the distal anastomosis. Exacerbating this issue was tension placed on the proximal anastomosis during eversion at the distal site. The xenograft was prone to tearing around the proximal anastomotic clips or staples if too much tension was placed on the graft during distal anastomosis. In some instances, excess graft length led to kinking and subsequent stenosis in the graft after closure, but the significance of this kinking relative to evaluation of the anastomatic device is unclear. Patency within the xenograft in this study was unaffected by closure type, approximately 50% after 6 months, and similar to that seen in a previous study with small diameter L+D-hydro processed bovine xenografts.1

Fewer C-staples than Anastoclips were required for xenograft anastomoses. This could potentially result in a time savings if the C-staples are delivered through a preloaded applicator, like the Anastoclips. Implantations with C-staples took, on average, ~57% more cross clamp time than implantations with Anastoclips. Further, total operative time for implantations using C-staples was, on average, ~26% greater than that for implantations using Anastoclips. This increased time is most likely related to the C-staple applicator and platform system design where C-staples are loaded onto a platform and picked up individually by the applicator. In contrast, Anastoclip applicators were preloaded with clips, a more efficient delivery system that likely contributed to lower cross clamp and total operative times for anastomoses using this device. Anastoclips have consistently shown to require less surgical time for anastomosis,2 and their long-term efficacy has been demonstrated,7,8 due in part to the fact that the Anastoclips do not penetrate the lumen. The bovine xenograft may prove useful for temporary revascularization with better blood infusion, enabling autologous veins or synthetic grafts to be used long-term after growing autologous cells along the graft.9,10 In addition, the decellularized, non-antigenic bovine xenograft could serve as a scaffold for autologous cell development; xenograft recellularization was not affected by vascular closure type.

CONCLUSIONS

Biocompatibility of the C-staples and Anastoclips was similar, with minimal host-response, and neither clip affected neointima formation within the L+D hydro-treated bovine xenograft. Patency was similar to that seen in a previous study with small-bore bovine xenografts in a sheep model.1

Revision of the C-staple delivery system is needed in which staples are premounted in the applicator. This modification would be expected to substantively shorten the surgical times to achieve successful anastomoses, the major shortcoming of the present clip delivery system design.

ACKNOWLEDGMENT

The assistance of Timothy Descamps, BS, CRA is acknowledged. No additional compensation was received for his contributions.

REFERENCES

The ABC’s of Pancreatic Trauma: Airway, Breathing, and Computerized Tomography Scan?

LT Matthew Vasquez, MC USN*; CPT Cassandra Cardarelli, MC USA†; LCDR Jacob Glaser, MC USN‡; Sarah Murthi, MD†; Deborah Stein, MD, MPH†; Thomas Scalea, MD†

ABSTRACT Objectives: Missed pancreatic injury carries significant morbidity. Computerized tomography (CT) imaging is useful, but may lack sensitivity to identify pancreatic injury. New-generation CT scanners should improve sensitivity, but this has not been studied. A previous study published in 2002 evaluating the sensitivity for identifying pancreatic injury with single-slice CT scanners yielded a 68% correlation between operative and CT findings. We aim to study the accuracy of modern CT for diagnosis and grading of pancreatic injury. Methods: All trauma admissions from 2008 and 2012 were retrospectively reviewed. Patients with a pancreatic injury, either on CT or intraoperatively, were included (n = 96). Sensitivity and specificity were calculated using Student’s t test. Results: 48 patients had injuries noted on CT and in the operating room. In this group, 68.8% had CT findings discordant with operative findings. Of these, 78.8% had no injury noted on CT, of which 26.9% required surgical intervention. Seven patients with injury on CT had none identified in the operating room. Based on these results, the sensitivity for CT imaging to identify an injury is 36.4% with a positive predictive value of 68.2%. Conclusions: Our results indicate that despite advances in CT technology, the sensitivity and specificity for identifying pancreatic injury remains low. Although CT scans remain critical in trauma evaluation, awareness of this diagnostic gap is important. Further analysis is required to determine any impact on patient outcomes.

INTRODUCTION Computerized tomography (CT) scanning has become a mainstay in the evaluation of the trauma patient. Advances in the quality and speed of imaging have made CT a critical adjunct to the physical examination in the obtunded but hemodynamically stable patient. As the ABC’s of trauma move towards “airway, breathing, and CT scan,” it is important to recognize that CT imaging still has limitations, and should not be relied completely in the evaluation of the trauma patient. This is particularly true with regard to the diagnosis of pancreatic injury. Pancreatic injury diagnosis by CT historically has poor sensitivity, ranging anywhere from 28% to 68%. Although pancreatic injury in trauma is relatively rare, especially in blunt trauma, missed pancreatic injuries can result in significant morbidity and mortality (20–42% morbidity, and mortality approaching 20%).

The clinical status of the trauma patient, along with high-quality CT imaging, can be used together to determine the management strategies in these patients. An operative vs. nonoperative strategy often depends on the nature of the identified injury, with high-grade injuries (grade 3–5) being addressed operatively and low-grade injuries being managed with further imaging, drainage, or observation (Figs. 1 and 2). Several studies show that nonoperative management may increase hospital stays, dependence on total parenteral nutrition, and increase complications, including pseudocyst, abscess, pancreatic leak, and fistula.

The R Adams Cowley Shock Trauma Center (STC) retrospectively looked at the accuracy of CT diagnosis for blunt pancreatic injury between 1996 and 2000. Operative and CT findings correlated only 68% of the time. This previous work evaluated pancreatic injuries using older CT technology, with a portion of the data set including images from single-slice CT scanners. Since this study, CT technology has advanced, and currently the Trauma Resuscitation Unit uses 64- and 40-slice CT scanners (as of 2008). We performed a retrospective cohort study, using a modern cohort with modern technology, to determine whether advancement in CT imaging has led to improved diagnosis of pancreatic injury. We hypothesize that newer generation, multidetector CT scanners will better identify pancreatic injuries in patients with either penetrating or blunt abdominal trauma.

METHODS With institutional review board approval, the American Association for the Surgery of Trauma (AAST) multi-institutional pancreatic trauma database (AAST Pancreatic Trauma Study) was accessed for retrospective review. The subset of patients from the database treated at the R Adams Cowley STC was
identified. We included a 5-year interval, from March 2008 to December 2012. All patients with a radiographically or operatively identified pancreatic injury were included. Ninety-six patients were identified. After a thorough chart review, patients with pancreatic injury who received CT imaging and no operative exploration \((n = 18)\), and those who went to the operating room (OR) without CT imaging \((n = 30)\) were excluded. This left 48 patients who received both a preoperative CT and surgical exploration, comprising our study group (Fig. 3). The date ranges of the data set were chosen to align with the acquisition of the upgraded 40- and 64-slice CT scanners.

The purpose of this study was to determine sensitivity of CT scans for identifying pancreatic injury, and was not directed at injury management. As such, both hemodynamically stable blunt and penetrating abdominal trauma patients were included. All patients underwent an abdominal CT scan (using the standardized protocol at the Shock Trauma Center, which includes both arterial and venous phase imaging). All patients underwent a CT scan within 6 hours of arrival and none of the patients underwent repeat imaging before the OR. Decision to proceed to the OR was based on clinical judgment of the board-certified trauma/critical care attending managing the patient upon admission.

Patient records were abstracted from the electronic record, using radiology and operative reports, and were sorted by grade of pancreatic injury. Other data extracted included age, injury severity score, mechanism of injury, hospital course, interventions performed (drains, pancreatic resection, etc.), and mortality.

All CT scan interpretations were performed by dedicated trauma radiologists using the standard AAST Pancreas Injury Scale (Table I). Board-certified trauma surgeons used the same scale to describe intraoperative findings. Positive findings included pancreatic laceration, pancreatic transection, edema, and hematoma.

All scans were performed on one of two dedicated trauma CT scanners (64 or 40 slice Philips Brilliance CT, Andover, MA). All records were reviewed and graded by two reviewers (surgical residents, with oversight from a board-certified critical care surgeon) based on radiologic and OR descriptions using the AAST Pancreas Injury Scale. Student t test and chi-square analysis were used for statistical analysis. The specificity and negative predictive values could not be calculated due to the inclusion criterion (all patients had either a positive CT or operative finding—there were no true negatives).

RESULTS
Over the 5-year study period, a total of 26,365 trauma admissions were identified, and the cohort with pancreatic injury were extracted as described in Figure 3. In the study cohort \((n = 48)\), the average age was 41.1 (SD ± 19.2 years.). Thirty-three patients had a blunt mechanism of injury (78.8%). Fourteen had a penetrating injury and there was one crush injury. Penetrating injuries comprised 29.2% of these injuries \((n = 14)\); all were assaults with 57.1% involving gunshot wounds \((n = 8)\).

The mean injury severity score for the cohort was 32.0 (median = 34, interquartile range [IQR] = 20–43). The average number of intensive care unit days was 10.5 (median = 4.8, IQR = 0–15.1) and total hospital days was 17.7 (median = 12.5, IQR = 7.3–22.4). The overall mortality in this cohort was 18.8%, with no mortality within the first 24 hours. All mortalities were included in the study.

Using the AAST pancreas injury scale as the reference standard, only 31.3% of patients had pancreatic injuries on CT that correlated with operative findings \((n = 15)\), whereas 68.8% \((n = 33)\) did not correlate. The reviewers used radiology reports and operative reports to sort injuries according

---

**FIGURE 1.** Grade 3 pancreatic injury. Arrow identifies the grade 3 hematoma.

**FIGURE 2.** Grade 5 pancreatic injury. Arrow indicates disruption and injury at the pancreatic head.
to the AAST injury scale. When category was not explicitly stated, the reviewers and attending surgeon used injury descriptions in the operative report to stratify the injuries according to the AAST injury scale.

We further evaluated the subset of patients with incongruent CT and operative findings. Among them, 78.8% (n = 26) had pancreatic injuries identified in the OR that were reported as normal on CT imaging. These patients went to the OR due to concern for other intra-abdominal solid organ injury or hemorrhage, at the discretion of the attending surgeon. The remaining 21.2% (n = 7) had injuries reported on CT imaging without an injury noted during exploration. Stated in another way, more than half (54.2%, n = 26) of patients had pancreatic injuries that were missed on CT scan. The AAST injury grades of this group included 18 grade 1, 6 grade 2, 1 grade 3, and 1 grade 5 injuries (Fig. 4).

Of the patients with missed pancreatic injury on CT, 26.9% (n = 7) required a pancreatic resection, including 6 distal pancreatectomies and one trauma pancreaticoduodenectomy. The remainder were managed with drains (n = 15) or without any intervention (n = 4). This represents an intervention in 84.6% of patients that would not have occurred if management had been directed by

![FIGURE 3. Study design and criteria.](image-url)

### TABLE I. AAST Pancreas Injury Scale

<table>
<thead>
<tr>
<th>Grade&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Type of Injury</th>
<th>Description of Injury</th>
<th>AIS-90</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Hematoma</td>
<td>Minor Contusion Without Duct Injury</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Laceration</td>
<td>Superficial Laceration Without Duct Injury</td>
<td>2</td>
</tr>
<tr>
<td>II</td>
<td>Hematoma</td>
<td>Major Contusion Without Duct Injury or Tissue Loss</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Laceration</td>
<td>Major Laceration Without Duct Injury or Tissue Loss</td>
<td>3</td>
</tr>
<tr>
<td>III</td>
<td>Laceration</td>
<td>Distal Transection or Parenchymal Injury With Duct Injury</td>
<td>3</td>
</tr>
<tr>
<td>IV</td>
<td>Laceration</td>
<td>Proximal Transection or Parenchymal Injury Involving Ampulla&lt;sup&gt;b&lt;/sup&gt;</td>
<td>4</td>
</tr>
<tr>
<td>V</td>
<td>Laceration</td>
<td>Massive Disruption of Pancreatic Head</td>
<td>5</td>
</tr>
</tbody>
</table>

AIS, abbreviated injury score. This table is available at [http://www.aast.org/library/traumatools/injuryscoringscales.aspx#pancreas](http://www.aast.org/library/traumatools/injuryscoringscales.aspx#pancreas); accessed September 19, 2016. <sup>a</sup>Advance one grade for multiple injuries up to grade III. <sup>b</sup>Proximal pancreas is to the patient’s right of the superior mesenteric vein.
CT findings alone. The overall mortality in this group was 26.9%.

Within the subset of patients with CT scans that were concordant with operative findings (n = 15), 20% (n = 3) required a distal pancreatectomy, 60% (n = 9) were managed with operatively placed drains alone, and 20% (n = 3) had no operative intervention. The overall mortality in this group was 13.3% (p = 0.31). This group included 8 grade 1 and 7 grade 2 injuries. Patients with a correctly diagnosed injury on CT required an operative intervention 80% of the time.

The calculated sensitivity of CT for identifying pancreatic injury in trauma was 36.4% (95% confidence interval [CI] = 22.1–53.1), with a positive predictive value at 68.2% (95% CI = 45.1–86.1). Of note, none of the patients in the cohort underwent preoperative endoscopic retrograde cholangiopancreatography or magnetic resonance cholangiopancreatography for further evaluation of pancreatic injury.

**DISCUSSION**

This study indicates that improvement in CT imaging has not necessarily improved the diagnostic accuracy in detecting pancreatic injury. The overall sensitivity and specificity was 36.4% and 68.2%, respectively, compared with the previous study at STC that showed a sensitivity of 68%. Despite the advances in CT technology, our cohort had missed pancreatic injuries 78.8% of the time. Of these missed injuries, 26.9% required pancreatic resection, whereas 57.7% were managed with drainage only. In total, nearly 85% of the patients with a normal pancreas on CT (missed injuries) had a surgical intervention.

While it can be argued that many of these injuries were low grade (24/26 missed injuries were grade 1 or 2), and thus would safely be managed nonoperatively, we do not know the actual outcomes if nonoperative management had been possible. In this cohort, management in the OR was based on the operative injury identified, and these injuries were not identified on CT. The one grade 5 injury missed on CT bears specific mention. This patient’s imaging was severely compromised by metallic scatter artifact, precluding the accurate identification of this severe injury. Artifact such as this is unlikely to be overcome with any advance in technology, but still remains a true limitation of CT imaging.
Pancreatic trauma is rare and difficult to identify, but carries a high morbidity and mortality if missed. At our high-volume institution, pancreatic trauma comprised only 0.3% of patients. Dreizin et al9 have reported blunt pancreatic trauma incidence anywhere from 3% to 12%, and nearly all of these injuries were associated with other solid organ injuries. 90% of pancreatic injuries had at least three to four coexisting injuries. Although there is a low incidence of injury, these are often severely injured patients, and the complications and morbidity from missed injury can be lethal.

Additionally, nonoperative management of pancreatic injury can have a high morbidity. Beres et al8 looked at nonoperative treatment of high-grade pancreatic injuries (grade 3 or higher). They demonstrated that nonoperative management of high-grade injuries led to longer hospitalizations (27.5 vs. 15.1 days), more days on total parenteral nutrition (21.8 vs. 7.6 days) and a 74% rate of complications (pancreatic pseudocyst being the most common). Wood et al7 evaluated operative vs. nonoperative management of blunt pancreatic trauma and showed an increase in complications in the nonoperative group (73% vs. 21%, p = 0.02). Early operative intervention, if needed, can decrease length of hospital stay, decrease complications, and improve patient outcomes. Being able to identify pancreatic injuries accurately is critical, and surgical decision-making based on CT must be done with caution.

Our study validates and expands upon prior studies. Phelan et al10 compared the sensitivity of 16- to 64-slice CT scanners in detecting pancreatic injury. Both the 64- and 16-slice scanners showed a relatively low sensitivity (47.2% vs. 60.1%; the 64-slice scanner was less sensitive than the 16). These authors concluded that using CT scan as a decision-making tool for the nonoperative management of pancreatic imaging is limited. Previous work at the STC1 reported a sensitivity of 68%, as compared with 36.4% in our cohort.

The current cohort included only patients with a preoperative CT and subsequently went to the OR. If our current criteria was applied to this prior work, only 19 of 40 patients in their cohort would have been included, dropping the sensitivity from 68% to 47.4% (95% CI = 24.45–71.14), statistically similar to the modern cohort. In all, these data suggest that there is no improved pancreatic injury identification with improved CT technology.

Imaging protocols can influence the diagnostic accuracy. Wong et al4 found that the equilibrium-phase CT scan underdiagnosed pancreatic injury 56% of the time, whereas portal venous phase CT imaging was reported as nearly 100% specific, with a negative predictive value of 100%. This study was conducted over a 17-month period and only included 95 patients with blunt abdominal trauma. The STC incorporates a portal venous phase scan as part of its abdominal imaging for trauma. At this high-volume trauma center, our specificity and negative predictive value were not equivalent to the results obtained by Wong et al.

The current study is limited in power by small sample size, despite a high volume of trauma. It is also limited by a high number of low-grade injuries which potentially could be managed nonoperatively. However, even with the low rate and grade of pancreatic injury in trauma, our data still represent a large single institutional cohort evaluating pancreatic injury. We do not have a comparison group of nonoperatively managed patients with pancreatic injury, and therefore cannot compare differences in outcome. We also do not have follow-up data to analyze any possible complications from missed injuries.

Surgeon bias also likely plays a role in outcomes. Historically, this institution has a low threshold for operative management of abdominal trauma. As such, nonoperative therapy for any suspected pancreatic injury is less likely, leading to operative exploration and often leading to, at minimum, placement of a drain. This common strategy will alter the intervention numbers, but should not change the missed pancreatic injury rates of CT. In contrast, we also have highly experienced radiologists, dedicated exclusively to trauma radiology. We believe that pancreatic injuries missed by CT are unlikely to be a gap in radiologist identification, but more likely to be a weakness in the technology for finding early pancreatic injury.

CONCLUSIONS

This study was designed to assess whether advancement in CT technology has improved diagnostic capabilities, not to comment on various management algorithms for pancreatic trauma. Our data suggest that despite advances in CT technology, the diagnostic gap for identifying pancreatic injury remains. Although CT imaging plays a central role in modern trauma management, it is critical that surgeons continue to continue to rely on clinical judgment based on patient status and mechanism of injury, and be mindful of the limitations in CT technology.

REFERENCES

The ABC’s of Pancreatic Trauma


ABSTRACT  Possible traumatic brain injury victims would greatly benefit from a handheld, noninvasive intracranial pressure (ICP) monitoring tool, which a medic could operate in a remote area. Such a device would also benefit the transport of injured soldiers during en route medical care and critical care air transport. This study demonstrates the use of noninvasive blood flow measurements in the eye by ultrasound as a proxy for ICP. ICP was artificially raised in a porcine model and resultant blood flow change in the ophthalmic artery was measured. In addition, the ultrasound transducer itself was used to compress the eye further altering ophthalmic hemodynamics. Blood flow velocities at a range of applied forces and ICP were compared. It was found that 3.25 N of force applied to the cornea was sufficient to produce significant changes in ophthalmic artery blood dynamics regardless of the ICP value. Specifically, the change in resistivity index (RI) and pulsatility index (PI) as force was applied to the cornea correlated with ICP levels. In multiple animal experiments, the magnitude of PI/RI percent change was inversely related to differences in ICP. Force applied to the cornea at baseline ICP resulted in a 15% increase in PI/RI. Results indicate that as ICP increases, the percent change in PI/RI while force is applied decreases. The consistency of data collected indicates that a trend line developed with this data and from similar experiments could be used as a predictive measurement of ICP.

INTRODUCTION

Intracranial pressure (ICP) is the pressure within the cranium, supported by and transmitted through brain tissue and cerebrospinal fluid (CSF). Since CSF surrounds the optic nerve in the subarachnoid space, elevated ICP can damage the optic nerve head. In addition, increased ICP can cause altered blood flow within the ophthalmic artery, and in some cases, the posterior eye globe can flatten. Recent literature also suggests that optic nerve injuries are strongly related to elevated ICP effects on the optic nerve head. Changes in transmucosal pressure across the optic nerve head may play a role in the pathogenesis of optic disk diseases.

A reliable, noninvasive tool for detecting relative changes in ICP would have a dramatic impact on neurointensive care, where invasive measurement (i.e., parenchymal fiberoptic ICP monitor, intracranial transducer, lumbar puncture) is presently the standard of care. Currently, there is no adequate noninvasive standard for ICP. Specifically, such a device is needed for ICP monitoring in remote environments (i.e., forward military locations). Possible traumatic brain injury victims could be quickly diagnosed with a handheld device that medical personnel could operate in a forward medical site. Such a device would also benefit the transport of injured soldiers during en route medical care and critical care air transport.

Transcranial Doppler ultrasonography has been used to monitor changes in neurological health as it provides real-time access to changes in cerebral hemodynamics. Both orbital and intracranial vascular flow patterns, measured using the pulsatility index (PI), have been reported to be affected by ICP. The acoustic axis of the transducer can be aligned nearly parallel with the ophthalmic artery behind the eye globe close to the anterior portion of the optic nerve head to optimize signal strength. Metrics of Doppler spectra have been introduced as a means of unitless measurement and comparison, including the resistivity index (RI) and the PI. Three values of the cardiac cycle are part of these indices: the peak systolic velocity (Vs), the end diastolic velocity (Vd), and the mean cardiac velocity (Vm). Vs and Vd can be seen in Figure 3 where the maximum and minimum velocities in the velocity envelope are circled. Vm is calculated as the time average of the velocity waveform over a single cycle. These 3 metrics can then be used to calculate RI and PI as shown in equations 1 and 2.

\[
RI = \frac{V_s - V_d}{V_s} \quad (1)
\]

\[
PI = \frac{V_s - V_d}{V_m} \quad (2)
\]

Researchers have used methods combining Doppler ultrasonography and ocular blood flow. These methods are a variation of ophthalmodynamometry, which uses ophthalmic artery blood flow dynamics to determine ICP. For example, Ragauskas examined an extracranial portion of the ophthalmic...
artery that travels in the extradural intraorbital space. This portion of the ophthalmic artery is not exposed to ICP because it is not surrounded by CSF. The methodology presented in the present article differs by measuring a single segment of the ophthalmic artery.

The development of a proof-of-concept data collection prototype was described in our previous work. A similar device was used to collect ocular hemodynamic data during in vivo porcine animal experiments presented in this work. The prototype device is equipped with a single-element, nonimaging Doppler ultrasound transducer, which was instrumented with a load cell to measure the amount of force applied to the cornea during the testing. Ocular hemodynamics were monitored while small forces are applied to the cornea. Force was gently applied by the device’s operator. The response in ocular blood flow dynamics is correlated with ICP levels.

This article describes in vivo animal test results to evaluate a methodology for a simple, fast, and reliable noninvasive method to monitor relative changes in ICP for a patient. The methodology could lead to a small, handheld medical device initially focusing on measuring relative changes in ICP instead of calculating an absolute value. Monitoring changes with a small device is a promising step toward ICP monitoring in remote environments.

**METHODOLOGY**

**Animals**

The in vivo data collection was demonstrated under an Institutional Animal Care and Use Committee approved protocol using porcine models at the University of Nebraska Medical Center. The female swine was selected as the most appropriate animal model due to familiarity by the operating surgeons, access to the CSF space, eye, and skull structure as well as the comparable anatomy of the cranial vessels. Eight in vivo experiments were used to demonstrate the proposed methodology’s ability to monitor relative changes in ICP levels.

Eight in vivo porcine animal tests were used to demonstrate the methodology, with the animals weighing between 65 and 75 kg. The first three animal experiments were used to demonstrate a specific procedure to accurately control and manipulate the ICP level of the porcine model. During these experiments, the efficacy of accessing and monitoring the CSF cavity was evaluated before blood flow data were collected for methodology validation.

**Anesthesia and Perioperative Care**

Animals were premedicated with a combination of Telazol (Zoetis, Parsippany, New Jersey), ketamine, and xylazine, given as a single intramuscular shot. An intravenous line was then established in a marginal ear vein to provide supplemental medication (Telazol, ketamine, and xylazine intravenous as needed), and euthanasia solution at the end of the procedure. The larynx was then sprayed with a topical anesthetic and the animal was masked with isoflurane and supplemented with oxygen before intubation. After intubation, the animals were maintained with 1% to 2% isoflurane supplemented with oxygen throughout the procedure. Mechanical ventilation was provided at a rate of 12 to 15 breaths per minute with a tidal volume of 5 to 10 mL/kg.

**Surgery**

After the animals were under general anesthesia and intubated, they were placed prone on the operating room table and kept flat for the duration of the experiment. The operative surfaces were prepped and draped in the usual fashion. The lumbar drain catheter was inserted 10 to 15 cm from the base of the tail in the midline. A 14-gauge Tuohy needle was used to access the lumbar intrathecal cistern and a catheter was placed between 20 and 30 cm of length at the skin. This was then hooked to a bag of saline. Free-flowing saline was confirmed.

Three centimeters anterior to the ridge of bone between the ears was marked. At 1 cm off the midline, a twist drill hole was placed and an ICP monitor (Integra fiberoptic...
ICP monitor) was inserted. Once a waveform with sufficient signal was obtained and confirmed, this was monitored continuously.

**ICP Management**

A saline drip bag was attached to the lumbar catheter and hung in order to equilibrate the ICP of the animal and the pressure associated with the introduction of saline (Figs. 1–2). The saline drip valve could be further opened or raised to increase the pressure in subarachnoid space.

The ICP was incrementally raised at 2 mm Hg intervals during each animal experiment. At every 2 mm Hg of ICP increase, a researcher manually applied force to the cornea using the ultrasound transducer in contact with the cornea. This stepwise increase was repeated until ICP reached 40 mm Hg. The data collection tests at each increment took between 30 and 60 seconds to complete.

**Doppler Measurement**

Blood flow dynamics of the ophthalmic artery were monitored with a single element ultrasound transducer (Doppler BoxX; Compumedics DWL, Singen, Germany) at four sample axial depths along the acoustic axis. Each of these volumes is gated in time, allowing for data at four depths to be collected along the same axis, centered at 32, 38, 44, and 50 mm from the transducer. This method allowed continuous monitoring and simplified data processing. A baseline measurement was taken at each ICP value, lasting 10 to 15 seconds. Pressure was then applied to the cornea with the ultrasound transducer while maintaining signal strength. The applied force was increased until a new steady state in blood flow velocities was reached. Blood flow spectra were recorded before and after the applied force.

**Data Processing**

At each level of ICP, data were processed before and after the push (Fig. 3). Initial and final forces applied were recorded. Gates were selected based on individual signal strength and correlation coefficient with known signals. Signals that did not meet a threshold of power (indicating insufficient Doppler strength within the gate) or signals where the cardiac cycle was not measured completely were removed from further

---

**FIGURE 3.** Velocity envelope of the ophthalmic artery and plot of applied force throughout a single push at an ICP of 28 with enlarged figures on the right.
processing. If no gate had a sufficient before-push and after-push signal for analysis, the trial was removed. Following gate selection, RI and PI were calculated, and percent change of each of the force values and blood flow metrics were calculated. Each of these values was then compared to ICP value, where linear and quadratic regression and mean square error were calculated.

RESULTS
The three initial experiments were used to optimize ICP adjustment protocol. No data were collected to determine a correlation between blood flow dynamics, but procedural insight was gained. Methodological changes were made such as instead of injecting a bolus, an elevated saline drip bag was used. More importantly, throughout the experiments, as the ICP began to reach around 40 mm Hg, blood flow dynamics responded unusually when ocular pressure was applied. At this threshold, maintaining a particular ICP also became drastically more unpredictable. The exact threshold varied between animal. Because of this phenomenon, data presented does not represent ICP values above the normally behaving range for each animal.

Blood flow data were recorded for the final five experiments, and are summarized in Figures 4 to 6. The percentage change from when no force was applied to the cornea to after steady state was reached when force is applied to the cornea was calculated from PI and RI values. These values are shown in Figure 4 with respect to ICP along with an
Noninvasive Intracranial Pressure Monitoring

analysis of variance. Each plot is designated a distinct symbol and represents a different animal trial. The legend displays each trials deviation from the overall mean using the trials’ mean squared error. The linear regression is plotted with a single standard deviation of the data in either direction. It can be seen from these plots that the first three animals behaved similarly. The fourth and fifth plots show a different slope in their individual regression lines (not shown) and deviate from the overall regression plot.

Figure 5 shows the trinomial regression for each animal trial for percentage change of PI and RI. The legend represents the same five trials presented in Figure 4 in the same order. These plots show the predictive trinomial regression curve. Figure 6 illustrates the maximum amount of force data for a single experiment.

DISCUSSION

After the initial three experiments, the unexpected phenomena reported in the results section occurred. To the authors’ knowledge, this is not something previously reported. It is speculated that above around 40 mm Hg, a mechanical threshold of dural tissue is reached. This may create a break somewhere in the CSF space and possibly initiate a secondary physiological mechanism for maintaining ICP in the animal model. This point is only important in relation to experiments of this kind (where ICP is artificially raised), as the normal range for human ICP is between 5 and 15 mm Hg. An ICP value above 15 is often sufficient to require hospitalization for intracranial hypertension, and cases of ICP values over 40 are even rarer. Because of this, it is interesting that the system seemed to break down around 40 mm Hg, but its particular mechanism is beyond the scope of the current study.

Another phenomenon that was noticed, and is important to analysis and interpretation of the data, is the initial success of placing the lumbar catheter and parenchymal bolt seemed to affect the overall outcome of the particular experiment. If both procedures had no complications, the data seemed to be very similar; i.e., all of the data in the first three experiments presented agreed. The fourth and fifth experiments involved complications associated with maintaining an accurate ICP waveform. These experiments involved either multiple lumbar catheter insertions or an adjustment to the parenchymal bolt. This can be seen in the data and may be the reason the fourth and fifth experiments skewed the regression analysis, and had lower MSE values.

It is interesting to note that the percentage change in the PI and RI began around 0.5 at lower ICP values and ended around −0.5. The exact values are less consequential as they could be due to coincidence or biological variability. PI and RI mathematically follow the same trends, so that was expected. The physiological meaning of the two is a subject of study and slight dispute. It is generally accepted that the RI is correlated to arterial resistance downstream the measured location. The location that is measured in this experiment is in the arterial system, and is only slightly upstream (proximal) to the orbit, meaning that the downstream location that RI could be an indicator of may in fact be the venous outflow routes of the orbit, through the facial veins, nasofrontalis, ophthalmic veins, and partially the cavernous sinus. This makes it interesting idea due to the inherent pulsatility in arteries that is much lower in veins. The fact that as PI or RI has a positive trend with added pressure at low ICP, and has a negative trend with added pressure at high ICP (Fig. 4) indicates one of two possibilities. It is conceivable that the physical properties of the arteries (or, as just discussed, veins) are highly influenced by outside pressure. In this scenario, added pressure to the outside of the eye would compress the intraorbital structures including the lamina cribosa, in turn compressing the optic nerve and surrounding tissues. This would create more resistance in the associated vessels. Another explanation to our data is that the relative difference in pressure is the reason for the change in PI and RI. If this is the case, the RI and PI would not be able to be used as proxy measurements for absolute ICP unless the original ICP, and the added force was known. It is more reasonable that RI and PI could be used with a method such as the one presented to monitor a relative change in ICP.

Figure 5 illustrates the results of an analysis of variance. It is shown that a trend line could be created quite easily with a confidence value of greater than 99%. This trend line becomes misrepresentative in an in vivo experiment such as the one presented if almost anything goes wrong, whether that is related to the lumbar catheter or parenchymal bolt placement. This is shown in the fourth and fifth experiments.

Future Directions

More in vivo animal experiments are scheduled. Similar tests to monitor ocular hemodynamics will take place in
addition to the use of optical coherence tomography. Opti-
cal coherence tomography imaging will be used to measure
displacement in the lamina cribrosa as ICP increases. In
addition to the animal testing, human subject testing is also
planned. An Internal Review Board protocol has been
approved to monitor changes in RI and PI in the ophthalmic
artery while force is applied to the cornea. Specifically,
patients will be placed in a head-down tilt position to simu-
late cephalad fluid shifts that could slightly elevate ICP in
healthy patients. Finally, a proposal to collect data on human
subjects suffering from intracranial hypertension has been
submitted. Patients for the proposed study would already
have invasive ICP monitors in place, and the methodology
discussed in this manuscript will be tested on human subjects
with elevated ICP.

ACKNOWLEDGMENTS

This work was funded by NASA Nebraska Space Grant Consortium;
Layman Fund, University of Nebraska Foundation; and University of
Nebraska Medical Center Core Facilities Grant.

REFERENCES

1. Morgan WH, Yu DY, Balaratnasingam C: The role of cerebrospinal
fluid pressure in glaucoma pathophysiology: the dark side of the optic
fluid pressure in glaucoma: a prospective study. Ophthalmology 2010;
4. Jonas JB, Wang N: Association between arterial blood pressure, cere-
brosinal fluid pressure and intraocular pressure in the pathophysiology
233–4.
5. Taibbi G, Cromwell RL, Kapoor KG, Godley BF, Vizzeri G: The effect of
microgravity on ocular structures and visual function: a review. Surv
7. Bhatia A, Grupta AK: Neumonitoring in the intensive care unit. Intra-
cranial pressure and cerebral blood flow monitoring. Intensive Care
8. Miller MM, Chang T, Keating R, Crouch E, Sable C: Blood flow veloc-
ities are reduced in the optic nerve of children with elevated intracranial
9. Pourcelot L: Velocimetrie ultrasonore Doppler Seminaire INSERM,
.nih.gov/pmc/articles/PMC1645777; accessed September 8, 2016.
Piper I, et al: Clinical assessment of noninvasive intracranial pressure
non-invasive absolute intracranial pressure measurement. Aviat Space
13. Ketchem T, Twedt M, Lim D, Bashford G, Hawks JA: Proof-
of-concept prototype for noninvasive ICP monitoring using ocular
024502–024502-4.
et al: Ocular blood flow response to applied force: in vivo tests toward
020911–020911-03.
15. Ashwini C: Comparative anatomy of the circle of Willis in man, cow,
16. de Riva N, Budohoski KP, Smielewski P, Kasprowicz M, Zweifel C,
Steiner LA, et al: Transcranial doppler pulsatility index: what it is and
Anesthesia and Postoperative Respiratory Compromise Following Major Lower Extremity Surgery: Implications for Combat Casualties

Samuel M. Galvagno Jr., DO, PhD*; Jordan Brayanov, PhD†; George Williams, MD‡; Edward E. George, MD, PhD§

ABSTRACT Care of military casualties requires not only assessment of patient, injury, and setting, but also the consequences of care decisions on other organ systems. In contemporary conflicts, pelvic and lower extremity trauma are common injuries, yet the optimal perioperative anesthetic and analgesic care remains unclear. Residual anesthesia and opioids can cause respiratory depression, specifically postoperative respiratory depression and opioid-induced respiratory depression. This observational study quantified and compared the incidences of respiratory depression following general anesthesia (GA) and spinal anesthesia (SA) for lower extremity surgery. Respiratory data were collected from 173 patients receiving either GA (n = 43) or SA (n = 130) via a bioimpedance-based respiratory volume monitor. Patients were further subdivided by postoperative opioid administration. The overall incidence of respiratory depression was significantly higher in the SA group (48/130 vs. 6/43, p = 0.004). These findings suggest that, while SA may be considered the safer alternative, it may in fact introduce confounding factors, which increase the risk of respiratory depression. Ensuring adequate respiratory status is particularly critical for the military population, as combat casualties are often monitored in understaffed environments following surgery. Using an SA strategy instead of GA may not prevent postoperative respiratory depression, and respiratory volume monitor monitoring may be useful to optimize care.

INTRODUCTION Wounding patterns in contemporary warfare have changed, with recent operations resulting in a significant number of military personnel suffering severe perineal injuries, pelvic fractures, and lower extremity amputations.1–5 Pelviperineal injuries have become the “signature” injury pattern in Afghanistan, and these injuries are commonly sustained as the result of blast mechanisms, particularly from improvised explosive devices.1,2,6 Although several reports exist regarding the surgical management of these complex injuries,1,6–8 less is known about how to select the optimal anesthetic technique and how to safely manage these patients throughout the painful perioperative period.

Major orthopedic surgery on the lower extremities has historically involved the use of general anesthesia (GA), in which central nervous system depression is induced via administration of volatile or intravenous anesthetics. With improvements in anesthetic technique and orthopedic surgical procedure over the past 20 years, regional anesthesia has become a viable alternative, typically consisting of a neuraxial blockade, in which drugs are injected into the epidural space (epidural anesthesia) or subarachnoid space (spinal anesthesia [SA]).

This shift toward regional technique has been driven in part by a growing body of research aimed at assessing the comparative effectiveness of general and neuraxial anesthesia in surgical care. In a broad sense, these studies have associated neuraxial anesthesia with improved outcomes and reduced postoperative complications, in both the short term and long term, and across a variety of procedures and patient types.9–14 Particular focus has been given to anesthesia technique in joint replacement surgery. Large-scale randomized trials and meta-analyses of total hip and knee arthroplasty data report advantages for neuraxial anesthesia in mortality, resource utilization and incidence of major complications such as deep vein thrombosis, pulmonary embolism, and myocardial infarction.9,11–13 There have also been suggestions of reduced postoperative pain, opioid consumption, and incidence of opioid-related effects, including nausea and respiratory depression.9,13

A major variable in evaluating the risk and physiological impact of anesthesia is the administration of opioids for the management of pain, both intraoperatively and postoperatively. In combination with the effects of surgical insult and anesthesia itself, which can lead to postoperative respiratory depression (PORD), opioids are known to independently decrease respiratory drive and alter normal rhythm generation, leading to a variant of PORD described as
opioid-induced respiratory depression (OIRD).

Respiratory compromise is a significant patient safety concern in the postanesthesia care unit (PACU), intensive care unit, general hospital floor, and en route aeromedical care environment. OIRD and PORD present a well-recognized threat to respiratory competence after surgery, and can lead to serious adverse events. Management of respiratory status is complicated by the potential manifestation of obstructed breathing and the confounding effects of supplemental oxygen use. When considering military casualties, one must also consider the setting under which surgery is performed and conditions likely postoperatively. Care of military casualties requires not only assessment of patient, injury, and setting, but also of the consequences of care decisions on other organ systems. The optimal perioperative anesthetic and analgesic care remains unclear.

There is a need for adequate, continuous respiratory monitoring in these settings to address these concerns. Current monitoring in nonintubated patients typically consists of pulse oximetry, respiratory rate (RR) monitoring, and, rarely, capnography; however, these offer, at best, late indications of significant respiratory depression.

Recent data show that a noninvasive respiratory volume monitor (RVM) was able to detect respiratory depression in advance of a pulse oximeter recorded True Desaturation event by 12.8 ± 2.8 minutes from the immediately preceding respiratory depressive event. In many cases, there were repeated respiratory depressive events starting an average of 71.4 ± 16.5 minutes before each true desaturation.

In order to objectively identify and quantify the effects of OIRD and PORD after anesthesia and address the new American Society of Anesthesiologists practice guidelines for monitoring patients receiving neuraxial anesthesia, objective measurement of the depth of ventilation is needed, which can currently be accomplished continuously and objectively only by the aforementioned noninvasive RVM, which provides digital volume traces and accurately reports minute ventilation (MV), tidal volume (TV), and RR in nonintubated patients.

Continuous MV measurements provide a direct, real-time assessment of respiratory function, allowing for an objective and quantitative analysis of respiratory depression in the PACU and on the general hospital floor.

Using the RVM, the objective of this study was to directly compare incidence of OIRD and/or PORD in the PACU in orthopedic surgery patients receiving general versus neuraxial anesthesia. We hypothesized that the incidence of respiratory depression in the PACU among patients receiving GA would be significantly higher than in those receiving neuraxial anesthesia, given that both groups are treated under the same pain management protocols postoperatively. We also hypothesized that the results from this investigation would have important implications for the prevention of perioperative complications in service members requiring surgery for pelvi-perineal and complex lower limb injuries.

**METHODS**

**Subjects**

This study was approved by the Partners Institutional Review Board (Boston, Massachusetts). All subjects provided written informed consent before enrolment. Inclusion criteria were English-speaking males and females 18 to 99 years of age undergoing elective orthopedic surgery at Massachusetts General Hospital. Surgeries performed in positions other than lithotomy or supine were excluded, as were subjects undergoing hip arthroplasty in a position other than left lateral (electrodes need to be placed on the chest and right side). All pregnant females were excluded, as well as patients with pacemakers or other electronic implantable devices. Medical history and basic demographic data were obtained from all subjects. Vital signs (BP, HR, RR, and O₂ saturation) were recorded throughout the study, along with the time, dosage, and route of administered medications.

**Primary Protocol**

A bioimpedance-based RVM (RVM, ExSpiron, Respiratory Motion, Inc., Waltham, Massachusetts) was used to collect digital respiratory data from 173 patients undergoing elective total hip or total knee arthroplasty/replacement surgery (mean: 67.3, 44–89 years of age; mean body mass index [BMI]: 30, 19–49 kg/m²). The RVM PadSet electrodes were placed at the sternal notch, xiphoid, and right mid-axillary line. This placement is consistent with implementation protocols, for which strong correlations (0.96 ± 0.16, mean ± 95% confidence interval for regular and erratic breathing) and high accuracy (average MV and TV errors less than 10% and average RR error less than 2%) between RVM and spirometric measurements have been demonstrated.

Data were collected beginning in preoperative holding, continuing throughout surgery and for the length of the PACU stay. MV, TV, and RR measurements were calculated every 5 seconds for the duration of this period from 30-second segments collected in a sliding window. Care providers were blinded to RVM data. Being an observational study, the choice of anesthetic and PACU management were at the sole discretion of the patient, anesthesiology team, and PACU staff. Monitoring was discontinued at PACU discharge.

Intraoperatively, all patients were managed under either GA or neuraxial (spinal) anesthesia (SA), based on patient preference after consultation with clinical staff. Both anesthetic techniques were initiated intraoperatively, immediately before surgery. Patients undergoing GA were given various doses of a paralytic (rocuronium, vecuronium, or cisatracurium), in conjunction with sedatives (midazolam and propofol), and opioids (fentanyl and hydromorphone). Patients undergoing SA almost exclusively received an

<table>
<thead>
<tr>
<th>General Anesthesia</th>
<th>43</th>
<th>Mean Dose</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paralytic (One):</td>
<td>37</td>
<td>0.86</td>
<td></td>
</tr>
<tr>
<td>Rocuronium</td>
<td>22</td>
<td>0.512</td>
<td>75.0 mg</td>
</tr>
<tr>
<td>Cisatracurium</td>
<td>8</td>
<td>0.186</td>
<td>13.3 mg</td>
</tr>
<tr>
<td>Vecuronium</td>
<td>7</td>
<td>0.163</td>
<td>10.4 mg</td>
</tr>
<tr>
<td>LMA Insertion (No Paralytic)</td>
<td>4</td>
<td>0.093</td>
<td></td>
</tr>
<tr>
<td>Paralytic Not Specified</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inhalation Agent (One):</td>
<td>40</td>
<td>0.930</td>
<td></td>
</tr>
<tr>
<td>Sevoflurane</td>
<td>33</td>
<td>0.767</td>
<td></td>
</tr>
<tr>
<td>Isoflurane</td>
<td>7</td>
<td>0.163</td>
<td></td>
</tr>
<tr>
<td>Reversal Agent:</td>
<td>31</td>
<td>0.721</td>
<td></td>
</tr>
<tr>
<td>Neostigmine</td>
<td>31</td>
<td>0.721</td>
<td>3.10 mg</td>
</tr>
<tr>
<td>Femoral Block:</td>
<td>14</td>
<td>0.326</td>
<td>20 mL</td>
</tr>
<tr>
<td>Ropivacaine 0.2%</td>
<td>12</td>
<td>0.279</td>
<td>20 mL</td>
</tr>
<tr>
<td>Bupivacaine 0.25%</td>
<td>2</td>
<td>0.047</td>
<td></td>
</tr>
<tr>
<td>Sedatives (One or More):</td>
<td>43</td>
<td>0.907</td>
<td>2.23 mg</td>
</tr>
<tr>
<td>Midazolam</td>
<td>39</td>
<td>0.907</td>
<td>2.23 mg</td>
</tr>
<tr>
<td>Propofol (Total)</td>
<td>43</td>
<td>1.000</td>
<td>289 mg</td>
</tr>
<tr>
<td>Bolus</td>
<td>43</td>
<td>1.000</td>
<td>212 mg</td>
</tr>
<tr>
<td>Infusion</td>
<td>3</td>
<td>0.070</td>
<td>1098 mg</td>
</tr>
<tr>
<td>Ketamine</td>
<td>2</td>
<td>0.047</td>
<td></td>
</tr>
<tr>
<td>Opioids (One or More):</td>
<td>43</td>
<td>0.907</td>
<td>2.23 mg</td>
</tr>
<tr>
<td>Fentanyl</td>
<td>43</td>
<td>0.907</td>
<td>221 mcg</td>
</tr>
<tr>
<td>Hydromorphone</td>
<td>38</td>
<td>0.884</td>
<td>1.29 mg</td>
</tr>
<tr>
<td>Meperidine</td>
<td>1</td>
<td>0.023</td>
<td></td>
</tr>
<tr>
<td>Remifentanil</td>
<td>1</td>
<td>0.023</td>
<td></td>
</tr>
<tr>
<td>Morphine</td>
<td>1</td>
<td>0.023</td>
<td></td>
</tr>
<tr>
<td>Other:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Haloperidol</td>
<td>20</td>
<td>0.465</td>
<td>1.1 mg</td>
</tr>
<tr>
<td>Succinylcholine</td>
<td>13</td>
<td>0.302</td>
<td>100 mg</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Spinal Anesthesia</th>
<th>130</th>
<th>Mean Dose</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spinal:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bupivacaine 0.5% (ITHEC)</td>
<td>125</td>
<td>0.962</td>
<td>3.07 mL</td>
</tr>
<tr>
<td>Bupivacaine 0.75% (ITHEC)</td>
<td>2</td>
<td>0.015</td>
<td></td>
</tr>
<tr>
<td>Spinal Dose Not Specified</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Inhalation Agent Used</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Reversal Agents Used</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Femoral Block:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ropivacaine 0.2%</td>
<td>50</td>
<td>0.385</td>
<td>20 mL</td>
</tr>
<tr>
<td>Bupivacaine 0.25%</td>
<td>19</td>
<td>0.146</td>
<td>19.7 mL</td>
</tr>
<tr>
<td>Sedatives (One or More):</td>
<td>127</td>
<td>0.977</td>
<td></td>
</tr>
<tr>
<td>Midazolam</td>
<td>123</td>
<td>0.946</td>
<td>2.59 mg</td>
</tr>
<tr>
<td>Propofol (Total)</td>
<td>87</td>
<td>0.669</td>
<td>290 mg</td>
</tr>
<tr>
<td>Bolus</td>
<td>26</td>
<td>0.2</td>
<td>31.9 mg</td>
</tr>
<tr>
<td>Infusion</td>
<td>80</td>
<td>0.615</td>
<td>301 mg</td>
</tr>
<tr>
<td>Ketamine</td>
<td>1</td>
<td>0.008</td>
<td></td>
</tr>
<tr>
<td>Opioids (One or More)</td>
<td>114</td>
<td>0.877</td>
<td></td>
</tr>
<tr>
<td>Fentanyl</td>
<td>113</td>
<td>0.869</td>
<td>98.8 mcg</td>
</tr>
<tr>
<td>Hydromorphone</td>
<td>11</td>
<td>0.885</td>
<td>0.45 mg</td>
</tr>
<tr>
<td>Meperidine</td>
<td>2</td>
<td>0.015</td>
<td></td>
</tr>
<tr>
<td>Dexametomidine</td>
<td>1</td>
<td>0.008</td>
<td></td>
</tr>
<tr>
<td>Other:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Haloperidol</td>
<td>10</td>
<td>0.077</td>
<td>2.55 mg</td>
</tr>
<tr>
<td>Succinylcholine</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>
based on the DuBois formula for body surface area\textsuperscript{24} was the PACU in a patient-specific respiratory depression in marginal MV and defined as a period of at least 2 minutes within the 15 minutes following an initial opioid dose in the PACU. A standard formula based on the DuBois formula for body surface area\textsuperscript{24} was used to calculate a patient’s predicted MV (MV\textsubscript{PRED}), sufficient to maintain blood oxygen, and carbon dioxide levels under baseline conditions. Previous work suggests that MV below 80% of MV\textsubscript{PRED} places a patient at risk for a negative response to additional opioids, and identifies MV below 40% of MV\textsubscript{PRED} as low MV (LMV).\textsuperscript{25} This is similar to the ARDSnet volume criteria for extubation,\textsuperscript{26} and was considered a potential threat to patient safety if sustained. MV, as reported continuously by the RVM (MV\textsubscript{MEASURED}), was calculated as a percentage of the patient’s MV\textsubscript{PRED} (MV\textsubscript{MEASURED}/MV\textsubscript{PRED} \times 100\%) and used to assess adequacy of ventilation according to this rubric.

To begin to define clinically relevant thresholds for patient evaluation, we examined significant drops in MV below the LMV threshold under a variety of patient situations. Specifically, among patients receiving postoperative opioids, those showing LMV (MV \textless 40\% \textsubscript{MV\textsubscript{PRED}}) sustained for at least 2 minutes within the 15 minutes following an initial PCA opioid dose were defined as experiencing OIRD.\textsuperscript{25} Patients receiving no postoperative opioids but showing LMV sustained over 10 minutes at any point in their PACU stay were considered to be experiencing PORD.\textsuperscript{25} Finally, in order to evaluate patient respiratory status at the time of PACU discharge, patients showing LMV for at least 10 minutes during the last 30 minutes before discharge were designated as experiencing a LMV at discharge, regardless of whether opioids were administered.\textsuperscript{25} This designation served as an assessment of patient respiratory safety at discharge from the PACU. Patients with LMV at discharge may require RVM or other monitoring as they are transferred to the floor. All terms and their abbreviations are summarized in Table II.

A posthoc analysis examined incidence of postoperative apnea (POA) in the cohort. Apnea was defined as a period of at least 10 seconds with no detected breaths. Hypopnea was defined as a period of at least 30 seconds with greater than 50\% reduction in TV from the baseline TV observed during ventilator synchronization. Patients were classified as demonstrating POA if they averaged more than five apneic or hypopneic events per hour over their entire PACU stay.

Statistical analyses
The main goal of this study was to compare incidence of respiratory depression between patients receiving GA (GA, \(n = 43\), 25\% of the cohort) and those receiving SA (SA, \(n = 130\), 75\% of the cohort). We hypothesized that the incidence of respiratory depression would be significantly higher among GA patients. Given that opioid use in postoperative pain management has been associated with an increased risk of respiratory depression, it was necessary to account for systematic differences in opioid intake between the GA and SA groups. To accomplish this, patients in both the GA and SA groups were separated according to whether or not they received opioids in the PACU. Analyses were performed in the subgroups (nonopioid and opioid) as well as across the combined cohort.

The incidences of OIRD, PORD, LMV at discharge, and POA across different groups were compared using two-tailed Fisher’s Exact Test with the hypothesis that there would be a significant difference in incidence across groups.

**TABLE II.** Definitions of Terms and Their Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>SA</td>
<td>Spinal Anesthesia</td>
<td>Delivery of local anesthetic via fine needle into the subarachnoid space</td>
</tr>
<tr>
<td>GA</td>
<td>General Anesthesia</td>
<td>Delivery of anesthetics both intravenously and via inhalation</td>
</tr>
<tr>
<td>MV\textsubscript{MEASURED}</td>
<td>Measured Minute Ventilation</td>
<td>The real-time MV reported by the RVM</td>
</tr>
<tr>
<td>MV\textsubscript{PRED}</td>
<td>Predicted Minute Ventilation</td>
<td>Expected MV Under Baseline Conditions of Quiet Respiration in awake, non-intubated patients</td>
</tr>
<tr>
<td>(% \text{MV\textsubscript{PRED}})</td>
<td>Percent Predicted Minute Ventilation</td>
<td>In percentage, the degree of deviation MV\textsubscript{MEASURED} is from MV\textsubscript{PRED}</td>
</tr>
<tr>
<td>LMV</td>
<td>Low Minute Ventilation</td>
<td>MV &lt;40% MV\textsubscript{PRED}</td>
</tr>
<tr>
<td>OIRD</td>
<td>Opioid-Induced Respiratory Depression</td>
<td>LMV sustained for at least 2 minutes within the 15 minutes following an initial opioid dose in the PACU</td>
</tr>
<tr>
<td>PORD</td>
<td>Post-Operative Respiratory Depression</td>
<td>LMV sustained over 10 minutes at any point in the PACU in patients receiving no post-operative opioids</td>
</tr>
<tr>
<td>POA</td>
<td>Low Minute Ventilation at Discharge</td>
<td>LMV for at least 10 minutes during the last 30 minutes prior to discharge</td>
</tr>
<tr>
<td>Hypopnea</td>
<td>PostOperative Apnea</td>
<td>More than five apneic or hypopneic events per hour over their entire PACU stay</td>
</tr>
<tr>
<td>Apnea</td>
<td>A period of at least ten seconds with no detected breaths</td>
<td></td>
</tr>
</tbody>
</table>

intrathecal dose of bupivacaine 0.5\% (1–4 mL), typically supplemented with midazolam, propofol, and fentanyl, but in lower dosages than those used during GA. Additional intraoperative opioids like hydromorphone were rarely used in spinal cases. Patients undergoing knee surgery typically also received a femoral nerve block, consisting of either 20 mL ropivacaine 0.2\% or 20 mL bupivacaine 0.25\%, administered in preoperative holding. A detailed summary of relevant medications used intraoperatively, with frequency and dosing, is reported in Table I.

For the purposes of this study, we established thresholds for marginal MV and defined respiratory depression in the PACU in a patient-specific manner. A standard formula based on the DuBois formula for body surface area\textsuperscript{24} was used to calculate a patient’s predicted MV (MV\textsubscript{PRED}), sufficient to maintain blood oxygen, and carbon dioxide levels under baseline conditions. Previous work suggests that MV below 80\% of MV\textsubscript{PRED} places a patient at risk for a negative response to additional opioids, and identifies MV below 40\% of MV\textsubscript{PRED} as low MV (LMV).\textsuperscript{25} This is similar to the ARDSnet volume criteria for extubation,\textsuperscript{26} and was considered a potential threat to patient safety if sustained. MV, as reported continuously by the RVM (MV\textsubscript{MEASURED}), was calculated as a percentage of the patient’s MV\textsubscript{PRED} (MV\textsubscript{MEASURED}/MV\textsubscript{PRED} \times 100\%) and used to assess adequacy of ventilation according to this rubric.

To begin to define clinically relevant thresholds for patient evaluation, we examined significant drops in MV below the LMV threshold under a variety of patient situations. Specifically, among patients receiving postoperative opioids, those showing LMV (MV <40\% MV\textsubscript{PRED}) sustained for at least 2 minutes within the 15 minutes following an initial PCA opioid dose were defined as experiencing OIRD.\textsuperscript{25} Patients receiving no postoperative opioids but showing LMV sustained over 10 minutes at any point in their PACU stay were considered to be experiencing PORD.\textsuperscript{25} Finally, in order to evaluate patient respiratory status at the time of PACU discharge, patients showing LMV for at least 10 minutes during the last 30 minutes before discharge were designated as experiencing a LMV at discharge, regardless of whether opioids were administered.\textsuperscript{25} This designation served as an assessment of patient respiratory safety at discharge from the PACU. Patients with LMV at discharge may require RVM or other monitoring as they are transferred to the floor. All terms and their abbreviations are summarized in Table II.

A posthoc analysis examined incidence of postoperative apnea (POA) in the cohort. Apnea was defined as a period of at least 10 seconds with no detected breaths. Hypopnea was defined as a period of at least 30 seconds with greater than 50\% reduction in TV from the baseline TV observed during ventilator synchronization. Patients were classified as demonstrating POA if they averaged more than five apneic or hypopneic events per hour over their entire PACU stay.

Statistical analyses
The main goal of this study was to compare incidence of respiratory depression between patients receiving GA (GA, \(n = 43\), 25\% of the cohort) and those receiving SA (SA, \(n = 130\), 75\% of the cohort). We hypothesized that the incidence of respiratory depression would be significantly higher among GA patients. Given that opioid use in postoperative pain management has been associated with an increased risk of respiratory depression, it was necessary to account for systematic differences in opioid intake between the GA and SA groups. To accomplish this, patients in both the GA and SA groups were separated according to whether or not they received opioids in the PACU. Analyses were performed in the subgroups (nonopioid and opioid) as well as across the combined cohort.

The incidences of OIRD, PORD, LMV at discharge, and POA across different groups were compared using two-tailed Fisher’s Exact Test with the hypothesis that there would be a significant difference in incidence across groups.
The incidences across combined data from multiple groups were also compared with a two-tailed Fisher Exact Test, which was possible given that the groups were independent (nonoverlapping). Multi-factor analysis of variance was used to evaluate the effects of demographics (age, height, weight, BMI, and sex) across groups. All analyses were performed in Matlab 2012b (Mathworks, Natick, Massachusetts) and results were considered significant at $p < 0.05$.

RESULTS

Opioid-Induced Respiratory Depression

Of the 173 patients, 92 received opioids in the PACU (53.2%, Table I). Consistent with previous work,28,29 a significantly higher fraction of GA patients received opioids in the PACU than SA patients (GA: 30 of 43, 69.8%; SA: 62 of 130, 47.7%, $p < 0.05$, one-tailed Fisher’s exact test). Among patients who received PACU opioids, the mean morphine milligram equivalent received in the PACU was significantly higher in GA patients compared to SA patients (GA: 7.68 ± 0.98; SA: 3.80 ± 0.54, $p < 0.001$).

Of the 92 patients receiving opioids in the PACU, 30 demonstrated OIRD (32.6%). These 30 patients comprised 24 of the 62 SA patients receiving opioids (SA$^{\text{opioid}}$) and 6 of the 30 GA patients receiving opioids (GA$^{\text{opioid}}$). Thus, despite lower opioid intake on average, SA$^{\text{opioid}}$ patients showed a slightly higher incidence of OIRD than GA$^{\text{opioid}}$ patients (SA: 38.7%, GA: 20.0%, $p = 0.098$, two-tailed Fisher’s Exact Test), as shown in Figure 1A. No significant difference was observed in demographics (age, height, weight, BMI, and sex) between the SA$^{\text{opioid}}$ and GA$^{\text{opioid}}$ subgroups ($p > 0.05$ for all factors).

Postoperative Respiratory Depression

In patients who did not receive opioids, we used sustained LMV to define PORD. Incidence of PORD in patients not receiving opioids resembled that of OIRD in patients receiving opioids in both the collective cohort and in distribution across the GA/SA subgroups. Of 81 patients not receiving opioids in the PACU, 24 demonstrated PORD (29.6%) at some time during their PACU stay. This group comprised 24 of the 68 SA nonopioid patients (SA$^{\text{nonopioid}}$) and 0 of the 13 GA nonopioid patients (GA$^{\text{nonopioid}}$). Incidence of PORD in SA$^{\text{nonopioid}}$ patients was significantly higher than that in GA$^{\text{nonopioid}}$ patients (SA: 35.3%, GA: 0.0%, $p = 0.008$), as shown in Figure 1A. Once again, there was no significant difference in demographics between the two sub-groups ($p > 0.05$ for all factors).

Combining the 2 independent opioid-stratified subpopulations allowed us to compare the incidence of either form of respiratory depression (OIRD or PORD) in the PACU, regardless of opioid administration. We found that in the combined SA group (i.e., SA$^{\text{opioid}}$ and SA$^{\text{nonopioid}}$), incidence of respiratory depression was significantly higher than in the combined GA group (GA$^{\text{opioid}}$ and GA$^{\text{nonopioid}}$), as shown in Figure 1B (SA: 48/130, 36.9%, GA: 6/43, 14.0%, $p = 0.004$).

Assessment of MV at Discharge

Patient ventilation status was assessed at PACU discharge, with patients showing LMV for more than 10 minutes of the 30 minutes before discharge designated as having an LMV discharge. This designation is intended to signify a greater risk for respiratory complications in a subsequent unit that would likely have reduced patient monitoring. Cohorts were initially subdivided according to whether or not opioids were administered postoperatively. Of 92 opioid patients, 13 patients showed LMV at discharge (14.1%). These 13 comprised 5/30 GA$^{\text{opioid}}$ patients and 8/62 SA$^{\text{opioid}}$ patients. Thus, among patients receiving opioids, the frequency of LMV discharge was marginally higher in GA patients (GA: 16.7%, SA: 12.9%, $p = 0.751$), as shown in Figure 2A.

Among nonopioid patients, 11 of 81 patients displayed LMV at discharge (13.6%). All 11 of these were among the 68 SA$^{\text{nonopioid}}$ patients. None of the 13 GA$^{\text{nonopioid}}$ patients were found to exhibit LMV at discharge. Thus, among nonopioid patients, the frequency of LMV at discharge was marginally higher in SA patients (GA: 0.0%, SA: 16.2%, $p = 0.197$), as shown in Figure 2A.

As with the respiratory depression analysis, LMV at discharge was also examined across the combined cohort including opioid and nonopioid patients. We found comparable rates of LMV at discharge between the combined SA group (i.e., SA$^{\text{opioid}}$ and SA$^{\text{nonopioid}}$) and the combined GA group (GA$^{\text{opioid}}$ and GA$^{\text{nonopioid}}$), with marginally higher rates in SA patients, as shown in Figure 2B (SA: 14.6%, 19/130, GA: 11.6%, 5/43, $p = 0.800$).

Postoperative Apnea

POA was found significantly more frequently in patients receiving opioids, with 25 of 92 patients receiving PACU opioids showing POA versus 12 of 81 nonopioid patients (Figure 3, 27.2% vs 14.8%, $p < 0.05$). This suggests that opioid analgesia may be a risk factor for apnea. Further analysis revealed that the difference in POA incidence was driven primarily by differences in GA patients, and in particular, by an exceptionally high incidence rate among GA patients receiving opioids. Of 30 GA$^{\text{opioid}}$ patients 13 demonstrated POA in the PACU, as compared to 12 of 62 SA$^{\text{opioid}}$ patients (43.3% vs 19.4%, $p = 0.0237$). For comparison 0 of 13 GA$^{\text{nonopioid}}$ patients showed POA, as compared to 12 of 68 SA$^{\text{nonopioid}}$ patients (0% vs 17.6%, $p = 0.198$).

DISCUSSION

The distribution of injuries in contemporary warfare has changed compared to that seen in previous conventional conflicts.3,4,6,8,30 New wounding patterns, including a greater proportion of severe pelvi-perineal and lower extremity

82

MILITARY MEDICINE, Vol. 182, March/April Supplement 2017
musculoskeletal injuries, have led joint medical providers to consider novel methods for ensuring the safety of combat casualties throughout the continuum of care. This work expands upon previous studies that have shown that respiratory depression, unsafe respiratory patterns, and risk for hypoxemia may be detected early with a noninvasive RVM. Although the RVM has not yet been implemented directly in the military environment, the results from this study population suggest that developing a protocol for generic use of SA may not have the safety advantage anticipated for patients undergoing lower body surgery due to traumatic wounds. The study also strongly suggests a role for the RVM as a means to individualize care regardless of anesthetic technique. The RVM data permits rapid recognition and management of postoperative complications in the contemporary combat casualty environment, improving overall outcomes.

The consensus in recent years is that neuraxial anesthesia is by and large a safer and preferred alternative to GA. There is no context for which this is more accepted than major orthopedic surgery. In particular, major surgeries of the hip and knee have been the focus of a series of large-scale trials and meta-analyses, with the broad conclusion that SA is associated with reduced mortality and reduced complications postoperatively. However, recently published practice guidelines by the American Society of Anesthesiologists note that, while safer overall, neuraxial anesthesia may place patients at additional risk for OIRD. Therefore, it is recommended that all patients receiving neuraxial anesthesia should be monitored for “adequacy of ventilation (e.g., RR, depth of respiration [assessed without disturbing a sleeping patient]).”

As discussed at length by Gulur et al., the limitations of such expansive research studies are numerous and inevitable. Results tend to be conflicting, with similar large-scale studies reporting no statistical difference in outcomes from general versus neuraxial anesthesia. Surgical outcomes are also influenced by a wide variety of confounding factors, which can often go unaccounted for in such work. These

FIGURE 1. Postoperative respiratory compromise in patients receiving joint replacement surgery: general and spinal anesthesia. (A) A novel, bioimpedance-based respiratory volume monitor (RVM) was used to track minute ventilation (MV) during and after surgery in patients presenting for total hip or knee arthroplasty. Patients receiving opioids in the post-anesthesia care unit (PACU) and showing minute ventilation (MV) below 40% of predicted minute ventilation (MVPRED)- for more than 2 minutes within the 15 minutes following initial opioid administration were defined as experiencing opioid-induced respiratory depression (OIRD). Patients not receiving opioids and showing MV below 40% MVPRED over a period of 10 minutes in the PACU were defined as experiencing post-operative respiratory depression (PORD). OIRD (in opioid patients) showed marginally higher incidence in patients receiving spinal anesthesia. PORD (in nonopioid patients) showed significantly higher incidence in patients receiving spinal anesthesia (p = 0.008). (B) A combined analysis looked at incidence of either form of respiratory depression in the PACU across both opioid and nonopioid patients. Patients receiving spinal anesthesia were found to have significantly higher incidence of respiratory depression than patients receiving general anesthesia (p = 0.004).

FIGURE 2. Low minute ventilation (LMV) at post-anesthesia care unit (PACU) discharge in patients receiving general anesthesia versus spinal anesthesia. Patients showing MV below 40% MVPRED for 10 minutes during their final 30 minutes in the PACU were defined as exhibiting LMV at discharge. (A) In the cohort receiving postoperative opioids, frequency of LMV discharge was found to be marginally high in general anesthesia (GA) patients than spinal anesthesia (SA) patients, but not statistically different. The opposite was true of the nonopioid cohort. Neither result was statistically significant (Fisher’s Exact Test, p > 0.05). Zero of 13 GA patients receiving no postoperative opioids exhibited LMV at discharge. (B) A combined analysis of opioid and nonopioid patients showed comparable frequency of LMV discharge between GA and SA cohorts.
may largely explain the observed gaps in mortality and morbidity, which appear to be closing as standards of care continue to improve.\(^1\) Meta-analyses that reach back to cases 10 and 15 years ago to support their conclusions may well be misrepresenting comparative outcomes under current practice. Lastly, it is valuable to recognize that in the context of a specific surgical protocol and individual patient risk factors, 1 anesthetic technique can present clear advantages over the other. To reduce the question of technique to a sweeping, binary conclusion is perhaps to discount the complexity of anesthesia.

Our findings in the present study speak to this point. We sought to evaluate the incidence of respiratory depression in the PACU in relation to primary anesthetic technique, general or neuraxial. To accomplish this, we used novel RVM technology to directly monitor MV in a cohort of patients presenting for total hip and knee arthroplasty. Apneic events were defined as periods of at least 10 seconds with no detected breath. Patients averaging more than five apneic events per hour over the course of their stay in the post-anesthesia care unit (PACU) were defined as demonstrating post-operative apnea (POA). Higher rates of POA were observed in patients receiving PACU opioids than in nonopioid patients (\(p = 0.063\)). This difference was driven by POA rates in general anesthesia (GA) patients receiving opioids. GA opioid patients showed significantly higher incidence of POA (13/30, 43.3%) than either spinal anesthesia (SA) opioid patients (12/62, 19.4%) (\(p = 0.024\)) or GA nonopioid patients (0/13, 0%) (\(p = 0.004\)). Incidence rates were similar among SA opioid and SA nonopioid patients.

PACU after SA showed a substantially higher rate of OIRD onset (38.7%) than their GA counterparts (20.0%). Although not quite demonstrating a statistically significant advantage for GA, this result certainly challenges the blanket notion that SA corresponds to reduced complications postoperatively. This higher rate was observed in spite of the SA cohort receiving consistently lower dosages of intraoperative opioids and sedatives than the general cohort, and having a markedly lower mean morphine equivalence for opioids received in the PACU as well (SA: 3.80, GA: 7.67). In other words, GA patients responding to opioid analgesia appear to maintain ventilatory function at least as well, if not better than SA patients, despite receiving substantially higher opioid dosing both during and after surgery.

These OIRD results are mirrored by our examination of PORD incidence in the nonopioid division of the cohort. SA nonopioid patients show a noticeably higher rate of PORD (36.8%) than GA nonopioid patients (15.4%). This finding further implicates SA as a factor in postoperative respiratory compromise.

In light of similar results in the OIRD and PORD analyses, we attempted to increase the statistical power of our analysis by implementing a combined Fisher’s Exact Test analyzing incidence of any form of respiratory depression across the combined +opioid/−opioid cohort. As shown in Figure 1B, we found that incidence of respiratory depression was in fact significantly higher in SA patients than in GA patients for the combined cohort (37.7% vs. 18.6%, \(p < 0.025\)). Keeping in mind that opioid intake was markedly higher in GA patients both intraoperatively and postoperatively, this result is strongly suggestive of an association between SA and an increased risk for early PORD.

Our analysis of respiratory competence at patient discharge from the PACU (Figure 2) was less conclusive, with little difference observed between GA and SA patient ventilation in either opioid-subgroup or combination analysis. Still, among GA patients not receiving opioids, not one patient showed LMV at discharge (GA\(^{-}\)opioid: 11/67 LMV at discharge, 0%), giving this group a slight advantage over its SA counterpart (SA\(^{-}\)opioid: 11/67 LMV at discharge, 16.4%). Collectively, our findings suggest that SA provided no advantage to GA with respect to postoperative respiratory compromise in the PACU after major orthopedic surgery. Although surprising, the observation of a higher incidence of respiratory depression among SA patients has potential explanations. Pain is known to act as a physiological antagonist of the respiratory depressant effects of opioid analgesics.\(^3,4\) It is possible that by reducing pain stimulation relative to GA, patients with SA are more vulnerable to OIRD.

Our results also speak to the dangers of stratifying patient respiratory risk postoperatively according to any generalized set of assumptions. Respiratory compromise is a major concern for PACU staff. Patients may have residual anesthetic, neural muscular blockade, hypercarbia, hypoxemia, airway obstruction, reduced RR, reduced TV, reduced
MV, and varying degrees of sedation. These risks are further compounded by the effects of surgical insult, the use of opioid analgesia, and potential manifestations of latent respiratory risk factors such as undiagnosed obstructive sleep apnea.

In the context of such a complex threat to patient safety, it may be tempting to stratify patient risk for respiratory complications according to anesthesia type, BMI, comorbidities, or any number of other factors, in an attempt to compartmentalize the attention of caregivers. However, as the results of this study show, patterns of respiratory compromise can be unpredictable in the PACU, at times defying traditionally accepted indicators of risk. Even in what are considered to be the safest of patient populations, preventable respiratory complications may arise, leading to increased morbidity, and cost of care. Our findings suggest that there is a need for careful, individual monitoring of ventilation in all postoperative patients, regardless of perceived risk. This is consistent with growing demand for continuous respiratory monitoring in postoperative care over the last 5 years.

Disordered or obstructed breathing, as we have mentioned, is a significant contributor to respiratory risk in the PACU. In the interest of including this phenomenon in the scope of our discussion of respiratory compromise, we were able to conduct a posthoc analysis in which we evaluated the incidence of what we termed POA in patients receiving general and SA. Patients were classified as demonstrating POA if they averaged more than five apneic or hypopneic events per hour over their entire PACU stay. This definition was intentionally blind to cause, as apnea can manifest postoperatively for a variety of reasons, and in the general clinical environment it is difficult if not impossible to identify which is responsible in a given patient.

Here we see that GA patients receiving opioids show a dramatically higher rate of POA than SA patients receiving opioids or GA patients receiving no opioids, whereas POA rates in SA patients are similar regardless of whether opioids were given. This is perhaps suggestive of a synergistic effect between GA and opioid analgesia in eliciting apnea per se. However, the result may also be skewed by the higher opioid dosages among GA patients in this cohort. Further research is required to investigate the potentially interactive effects of anesthesia and opioid analgesia on breathing obstruction. As an aside, it is interesting to note that among GA patients receiving no postoperative opioids, not one patient had an incidence of POA or showed LMV at PACU discharge. The small sample size for such patients (n = 11) limits any conclusion, but this subgroup appears tentatively to be the least prone to respiratory compromise, and further research may be warranted.

This study was limited by an inability to continue monitoring on the hospital floor after discharge from the PACU, which would allow the assessment of respiratory status over a longer period of time than just the first 2 to 5 hours after surgery. Another major limitation of this study was the inability to follow these patients on the floor and determine the impact of a LMV at PACU discharge. In addition, this study was limited to a specific, elective surgical population, and consequently the distributions of age and ethnicity were not completely representative of the general populace.

CONCLUSION

RVM is a novel monitoring modality that was used to identify patients with respiratory depression following elective major lower extremity orthopedic surgery. A surprising finding was a greater association of respiratory depression in patients receiving SA compared to patients who received GA, in the context of recent work where SA was found to be generally associated with reduced complications. While these findings highlight a previously unrecognized potential risk of SA, this study also demonstrates the value of postoperative respiratory volume monitoring. We propose future studies using an RVM to detect and report unsafe respiratory status in combat casualties with major pelvi-perineal and/or lower extremity traumatic injuries.

REFERENCES

Anesthesia and Respiratory Compromise: Implications for Combat Casualties


Impact of Oxygenation Status on the Noninvasive Measurement of Hemoglobin

Dina Gomaa, BS RRT†; Dario Rodriguez Jr., MSc RRT†; Lt Col Michael Petro, USAF‡; Thomas C. Blakeman, MSc RRT†; Richard D. Branson, MSc RRT* 

ABSTRACT Background: Noninvasive monitoring of hemoglobin (SpHgb) via pulse oximetry has the potential to alert caregivers to blood loss. Previous studies have demonstrated that changes in oxygenation may impact accuracy. Methods: Twenty normal volunteers were monitored using SpHgb, at sea level, during ascent to 14,000 feet, at 14,000 feet with 100% oxygen delivery, and again at sea level. Each period consisted of 15 minutes of monitoring. SpHgb measurements were compared to a blood sample using Bland-Altman analysis. The loss of the SpHgb signal was also recorded. Results: The mean difference in measured hemoglobin (Hgb) between a venous sample and SpHgb was 2.6 ± 0.96 at 14,000 feet. Ascent to 14,000 feet resulted in a predictable fall in SpO2 and was associated with loss of the SpHgb signal for half the period of observation (7.4 minutes). In the other three conditions, SpHgb signal was missing 1 to 12.6% of the time. The nadir SpO2 was not predictive of the loss of SpHgb signal. Discussion: Changes in oxygenation in normal volunteers are associated with short-term SpHgb signal loss (<10 minutes), but no impact on the measured SpHgb.

INTRODUCTION

Pulse oximetry for the measurement of oxygen saturation (SpO2) is ubiquitous in a wide range of care environments and is a standard of care in the intensive care unit. In recent years, noninvasive oximetry has evolved to include the measurement of hemoglobin (SpHgb), carboxyhemoglobin, and methemoglobin.1 The utility and accuracy of these measurements have met with conflicting and often disappointing results.2–8 In the operating room and in the intensive care unit, the accuracy of SpHgb has not proven reliable enough to guide transfusion decisions.2,4–8 We found the use of SpHgb to be unreliable in reflecting actual hemoglobin in trauma and surgical patients at risk for hemorrhage.8 In normal volunteers and in noncritically ill patients, the results have been more promising.8–11

The accuracy of SpO2 measurements are altered by a number of common clinical consequences including skin pigmentation, perfusion, nail polish, motion, and external light interference.12 At least one study has demonstrated that decreased perfusion reduces the accuracy of SpHgb.13

In a previous study evaluating oxygen requirements for normal subjects exposed to hypobaric conditions (14,000 feet altitude), we noted that the SpHgb measurement was influenced by sudden hypoxemia.14 In these subjects, the SpO2 signal was lost for several minutes and when the signal returned, the measured SpHgb fell by nearly 2.0 grams per deciliter (g/dL) despite no blood loss. Concordantly, Gayat et al found that hyperoxia resulted in an increase in SpHgb equivalent to approximately 2.0 g/dL (9.9–12.2 g/dL) in patients in the operating room.15 Both studies and our current work used the RAD-7. We designed a study to evaluate the impact of sudden changes in oxygenation, both hypoxemia and hyperoxemia, on the accuracy of SpHgb in a group of normal subjects. We hypothesized that changes in oxygenation would alter the accuracy of SpHgb in spite of stable red blood cell volume.

METHODS

The study was approved by the institutional review boards of the University of Cincinnati Medical Center and Wright Patterson Air Force Base. A flyer was approved to recruit potential participants for the study. Flyers were provided to interested subjects compromising a pool of candidates with altitude chamber certification. Inclusion criteria included altitude chamber certification up to 22,000 feet and age ≥ 18 and ≤ 65 years of age.

Subjects were excluded if they had a medical history of migraine headaches or known sickle cell trait. Informed consent was obtained from each subject before any study-related procedure. Consent was obtained by one of the civilian study staff; USAF personnel were prohibited from administering consent in an effort to eliminate the perception of coercion.

On the morning of the study, venous blood samples of 3 to 5 mL were obtained from subjects with a single venipuncture from the antecubital space for analysis of hemoglobin. Subjects were then seated in the altitude chamber and pulse oximeter probes (Masimo, Irvine, California) were placed on the second or third digit of one hand. Oximeter probes were connected to the respective oximeters (Rad-7; Masimo). Probes were covered to prevent cross talk or interference

*Department of Surgery, University of Cincinnati, 231 Albert Sabin Way, Cincinnati, OH 45267. 
†Aeromedical Research Branch, USAFSAM, 2510 Fifth Street Building 840, Wright-Patterson AFB, OH 45433–7901.
‡Center of Sustainment and Trauma Readiness Skills, 234 Goodman Street, Cincinnati, OH 45229.

doi: 10.7205/MILMED-D-16-00137
from ambient light. Subjects remained at ambient barometric pressure for 15 minutes (normobaric normoxia). All oximeter data were recorded to the internal memory of the oximeter every 2 seconds.

Subjects ascended to 14,000 feet of simulated altitude (barometric pressure, 460 mm Hg) at 2,500 feet per minute. Predicted PaO₂ at this altitude using the alveolar air equation (PaO₂ = FIO₂ [PB mm Hg – 47 mm Hg] – PaCO₂/0.8), assuming a PaCO₂ of 30 mm Hg and an alveolar to arterial gradient (A – a gradient) of 5 mm Hg was 45 mm Hg. Subjects remained at this altitude for 15 minutes while data were continuously collected to the internal memory of the oximeters. At this altitude, the anticipated SpO₂ was 82 to 84% (hypobaric hypoxia). If SpO₂ fell below 82% or the subject demonstrated anxiety, discomfort, or signs of altitude sickness, supplemental oxygen was applied.

Following 15 minutes of hypobaric hypoxia, subjects were administered 100% oxygen via a nonrebreathing mask (hypobaric hyperoxia). Data were continuously recorded to the oximeter for 15 minutes. At the end of this period, patients were returned to sea level and oxygen was discontinued (normobaric normoxia) for 15 minutes and recordings were continued. The study conduct is shown in Figure 1.

During each phase of the study, manual recordings of data were made every 5 minutes as a backup to the electronic recordings. Throughout the study, subjects were monitored for signs of high altitude sickness and decompression sickness.

All data were reviewed for accuracy and missing values. Mean data at the end of each 15 minute period (normobaric normoxia, hypobaric hypoxia, hypobaric hyperoxia, and normobaric normoxia) were compared to the measured hemoglobin at sea level using a t test and Bland Altman analysis. The number of missing values during each 15-minute period were compared using nonparametric signed rank tests. Regression analysis was used to measure the rate of change of SpO₂ immediately before loss of signal. All data are mean ± SD.

RESULTS

Twenty subjects (17 men and 3 women) with a mean age of 31.4 ± 7.1 years were studied and all subjects completed the trial. There were 15 Caucasian, 3 African American, and 2 Asian subjects. Subjects were monitored using the RAD-7 oximeter. Differences in monitored subjects were related to availability of equipment at the time of the study. No subject required emergent oxygen administration outside the specified protocol to relieve severe hypoxemia or discomfort while at altitude. At baseline normobaric normoxia the mean SpHgb was 2.74 g/dL less than the measured blood Hgb (95% confidence interval = −3.20, −2.78, p < 0.01). This relationship was not appreciably different at each of the other 3 conditions (Table I).

At sea level, the mean SpO₂ was 97.1% (range, 90–99%) and the mean SpHgb was 12.4 g/dL. The percentage of time with missing data during this period was 1.1% (22 seconds). At 14,000 feet, the mean SpO₂ after 15 minutes was 87% (range, 82–89%), mean SpHgb was 12.8 g/dL, and 49% (9.8 minutes) of data was missing. Table I compares the mean SpO₂, mean SpHgb, and duration of missing data for each study period.

Subjects monitored with the lowest nadir SpO₂ values at 14,000 feet demonstrated a loss of SpHgb signal in 7 of

<table>
<thead>
<tr>
<th>TABLE I.</th>
<th>Measured Hemoglobin From a Venous Blood Sample Compared With the SpHgb Value and Missing Data During Each Study Phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study Condition</td>
<td>Mean Difference in Hgb (±SD)</td>
</tr>
<tr>
<td>Sea Level 1 (n = 20)</td>
<td>−2.74 ± 1.0</td>
</tr>
<tr>
<td>14,000 Feet (n = 20)</td>
<td>−2.63 ± 0.95</td>
</tr>
<tr>
<td>14,000 Feet w/ Oxygen (n = 20)</td>
<td>−2.45 ± 1.0</td>
</tr>
<tr>
<td>Sea Level 2 (n = 20)</td>
<td>−2.61 ± 0.82</td>
</tr>
<tr>
<td>Study Condition</td>
<td>Missing SpHgb (mins [%])</td>
</tr>
<tr>
<td>Sea Level 1</td>
<td>0.22 (1.1)</td>
</tr>
<tr>
<td>14,000 Feet</td>
<td>7.4 (49.1)</td>
</tr>
<tr>
<td>14,000 Feet w/ O₂</td>
<td>1.89 (12.6)</td>
</tr>
<tr>
<td>Sea Level 2</td>
<td>1.14 (7.6)</td>
</tr>
</tbody>
</table>

Mean differences between blood hemoglobin and hemoglobin measured by pulse oximetry (top). Duration of missing data at each study condition (bottom).
10 cases. In the 10 cases with nadir SpO₂ of >87%, only 5 of 10 subjects demonstrated loss of the SpHgb signal. This difference was not statistically significant ($p = 0.36$) but suggests that the greater the decrement in SpO₂, the greater risk of signal loss. Analysis of bias and precision demonstrates the underestimation of Hgb by SpHgb across the study conditions (Figs. 4–7).

**DISCUSSION**

We found that sudden hypoxia in normal subjects resulted in the loss of hemoglobin signal for an average of almost 10 minutes. This finding is counter to the work of Gayat et al and our own previous work. The SpHgb value was not significantly different at any of the four barometric and oxygenation conditions. We did, however, find that sudden hypoxia was associated with “loss of the SpHgb” signal for 7 minutes. Bland Altman analysis demonstrated that the mean SpHgb was $-2.45$ to $-2.75$ g/dL less than the measured Hgb from a venous sample regardless of SpO₂.

Hyperoxia at hypobaric conditions did not have a significant impact on SpHgb. The time of missing data during oxygen administration was 12.6% of 15 minutes or just less than 3 minutes. These findings are different than our earlier work and that of Gayat. Figure 2 depicts the average change in SpO₂ associated with changes in oxygenation from these two previous trials. Data from the current study are demonstrated in Figure 3.

The algorithms for monitoring noninvasive signals are constantly improving. We believe that the difference between this study and our previous work is an advancement in the
technology. In our previous work and that of Gayat et al, changes in oxygenation resulted in further inaccuracies in SpHgb measurements. In the case of hypoxemia, decreases in SpHgb could have been misinterpreted and led to inappropriate management decisions. The current generation of oximeters does not display erroneous data, but rather waits to display data when the signal has stabilized. We believe this is an advantage over previous devices.

The article by Gayat et al demonstrated that hyperoxia resulted in a measured increase in SpHgb despite no intervention that could explain this change. In our study, hyperoxia did not result in any appreciable change in SpHgb, although there was a short duration of missing data. Missing data were more frequently associated with sudden hypoxia than hyperoxia.

The difference in laboratory assessments of hemoglobin from a blood sample and SpHgb in this study was 2.5 to 2.75 g/dL. This difference was unaffected by the oxygenation status. The SpHgb values were consistent and perhaps could be used for trending. Several investigators have shown that data from SpHgb are unreliable for guiding treatment decisions. This is in concordance with the manufacturer’s claims that changes in SpHgb should be verified by laboratory analysis.

There are a number of limitations to this study. We only used normal volunteers with a normal range of Hgb. Only one blood sample was obtained at sea level, before exposure to altitude and oxygen. There are no data or physiologic reasons to suspect that venous Hgb would change with exposure to altitude or oxygen. The operation of an oximeter utilizes tissue bed determinations of oxygen and hemoglobin which can be slightly altered from blood measurements. We used three different RAD-7 oximeters in the subjects to obtain a representative sample of device operation.

FIGURE 4. Bland Altman analysis demonstrating bias and precision of SpHgb at sea level and oxygen saturation values.

FIGURE 6. Bland Altman analysis demonstrating bias and precision of SpHgb at 14,000 feet with oxygen and oxygen saturation values.

FIGURE 5. Bland Altman analysis demonstrating bias and precision of SpHgb at 14,000 feet and oxygen saturation values.

FIGURE 6. Bland Altman analysis demonstrating bias and precision of SpHgb at 14,000 feet with oxygen and oxygen saturation values.

FIGURE 7. Bland Altman analysis demonstrating bias and precision of SpHgb at sea level and oxygen saturation values at the end of the study.
However, this may not reflect the operation of all devices and software versions.

The current work demonstrates improved stability of SpHgb measurement in the face of both hypoxia and hyperoxia compared to previous work. The values at the three study conditions were not statistically or clinically different. The relationship to the measured blood Hgb was unchanged. The loss of the SpHgb signal is transient and following signal loss, the SpHgb value remains stable in a group of normal volunteers. Changes in SpHgb algorithms over the last few years appear to have improved the ability of the measurement to compensate for sudden changes in oxygenation.

REFERENCES

Wearable Pulse Oximetry Measurements on the Torso, Arms, and Legs: A Proof of Concept

Marcus Kramer, PhD; Aaron Lobbestael, MS; Emily Barten, MS; John Eian, MS; Gregory Rausch, MS

ABSTRACT  For decades pulse oximeters designed for use on the head, hands, or feet have provided invaluable estimates of oxygen saturation to medical personnel attending to combat casualties. However, traditional placement sites are not ideal for the relatively new paradigm of continuous battlefield telemonitoring. To assess the feasibility of oximetry on nontraditional body sites, 42 healthy volunteers were enrolled, consented, and underwent an industry standard induced-hypoxia study. During the study volunteers used prototype wearable oximeters, designed for the torso, arms, and legs. Subsets (size $n$) of the volunteers had the wearables placed at the following body sites, and achieved accuracies ($A_{RMS}$, root-mean-square difference) of the following: calf 1.7% ($n = 26$); bicep 3.1% ($n = 12$); forearm 3.4% ($n = 11$); pectoral 2.9% ($n = 42$); sternum 2.9% ($n = 13$). In keeping with regulatory guidance calibrations with an $A_{RMS}$ of less than 3.5% are acceptable for potential future development. Additionally, a new method was developed to enable accurate reporting of respiration rate from the pectoral oximeter. $A_{RMS}$ of 1.1 breaths per minute ($n = 10$). This study demonstrates the feasibility of monitoring oxygen saturation and respiration rate from nontraditional sites via a wearable pulse oximeter.

INTRODUCTION

Recent years have seen an upsurge in the investigation of wearable sensors to meet the growing civilian and military needs for continuous critical care monitoring. Such wearable devices have included measures of heart rate (HR), temperature, oxygen saturation, electrocardiogram, and respiration rate among others. 

In modern medical practice, peripheral oxygen saturation ($SpO_2$), HR, and respiratory rate have been identified as crucial metrics to protect against the onset of dangerous and potentially life-threatening conditions such as sepsis and pulmonary embolism. Traditional pulse oximetry provides both $SpO_2$ and HR, whereas respiration rate can be provided through a variety of means including impedance pneumography or capnography. Still, the possibility of accessing all three of these parameters from a single sensor has led to many notable efforts to derive respiration rate from traditional pulse oximeters.

Likewise, there have been a number of attempts to incorporate oximetry into new wearable sensors. Unfortunately, these efforts have largely been limited to traditional oximetry sites. For example, Mendelson et al has explored the implementation of a wearable oximeter on the forehead for military and civilian use. Likewise, He et al described a new wearable oximeter for behind the ear, citing the location as having good pulse quality, being discrete and stable with natural anchoring to the ear. 

Inevitably, to meet the growing medical needs of civilians and the military, new oximeter measurement technologies will need to be developed which can make a broader array of sites available.

In traditional pulse oximetry the sensors are placed on the head, hands, or feet as a result of their strong perfusion, accessibility, and lower intersubject variability. Indeed, the fingers, palms, face, and ears have significantly higher perfusion than the rest of the body sites such as the torso, arms, and legs. Likewise, with respect to physiologic characteristics such as adipose tissue, which accumulates disproportionately on the torso, legs, and upper arms, the traditional sites have significantly lower intersubject variability. Similarly, the accessibility and reliability of transmission oximetry on the fingers, toes, and ears have made these sites a mainstay in modern medical practice. Even forehead reflectance sensors have gained market acceptance as a replacement for the other traditional sites in cases of compromised circulation.

Nevertheless, soldiers on the battlefield require new solutions which free the head, hands, and feet, as these sites are often already in use or unavailable after injury. Notwithstanding the many advances in traditional pulse oximetry since its development in 1973 little progress has been made transitioning away from the traditional sites to new sites. This lack of progress has hampered the development of new wearable medical devices which could incorporate these valuable parameters.

Using our extensive experience designing small, low-power reflectance pulse, and tissue oximetry systems Nonin (Plymouth, Minnesota) has developed a wearable pulse oximeter for use on nontraditional sites. Presented here are the results of a pilot study to evaluate the accuracy of these new oximeters (Fig. 1A). The nontraditional body sites were selected from the torso, arms, and legs based on the technical feasibility, clinical acceptance, and access to the anatomical site. This new technology was assessed on a demographically diverse patient population in a standard-induced hypoxia clinical trial.
METHODS
This study was approved by the University of California San Francisco Human Research Protection Program and conducted at the University of California San Francisco Hypoxia Research Laboratory in compliance with the principles of the Declaration of Helsinki.

Study Design
Forty-two healthy volunteers were enrolled, consented, and had sensors placed on up to five sites. The wearable sensors were placed on the chest, sternum, bicep, forearm, or calf before the start of hypoxia as shown in Figure 2. A Nonin signal processor (Model X-100SP) with a modified 8100AA (not available for sale) fingertip sensor, a commercially available fingertip oximeter (used by the research lab to monitor the patient) and a radial artery catheter were placed for monitoring during the study. Demographic (skin tone and race) and anthropometric (skin fold thickness measurements, height, and weight) information was collected before the study. Volunteers were recruited to reach an approximately equal mix of males and females with varying skin pigmentation. Race was self-identified and was collected to ensure that a range of ethnicities were captured. Skin color was judged by the study coordinator and was collected to examine effects of skin pigmentation on oximeter accuracy. Given the risks involved in induced hypoxia studies, the volunteers were limited to the young (ages 18–45 years) and healthy (excluded for a body mass index [BMI] greater than 31, pregnancy, respiratory conditions, etc.).

The volunteers underwent induced hypoxia using an industry-standard protocol. During the study, end-tidal O₂, and CO₂ were monitored to assess subjects metabolic condition and aid in controlling the depth of hypoxia. Hypoxia was induced by lowering the oxygen content of the air respired by the volunteers to achieve arterial oxygen saturation (SaO₂) blood draws ranging from 70 to 100%. Volunteers were taken through seven saturation plateaus (Fig. 3) to a minimum SpO₂ at around 70%. After the volunteer’s saturation stabilized according to the fingertip pulse oximeter at each plateau arterial blood was collected concurrently with optical data from the oximeters, and evaluated with three ABL90 FLEX CO-oximeters (Radiometer, Brønshøj, Denmark). The average of the three CO-oximetry measurements was used as the truth value to determine the accuracy of the wearable oximeters.

System Design
Raw optical measurements were recorded using a custom data collection system consisting of a data storage and visualization system, Nonin signal processors (Model X-100SP), and custom wearable sensors. The sensors were designed to be mechanically optimized to the arms, legs, and chest (Fig. 1A). The sensor was attached to the patient through a flexible, medically safe, adhesive patch bonded to the sensor housing (Fig. 2). The spacing between the emitter and detector were optimized to a value between 3 and 20 millimeters to maximize the pulse signal, optical power, and accuracy of the sensor from each site (Fig. 1B). The emitter was composed of light emitting diodes (LED) with two wavelengths in the 600 to 1000 nanometer range, and the detector was a silicon photodiode. The sensors were connected to a commercially available Nonin signal processor. The signal processor was hardwired directly to a laptop through a proprietary universal serial bus converter. The laptop was used to store and visualize raw optical measurements from the signal processor. Data were stored on the laptop for future processing and calibration.

Data Analysis
Following the study, the raw optical measurements were processed using Nonin’s proprietary signal processing algorithm for pulse identification, qualification, and quantification. For each qualified pulse, the ratio of pulse amplitudes...
was output. Analysis was completed in SAS 9.3 (SAS Institute, Cary, North Carolina) and Matlab (Mathworks, Natick, Massachusetts) to calibrate the sensors and provide a preliminary assessment of the oximetry system performance.

For each site, a unique calibration was determined using the entire range of data collected from all subjects. Measurements of SaO2, averaged from the three CO-oximeters, and the average ratio of pulse amplitudes during the blood draw for each of the wearable sensors, were paired. SpO2 was calculated using a single a priori model common to pulse oximetry:

\[
\text{SpO}_2 = \frac{A \times R + B}{C \times R + D}
\]

where \( R \) is the average ratio of pulse amplitudes from two wavelengths of light during the blood draw, and \( A, B, C, \) and \( D \) were calibrated for each site to minimize the root-mean-square difference between SaO2 and SpO2 (\( A_{\text{RMS}} \)) using data across the entire range collected. \( A_{\text{RMS}} \) is a representation of the combination of systematic and random error. It is defined as the accuracy root-mean-square error difference between the measured value and the truth value calculated as:

\[
A_{\text{RMS}} = \sqrt{\frac{\sum_{i=1}^{n}(\text{SpO}_2^i)^2}{n}}
\]

where \( n \) is the number of observations, and both SpO2 and SaO2 are defined above.

Observations were dropped in accordance with food and drug administration (FDA) standards for data analysis,\(^{16}\) due to no good pulses being declared from Nonin’s proprietary signal processing or because they were outside of the standard range for reporting accuracy of pulse oximeters (70–100%).

The preliminary accuracy assessment, the mean bias is presented as the least squares mean for each site between 70% and 100% as estimated by a mixed model that accounts for the multiple measurements and multiple measurement sites per subject.\(^ {17}\) The Pearson correlation coefficient (PCC) is also presented by site for saturations between 70% and 100%. The 95% bias-corrected adjusted bootstrap confidence intervals are provided for the \( A_{\text{RMS}} \) using 2000 bootstrap samples.

Summary statistics and data plots were also generated for comparison among the potential anatomical body sites. To compare between sites confidence intervals of the difference in \( A_{\text{RMS}} \) between sites are provided. These intervals account for some subjects having more than one site tested at a time. Sites with an \( A_{\text{RMS}} \) value of less than 3.5% for SaO2 values between 70% and 100% are considered favorable for future development in keeping with regulatory guidance for reflectance oximeters.\(^ {19}\)

\( A_{\text{RMS}} \) statistics were also generated for the full range of data available to demonstrate and compare the accuracy of the device beyond the range used by the FDA to regulate
pulse oximetry. $A_{RMS}$ values are reported for the entire range, the 60% to 100% range and the 70% to 100% range in the manuscript.

Assessment of the effect of demographic and anthropometric characteristics on accuracy was completed through analysis of variance for the categorical variables (gender, skin tone, and race), and generalized linear regression for continuous variables (height, weight, BMI, and blood pressure) with a $p$ value of less than 0.05 signifying a statistically significant effect on bias after adjusting for using a false discovery rate adjustment for multiplicity.

A proprietary digital signal processing algorithm was developed to estimate respiration rate from the pectoral sensor’s optical data. The respiration rate algorithm was developed using data from subjects undergoing spontaneous breathing during the course of the hypoxia study described above. Paired capnography data were available for 10 of the 42 hypoxia subjects, and portions of data from a subset of those subjects were used to develop the respiration rate algorithm. Once developed the algorithm was applied to rolling 30 second windows of pectoral data from all 10 subjects to provide a mean respiration rate over the windows. Quality metrics for the respiration rate algorithm were defined to be the correlation, bias, and $A_{RMS}$ between the algorithm respiration rate and a respiration rate inferred from the paired capnography data.

RESULTS

Forty-two healthy, nonsmoking volunteers were enrolled in the study and underwent hypoxia. Sensors were placed on all 42 volunteers but not all volunteers received the same sensors. See Table I for details on the number of subjects each sensor was tested on.

The study sample had an average age of 26 years, with an average BMI of 23 kg/m² (range 19–31 kg/m²). Subjects were mainly Caucasian (27/42) and male (34/42) (see Table II for more details). Demographic and anthropomorphic characteristics were not found to have a significant impact on the accuracy of the tested devices as determined by the metrics set forth in the methods section.

Data points from the study were dropped in accordance with the methods described previously. Table I reports the number of expected data points per site, with the number of draws dropped for various reasons (missing co-oximeter value, device signal quality errors, saturation greater than or equal to 100, or saturation less than 70), and the accuracy in each of the 3 saturation ranges.

Raw optical measurements recorded using the data collection system were postprocessed to identify a photoplethysmograph which was used to derive $SpO_2$. Sensor performance on the legs and arms varied between sites. The calf was found to have an accuracy, $A_{RMS}$ of 1.7% (95% bootstrap confidence interval [1.5, 2.1], number of subjects = 26, mean bias = 0.17, PCC = 0.99, $SaO_2$ range = 70–100%) in line with the modified 8100AA fingertip oximeter, $A_{RMS}$ of 2.0% (95% bootstrap confidence interval [1.5, 3.5] number of subjects = 42). However, performance was significantly degraded (Tables I and III) on the bicep where the $A_{RMS}$ was 3.1% (95% bootstrap confidence interval [2.1, 5.6], number of subjects = 12, mean bias = 0.32, PCC = 0.95, $SaO_2$ range = 70–100%) or the forearm which had an $A_{RMS}$ of 3.4% (95% bootstrap confidence interval [2.4, 5.2], number of subjects = 11, mean bias = 0.48, PCC =

<table>
<thead>
<tr>
<th>TABLE I.</th>
<th>Accuracy of Wearable Pulse Oximeters Used on Nontraditional Body Sites From Single Site Specific Calibration Across All Subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Site</td>
<td>Number of Subjects</td>
</tr>
<tr>
<td>Bicep</td>
<td>12</td>
</tr>
<tr>
<td>Calf</td>
<td>26</td>
</tr>
<tr>
<td>Forearm</td>
<td>11</td>
</tr>
<tr>
<td>Pectoral</td>
<td>42</td>
</tr>
<tr>
<td>Sternum</td>
<td>13</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TABLE II.</th>
<th>Demographic and Anthropomorphic Characteristics of Volunteers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characteristic</td>
<td>Level</td>
</tr>
<tr>
<td>Age</td>
<td>26 ± 5</td>
</tr>
<tr>
<td>Gender</td>
<td>Male</td>
</tr>
<tr>
<td>Height</td>
<td>69 ± 3</td>
</tr>
<tr>
<td>Weight</td>
<td>153 ± 22</td>
</tr>
<tr>
<td>BMI</td>
<td>23 ± 3</td>
</tr>
<tr>
<td>Blood Pressure (mmHg)</td>
<td>133 ± 15</td>
</tr>
<tr>
<td>Race</td>
<td>White</td>
</tr>
<tr>
<td>Skin Tone</td>
<td>Intermediate</td>
</tr>
<tr>
<td></td>
<td>Light</td>
</tr>
<tr>
<td></td>
<td>Dark</td>
</tr>
<tr>
<td></td>
<td>Very Dark</td>
</tr>
</tbody>
</table>

Data are presented as mean ± standard deviation.
Wearable Pulse Oximetry Measurements on the Torso, Arms, and Legs

TABLE III. Comparison Between Accuracy ($A_{RMS}$, 70–100%) from Different Sites

<table>
<thead>
<tr>
<th>Site 1</th>
<th>Site 2</th>
<th>Difference in $A_{RMS}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bicep</td>
<td>Forearm</td>
<td>−0.4 (−1.4, 1.3)</td>
</tr>
<tr>
<td>Calf</td>
<td>Bicep</td>
<td>−1.3 (−3.9, −0.3)</td>
</tr>
<tr>
<td>Calf</td>
<td>Forearm</td>
<td>−1.7 (−3.5, −0.8)</td>
</tr>
<tr>
<td>Calf</td>
<td>Pectoral</td>
<td>−1.2 (−1.9, −0.7)</td>
</tr>
<tr>
<td>Calf</td>
<td>Sternum</td>
<td>−1.1 (−2.5, −0.2)</td>
</tr>
<tr>
<td>Pectoral</td>
<td>Bicep</td>
<td>−0.1 (−2.2, 0.7)</td>
</tr>
<tr>
<td>Pectoral</td>
<td>Forearm</td>
<td>−0.5 (−1.7, 0.4)</td>
</tr>
<tr>
<td>Pectoral</td>
<td>Sternum</td>
<td>0.1 (−1.6, 1.3)</td>
</tr>
<tr>
<td>Sternum</td>
<td>Bicep</td>
<td>−0.2 (−2.3, 1.6)</td>
</tr>
<tr>
<td>Sternum</td>
<td>Forearm</td>
<td>−0.6 (−2.5, 1.0)</td>
</tr>
</tbody>
</table>

0.95, $SaO_2$ range = 70–100%). $SpO_2$ accuracy results are summarized in Table I. Similar results are seen for the sensors when examining their full range of data collected, 60% to 100%, which is outside of FDA reporting guidelines, 70% to 100% (Table I and Fig. 4).

Even with the encumbrance of large motion from respiration, the sternum and pectoral oximeters both produced acceptable accuracies. During the study the pectoral oximeter was placed on 42 subjects with an accuracy of 2.9% (95% bootstrap confidence interval [2.4, 3.5], number of subjects = 42, mean bias = −0.12%, PCC = 0.96, $SaO_2$ range = 70–100%), whereas the sternum oximeter was only placed on 13 subjects but also produced an accuracy of 2.9% (95% bootstrap confidence interval [2.0, 4.8], number of subjects = 13, mean bias = 0.18, PCC = 0.96). Results on the torso are correspondingly worse than on the calf (Tables I and III) but still within guidelines set for the study. Similar results are seen for the sensors when examining their full range of data collected, 60% to 100%, which is outside of FDA reporting guidelines, 70% to 100% (Table I and Fig. 4).

Further analysis of the pectoral sensor data also uncovered a new method to determine respiration rate. End-tidal CO$_2$ and oximeter data could be paired for 10 of the 42 subjects. The other 32 subjects either did not have end-tidal CO$_2$ recordings completed or the data could not be paired with the oximeter data. An accuracy, $A_{RMS}$, of 1.1 breaths per minute, was achieved with a highly variable respiration rate between 10 and 45 breaths per minute (mean bias = 0.1%, PCC = 0.9988, Fig. 5).

DISCUSSION

The development of a wearable pulse oximeter for nontraditional sites is a challenging endeavor. In traditional transmission pulse oximetry the LED and detector are placed on opposing sides of the tissue bed. The light emitted by the LED is predominantly scattered forward through the tissue, therefore, the majority of the light is transmitted to the opposing detector. In reflectance pulse oximetry, the detector is placed on the same side of the tissue as the LED (Fig. 1B) and a small portion of the light scattered by the tissue will instead reflect back to the same side of the tissue it came from (Fig. 1C). Therefore, only the small fraction of light reflected by the tissue will reach the detector and be usable for measurement. Methods such as driving the LEDs harder or an annular photodiode could be used to produce an equivalent optical signal to that of transmission oximetry but require more power, or a larger sensor size.

During the development of these wearable oximeters Nonin leveraged their experience designing medical grade oximeters that are favored for their small size and low power. The current prototype oximeter incorporates Nonin’s commercially available low-power and low-noise signal processor (model X-100SP) in combination with anatomically and optically optimized reusable sensors. However, through minor modification the signal processor cable can be removed to convert the unit into the wireless, wearable oximeter module shown in Figure 1A. This system works in combination with proprietary digital signal processing techniques providing a high-quality, low-noise optical measurement.

The 10 blood draws per subject, expected for the calibration based on the study design, were not realized due to co-oximeter values missing, and signal quality errors. The raw optical measurements were digitally processed to improve signal quality, before calibration. This involves multiple stages of filtering and signal quality testing used to distinguish heart beats from noise or interfering signals. Inappropriate signals or distorted heartbeats are discarded by the signal quality tests. As a result, fewer blood draws are available for the.

TABLE I. Accuracy of Wearable Pulse Oximeters Used on Nontraditional Body Sites From Single Site Specific Calibration Across All Subjects

<table>
<thead>
<tr>
<th>Number of Draws With $SaO_2 ≥ 100$</th>
<th>Number of Draws With $SaO_2 (60-100)$</th>
<th>$A_{RMS}$ With $SaO_2 (60-100)$</th>
<th>Number of Draws With $SaO_2 &lt;70$</th>
<th>Number of Draws With $SaO_2 (70-100)$</th>
<th>Least Squares Mean Bias With $SaO_2 (70-100)$</th>
<th>PCC</th>
<th>$A_{RMS}$ With $SaO_2 (70-100)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>94</td>
<td>3.2 (2.1, 6.2)</td>
<td>9</td>
<td>85</td>
<td>0.32 (−0.52, 1.17)</td>
<td>0.95</td>
<td>3.1 (2.1, 5.6)</td>
</tr>
<tr>
<td>26</td>
<td>229</td>
<td>1.8 (1.5, 2.2)</td>
<td>25</td>
<td>204</td>
<td>0.17 (−0.14, 0.47)</td>
<td>0.99</td>
<td>1.7 (1.5, 2.1)</td>
</tr>
<tr>
<td>10</td>
<td>82</td>
<td>3.6 (2.5, 5.4)</td>
<td>9</td>
<td>73</td>
<td>0.48 (−0.27, 1.23)</td>
<td>0.95</td>
<td>3.4 (2.4, 5.2)</td>
</tr>
<tr>
<td>38</td>
<td>351</td>
<td>2.8 (2.3, 3.6)</td>
<td>30</td>
<td>321</td>
<td>−0.12 (−0.41, 0.16)</td>
<td>0.96</td>
<td>2.9 (2.4, 3.5)</td>
</tr>
<tr>
<td>12</td>
<td>116</td>
<td>3.0 (2.1, 5.0)</td>
<td>5</td>
<td>111</td>
<td>0.18 (−0.54, 0.90)</td>
<td>0.96</td>
<td>2.9 (2.0, 4.8)</td>
</tr>
</tbody>
</table>
calibration model used to assess the accuracy ($A_{RMS}$) of the oximeter system. This can be seen in Table I where, e.g., the pectoral sensor should have 420 blood draws but only 389 were used in the calibration.

In accordance with regulatory standards a minimum of 200 data points from at least 10 subjects, and $A_{RMS}$ better than 3.5% between 70% to 100% are required for a reflectance oximeter to be cleared for medical use in the United States. However, due to the inaccuracy of the reference pulse oximeter (and co-oximeter) and challenges in precisely controlling the desaturation between subjects; the true saturation ($SaO_2$, as measured from the co-oximeter) can inadvertently go outside these guidelines. Therefore, we have reported results for all sensors based on the 70% to 100% criterion to provide a direct comparison of this study’s $A_{RMS}$ to those reported by medical device manufacturers for the fingers, ears, and forehead (Table I). Of the sites tested in this study, the upper pectoral and calf both were better than the minimum FDA requirement. As well, the reported accuracy on the calf is equivalent with commercially available fingertip pulse oximeter accuracies of ± 2% ($A_{RMS}$) and superior to all other sites reported in this study (Table III). It should be recognized that the FDA rounds accuracies to the nearest whole number so both the 1.7% (calf) and 2.0% (8100AA fingertip oximeter) accuracies would be reported as an accuracy of ± 2% by a medical device manufacturer. These sites are therefore very promising for further development. Whereas, the forearm, bicep, and sternum lack enough data points, they do meet the basic accuracy requirements (between 70% and 100%) for further development.

Today the standard of care for patients uses the finger-tip pulse oximeter that monitors $SpO_2$ and HR. However, in their comparative study on key contributors of death, Lynn et al showed that these metrics alone do not provide a holistic picture of patient health. Instead, with the addition of respiration rate, these three metrics may serve to minimize the occurrences of sepsis and pulmonary embolism while aiding in the detection of respiratory distress or failure, assisting in triage efforts. This new wearable oximeter technology could potentially provide $SpO_2$, HR, and respiration rate to military and civilian medical personnel, delivering essential metrics during critical care.

**LIMITATIONS**

This study was limited to young healthy adults. The devices described are prototypes that have not been cleared by the FDA. This study did not include an appropriate truth (e.g. electrocardiography) for the HR parameter derived from the wearable oximeters and as such, accuracy was not presented.
for the parameter. Likewise, the results of this study only estimate the accuracy of SpO₂ and respiration rate not the overall performance of these devices across patient populations and use cases. Furthermore, because these data were used to calibrate the oximeter, we anticipate further study is required to validate the device accuracy.

CONCLUSIONS
This study demonstrates the feasibility of accurately monitoring oxygen saturation from the torso, arms, and legs via a miniaturized wearable pulse oximeter system which combines low-power and low noise hardware with cutting-edge signal processing. The unique calibration, patient interface, and sensor design are some of the many reasons reflectance sensors work effectively only on sites to which they have been optimized. These newly developed wearable oximeters have the potential to perform across nontraditional body sites enabling the monitoring of wounded warriors during the continuum of care. Additionally, the torso oximeter has been enabled to improve critical casualty care by adding accurate respiration rate to its parameter set. Further development will focus on consolidating the design into a single wearable module, validating the accuracy of this new wearable oximetry system and expanding the performance across use cases. This technological evolution represents the first step of developing wearable oximeters for use during the continuum of battle and critical care.

ACKNOWLEDGMENTS
This study was funded by Nonin Medical, Inc. as part of ongoing research and development.

REFERENCES
tance%20pulse%20oximeter%20for%20remote%20physiological%20monitor-
ing.%20Proceedings%20of%20the%2028th%20IEEE; accessed April 5, 2016.
Clinical and Magnetic Resonance Spectroscopic Imaging Findings in Veterans With Blast Mild Traumatic Brain Injury and Post-Traumatic Stress Disorder

Anthony P. Kontos, PhD*; Anne C. Van Cott, MD†‡; Jodilyn Roberts, MS†; Jullie W. Pan, MD, PhD‡; Monique B. Kelly, PhD†; Jamie McAllister-Deitrick, PhD*; Hoby P. Hetherington, PhD§

ABSTRACT Objectives: To compare magnetic resonance spectroscopic imaging (MRSI) findings from the hippocampal regions of military veterans with blast-related mild traumatic brain injury (blast mTBI) and post-traumatic stress disorder (PTSD) to those with PTSD only; and to examine the relationship of MRSI findings to cognitive and neuromotor impairment. Methods: 35 military veterans—23 with blast mTBI and PTSD (blast mTBI/PTSD) and 12 with PTSD only participated in the study. Whole plane MRSI data including N-acetyl aspartate (NAA) and choline (Ch) were acquired at 7T for the hippocampus. Concurrent cognitive and neuromotor data were collected using established assessments. General linear models (GLMs) with Bonferroni correction were used to compare the two groups on NAA/Ch ratios across regions of the hippocampus. Spearman’s correlations were used to examine correlations between NAA/Ch and cognitive and neuromotor impairment. Results: The NAA/Ch results for the left hippocampus were lower in the blast mTBI/PTSD group than the PTSD-only group. The blast mTBI/PTSD group also scored worse on the WAIS-IV-vocabulary. Significant correlations between NAA/Ch and neuromotor outcomes—including vestibular impairment—were supported. Conclusions: Combined MRSI and cognitive and neuromotor data may help inform more objective and accurate diagnoses and effective treatments for patients with blast mTBI and PTSD.

INTRODUCTION
Since 2000, approximately 280,000 U.S. military personnel have sustained a mild traumatic brain injury (mTBI).1 It is estimated that annual health care costs for U.S. military personnel with mTBI are $12,990 compared to $7,377 for non-TBI controls.2 Nearly one in five veterans experience lingering symptoms or impairment related to their mTBI.3 Between 20% and 40% of military personnel experience post-traumatic stress (PTS) symptoms or post-traumatic stress disorder (PTSD) following exposure to mTBI.4,5 Long-term effects of mTBI and PTSD include residual mental and cognitive impairment, increased suicide risk, poor physical health, unemployment, homelessness, and substance abuse.6 Exposure to blast forces such as those from improvised explosive devices, exploding munitions, and other sources is particularly problematic for military personnel, and may result in complicated polytrauma presentations that are difficult to diagnose and treat.7,8 It is challenging for clinicians to distinguish PTSD from lingering mTBI symptoms in this at-risk group, as there is considerable overlap between them. Both PTSD and mTBI symptoms include anxiety, irritability, memory problems, and sleep disturbances.9 This overlap in symptoms presents a challenge in determining the etiology of mTBI versus PTSD symptoms following exposure to mTBI. Misattribution of symptoms to PTSD or mTBI may result in ineffective treatment and prolonged symptoms and morbidity. For example, military personnel misdiagnosed as having mTBI without PTSD may not receive the indicated trauma-focused exposure therapy that is effective with PTSD.10 In contrast, military personnel misdiagnosed with PTSD only may engage in this therapy with limited success as a result of mTBI-related impaired consciousness.11 If an incorrect diagnosis and treatment were provided, military personnel may return to duty while still experiencing symptoms that were not properly treated, thereby increasing their risk for continued morbidity associated with the untreated condition. This misidentification may, in part, influence the residual PTS symptoms associated with prior exposure to blast mTBI reported in the literature.12 A better approach to differentiating PTSD from mTBI is warranted, as there is currently no objective method for extricating the effects of one from the other.

Researchers have reported that magnetic resonance spectroscopic imaging (MRSI) may be sensitive to detecting localized neuronal injury in the hippocampus associated with blast-related brain injury.13 Previously, we demonstrated that in veterans with a history of mTBI due to blast with self-reported persistent memory impairment, significant declines in the ratios of N-acetyl aspartate to choline (NAA/Ch) were seen in the anterior portions of both hippocampi more than 1 year postinjury in comparison to age-matched control subjects.14 These changes are consistent with both neuronal loss and impairment (NAA is synthesized only in neuronal mitochondria)15 and axonal injury (Ch is known to be increased
following TBI\textsuperscript{16} and multiple sclerosis).\textsuperscript{17} The changes seen were independent of the presence or absence of PTSD and depression. In an extension study that added a small group of veterans with PTSD without a history of mTBI, de Lanerolle et al\textsuperscript{13} reported that hippocampal levels of NAA/Ch were generally intermediate between control and veterans with blast-related mTBI. These preliminary findings suggest that MRSI may offer a useful neuroimaging tool to help differentiate mTBI from PTSD and corroborate current clinical delineations based on functional assessments.

The primary purpose of the current study was to identify clinical and imaging biomarkers that might help differentiate veterans with blast mTBI and PTSD from those with PTSD only. To that end, we compared MRSI findings from the hippocampal regions of military veterans with blast mTBI and PTSD to those with PTSD only. We expected that NAA/Ch ratios would be lower in veterans with blast mTBI and PTSD compared to PTSD. A secondary purpose of this study was to examine the relationship of cognitive and neuromotor impairment to MRSI findings across the two groups. We expected a negative correlation between cognitive and neuromotor impairment and NAA/Ch ratios.

**METHODS**

**Participants**

U.S. military veterans between the ages of 18 and 65 were identified for the study through clinics within the Veterans Administration Pittsburgh Healthcare System (VAPHS) and referred for screening by their providers. All participants were deemed cognitively able to provide informed consent and provided written informed consent before the initiation of study procedures. Exclusion criteria included 1) ferrous metal in their body and 2) claustrophobia or a body mass index in excess of 34. A total of 50 veterans from the VAPHS met study exclusion criteria and were enrolled in the study. However, only 35/50 (70\%) met the following study inclusion criteria: 1) a complete 7T MRI scan, 2) complete cognitive and neuromotor testing, and 3) a current medical diagnosis (per VA medical records) of either: a) blast mTBI and PTSD (n = 23), or b) PTSD only (i.e., no current, symptomatic mTBI) (n = 12). The sample ranged in age from 25 to 65 years. A comparison of select characteristics for the two groups is provided in Table I. Please note that TBI history data presented in Table I are self-reported. As expected, significantly more of the participants in the blast mTBI/PTSD group reported a history of mTBI than in the PTSD-only group. No other group differences were noted.

**Measures**

**Magnetic Resonance Spectroscopic Imaging**

Whole-plane MRSI data including NAA and Ch were acquired at 7T using a single-slice acquisition (10 mm thickness) with 24 × 24 encodes over the entire field of view of 240 × 240 mm. A transceiver array (Resonance Research Inc)\textsuperscript{18} was used for both transmission and reception. Extra-cerebral lipid contamination was minimized using radio frequency-based outer volume suppression.\textsuperscript{19} A moderate echo time (40 milliseconds) was used to minimize contributions from macromolecule and amino acid resonances underlying the NAA and creatine (Cr) resonances. With TR = 1,500 ms, the acquisition time was 14.5 minutes. To maximize spectral quality, a very high order shim insert\textsuperscript{20} was used, which provided third- and fourth-degree shims. The MRSI data from each channel of the transceiver array was phased and scaled using a gradient echo reference image before summation of the individual channel data.\textsuperscript{21} To minimize the effects of metabolite heterogeneity along the hippocampal formation, the data were reconstructed using single voxel reconstructions.\textsuperscript{19} Briefly, hippocampi were manually outlined on T1-weighted images, and a midline running along the length of the hippocampi was calculated. The aqueduct was identified and spectra starting at the level of the aqueduct were reconstructed using voxel shifting methods to provide six loci along the hippocampal formation. The loci were numbered from 1 (most posterior) to 6 (most anterior) with three loci anterior to the aqueduct and two posterior (Fig. 1). Spectra were analyzed by fitting the data with three Gaussian lines (NAA, Cr, Ch) and calculating the ratio Ch/NAA from the integrated areas of Ch and NAA.\textsuperscript{22}

Spectra were reconstructed from six loci spanning the hippocampi. The right and left hippocampi were manually delineated (green and red tracings) and a midline (blue) was automatically calculated. The aqueduct was identified (white cross hairs) and spectra were reconstructed in 9-mm increments starting at the intersection of the midline and the white horizontal line moving along the midline (three anterior and two posterior). The loci were numbered 1 (most posterior) to 6 (most anterior).

**PTSD Assessment**

The PTSD Checklist (PCL) was used to assess self-reported PTSD symptoms. The PCL encompasses 17 items and requires individuals to indicate on a scale of 1 (not at all) to 5 (extremely) how much each item bothered them during the past month. The PCL yields a PTSD symptom score that ranges from 17 to 85. Higher scores represent higher levels of PTS symptomology.
Cognitive and Neuromotor Assessments

Cognitive assessments included 1) Wechsler Adult Intelligence Scale-IV Vocabulary (WAIS-IV Vocabulary) test, which measures a subjects’ ability to explain the meaning of words; 2) Wechsler Test of Adult Reading (WTAR), which measures intelligence using word pronunciation; 3) Wisconsin Card Sorting Test (WCST), which measures executive function using a card sorting task based on numbers, form, and color categories; 4) Delis-Kaplan Executive Function (D-KEFS) verbal fluency test, which measures letter and category fluency; and 5) Trails A and B, which measure attention and task switching during a timed scanning task.

Neuromotor assessments included 1) grooved pegboard test, which measures manipulative dexterity in a timed dominant and nondominant hand pegboard task and 2) the Vestibular/Ocular Motor Screening (VOMS) tool, which measures vestibular and oculomotor symptoms and impairment following the performance of smooth pursuit, horizontal and vertical saccades, vestibular ocular reflex (VOR), and visual motion sensitivity (VMS) tasks.23 The VOMS also includes a measurement of near-point convergence (NPC) distance in centimeters.

Procedures

Institutional Review Board and Office of Research and Development approvals were obtained from the VAPHS and University of Pittsburgh before study initiation. Following these approvals and written informed consent procedures, all participants completed the cognitive and neuromotor tests in a dedicated interview room at the VAPHS. The cognitive and neuromotor tests were administered by a member of the research team trained in neuropsychological test administration in the following order: PCL-C, Trails A&B, WAIS-IV, Grooved Peg Board, VOMS, WCST, D-KEFS, and WTAR. Other clinical tests were administered subsequently as part of an ongoing National Institutes of Health-funded study (R01NS081772), but are not reported here. The cognitive and neuromotor components required approximately 2.5 to 3 hours to complete. After participants completed the cognitive and neuromotor portions of the study, they were given a 60-minute break and then escorted from the VAPHS to the University of Pittsburgh Magnetic Resonance Research Center for the 7T MRSI scan. The MRSI scan process required an additional hour to complete. The combined total testing time for each participant was approximately 4.5 to 5 hours. Participants were compensated for their participation.

Analysis

We compared the mTBI/PTSD- and PTSD-only groups on NAA/Ch across the six different loci of both the right and left hippocampus and both cognitive and neuromotor assessments using a series of general linear models (GLMs) with Bonferroni correction. The GLMs were used to account for the unequal cell sizes. We also conducted a series of Spearman correlations among ratios of NAA/Ch in the hippocampus and cognitive and neuromotor assessments. Spearman correlations were used because the data were not normally distributed. All statistical analyses were performed using Statistical Analysis System, version 9.3 (SAS Institute, Inc, Cary, North Carolina), with \( p < 0.05 \).

RESULTS

Mean Comparisons of Blast mTBI/PTSD and PTSD Groups

The results of the GLMs with Bonferroni correction supported significant differences in the ratio of NAA/Ch in three left anterior loci of the hippocampus (Fig. 2). Specifically, the ratio of NAA/Ch in the left anterior locus (Number 4) was significantly \( (F = 8.24, p = 0.007) \) lower in the blast mTBI/PTSD group \( (M = 1.20, SD = 0.22) \) compared to the...
PTSD-only group (M = 1.41, SD = 0.20). Likewise, the ratio of NAA/Ch in the left anterior locus (#5) was significantly (F = 7.74, p = 0.009) lower in the blast mTBI/PTSD group (M = 1.14, SD = 0.17) compared to the PTSD-only group (M = 1.29, SD = 0.11), and the ratio in the left anterior locus (#6) was significantly (F = 6.82, p = 0.01) lower in the blast mTBI/PTSD group (M = 1.01, SD = 0.17) compared to the PTSD-only group (M = 1.15, SD = 0.10). With regard to cognitive function, the blast mTBI/PTSD group (M = 29.91, SD = 9.69) scored significantly (F = 5.99, p = 0.02) lower on the WAIS-IV vocabulary than the PTSD-only group (M = 38.00, SD = 8.39). No other significant group differences were noted.

Correlations Between Cognitive and Neuromotor Impairment and MRSI Findings

Results of a series of Spearman correlations for NAA/Ch in the hippocampus and neuromotor impairments did not support any significant relationships between the WAIS-IV vocabulary and arithmetic, WATAR, WCST, D-KEFS verbal fluency, Trails A and B, and NAA/Ch ratios for any of the loci of the hippocampus. Results of a series of Spearman correlations for NAA/Ch in the hippocampus and neuromotor impairment revealed several significant relationships (Table II). Specifically, higher symptom provocation on smooth pursuit, saccades, VOR, and VMS were all correlated with decreased NAA/Ch ratios across multiple left and right loci of the hippocampus. Slower performance on the grooved peg board test nondominant hand was correlated with decreased NAA/Ch in the left #2 locus (Spearman’s rho = −0.36, p = 0.04) of the hippocampus. The average NPC distance from the VOMS was not related to any of the NAA/Ch ratios in the loci of the hippocampus.

DISCUSSION

To our knowledge the current study is the first to compare MRSI findings and cognitive and neuromotor impairments between military veterans with blast mTBI/PTSD and PTSD only. Consistent with previous findings13 and our hypothesis, we reported lower NAA/Ch ratios in anterior regions of the hippocampus in the blast mTBI/PTSD group compared to the PTSD-only group. The current findings support significant differences in the left anterior hippocampus, whereas, the previous findings, although reporting reductions in NAA/Ch in the anterior loci (#5, #6) of both hippocampi, did not reach statistical significance. This discrepancy may be due to differences between the two veteran populations studied. The previous blast mTBI and PTSD population included symptomatic patients who had subjective cognitive complaints

<table>
<thead>
<tr>
<th>NAA/Ch Loci</th>
<th>Smooth Pursuits</th>
<th>Horizontal Sac</th>
<th>Vertical Sac</th>
<th>Horizontal VOR</th>
<th>Vertical VOR</th>
<th>VMS</th>
<th>NPC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Right #6Left #6</td>
<td>−0.62***−0.38*</td>
<td>−0.61***−0.37*</td>
<td>−0.61***−0.34*</td>
<td>−0.58***−0.39*</td>
<td>−0.58***−0.32</td>
<td>−0.07−0.33</td>
<td>0.170.10</td>
</tr>
<tr>
<td>Right #5Left #5</td>
<td>−0.41*−0.34*</td>
<td>−0.42*−0.35*</td>
<td>−0.42**−0.35*</td>
<td>−0.58***−0.34*</td>
<td>−0.58***−0.31</td>
<td>−0.21−0.25</td>
<td>−0.080.01</td>
</tr>
<tr>
<td>Right #4Left #4</td>
<td>−0.36*−0.16</td>
<td>−0.38*−0.17</td>
<td>−0.36*−0.14</td>
<td>−0.38*−0.16</td>
<td>−0.38*−0.10</td>
<td>−0.23−0.36*</td>
<td>−0.060.14</td>
</tr>
<tr>
<td>Right #3Left #3</td>
<td>−0.27−0.14</td>
<td>−0.26−0.15</td>
<td>−0.24−0.12</td>
<td>−0.22−0.10</td>
<td>−0.22−0.10</td>
<td>−0.35−0.45**</td>
<td>0.050.10</td>
</tr>
<tr>
<td>Right #2Left #2</td>
<td>−0.45**−0.14</td>
<td>−0.44*−0.15</td>
<td>−0.43**−0.11</td>
<td>−0.42**−0.10</td>
<td>−0.42**−0.07</td>
<td>−0.29−0.40*</td>
<td>−0.100.16</td>
</tr>
<tr>
<td>Right #1Left #1</td>
<td>−0.43−0.27</td>
<td>−0.43*−0.29</td>
<td>−0.42**−0.28</td>
<td>−0.46**−0.32</td>
<td>−0.46**−0.25</td>
<td>−0.12−0.17</td>
<td>−0.190.15</td>
</tr>
</tbody>
</table>

NPC, near point convergence; Sacc, saccades; VMS, visual motion sensitivity; VOR, vestibular ocular reflex. *p < 0.05; **p < 0.01; ***p < 0.001.
Clinical and Magnetic Resonance Spectroscopic Imaging Findings

(e.g., memory problems). In contrast, participants in the current study were not selected based on subjective cognitive complaints. In addition, the current sample completed testing at a later time with respect to blast mTBI than the previous study. The PTSD-only population was older in the deLanerolle et al study, (43 ± 9 years), as compared to the current study (38 ± 14 years). However, the blast mTBI/PTSD groups in the two studies were nearly identical in age (deLanerolle et al sample = 34.1 ± 8 years; current sample = 34 ± 9 years). Also, given that NAA levels in the hippocampus are known to decline with age in normal aging, additional research focusing on both younger and older age groups is warranted to examine the effect of ageing on the current findings.

The blast mTBI/PTSD group in the current study scored worse on the WAIS-IV vocabulary test than the PTSD-only group. However, surprisingly, our findings did not support any differences in neuromotor impairment between the groups. Researchers have previously reported cognitive deficits, including poor performance on visual memory and reaction time, as well as increased vestibular impairment and related symptoms following blast mTBI. However, some researchers have reported no differences on neuropsychological tests following blast-related mTBI between soldiers with and without PTSD, suggesting the presence of PTSD may not impact neuropsychological performance. The current study expands this previous research by exploring the impact of PTSD on neuropsychological function in those with a history of blast-related mTBI. The findings supported differences in the WAIS-IV vocabulary test that suggest that military veterans with blast mTBI/PTSD performed significantly worse than those with PTSD only. However, no other cognitive or neuromotor differences were noted in the current study, supporting the previous research that did not find an exacerbating effect of PTSD on cognitive functioning following mTBI. The current findings with regard to limited cognitive and no neuromotor impairment differences between the groups highlights the importance of combining neuroimaging such as MRSI with clinical assessments to help distinguish the effects of blast mTBI/PTSD from those of PTSD only.

In the current study, we also explored the relationship between cognitive and neuromotor impairment and MRSI findings. Our results supported correlations between decreased performance on the D-KEFS verbal fluency and category tests and decreased NAA/Cr ratios in the left anterior hippocampus. Our findings also supported strong correlations between decreased NAA/Ch ratios in the left anterior hippocampus and increased vestibular impairment and symptoms as measured by the VOMS. The magnitude of the correlations between the VOMS items and MRSI findings was moderate to high, suggesting a more robust relationship than for cognitive outcomes. These preliminary findings suggest that MRSI findings may be useful in corroborating vestibular impairment and may hold promise to image improvement over time. Future research should include repeated assessments of MRSI and vestibular impairment during the recovery process.

Strengths and Limitations

The current study examined two groups of patients from the VAPHS that represent a substantial challenge for VA and other clinicians to assess and subsequently treat. As a result, the current findings may help to inform clinical practice with these patients. The study also combined MRSI with cognitive and neuromotor data to allow for direct comparison of neuroimaging and clinical findings. The addition of assessments of vestibular/oculomotor impairment and symptoms, which had not been examined in this population before, represents another strength of the current study. We assumed that all participants completed the cognitive and neuromotor assessments in an honest and accurate manner and with full effort. To ensure genuine effort on the tests, we used the Test of Memory Malingering, the results of which indicated all participants gave adequate effort during the study. The sample was relatively small and included veterans from a wide age range. The sample may have also included veterans with undocumented co-morbid diseases and disorders such as substance abuse and personality disorders that might have influenced the results. In addition, very few females were included in the study, limiting the generalizability of the findings to male veterans. Another limitation of the study is that details regarding the nature of the blast mTBI, including severity of symptoms at the time of injury, were not available due to limitations in the information that was provided in veterans’ medical records. A final limitation of the study is that the MRSI data were limited to the hippocampal regions, and as such, our inferences about the effects of blast mTBI and PTSD are limited to this region of the brain.

CONCLUSION

A reduced ratio of NAA/Ch, particularly in left anterior regions of the hippocampus, combined with certain neuromotor impairments may be useful in differentiating military personnel with blast mTBI and PTSD from those with PTSD only. Neuromotor and, in particular, vestibular/oculomotor impairments and symptoms were correlated with decreased ratios of NAA/Ch, suggesting that functional impairment in these areas may correspond to localized injury in the hippocampus following blast injury. Additional research is needed to determine if the combined MRSI, neuromotor, and cognitive measurements in the current study can be used to monitor recovery and the effectiveness of interventions following exposure to blast mTBI as in epilepsy. Regardless, MRSI, a novel neuroimaging technique, again supports a biochemical difference in hippocampal regions of the veterans with blast mTBI/PTSD. Specific clinical tests such as cognitive and neuromotor assessments may also serve as indirect assessments of brain dysfunction related to blast mTBI and PTSD.
In the future, the use of MRSI in combination with clinical assessments of neuromotor and cognitive function may help inform more accurate diagnosis of mTBI in specific veteran patient populations especially those exposed to blast forces.

ACKNOWLEDGMENTS

This research was supported in part by a grant to the University of Pittsburgh to Dr. Hetherington and Dr. VanCott from the National Institutes of Health/National Institute of Neurological Disorders and Stroke-NIH-R01NS081772; and by a grant to the University of Pittsburgh to Dr. Kontos from the National Institute on Deafness and Other Communication Disorders (1K01DC012332-01A1); and a grant to Dr Jullie W. Pan from the National Institutes of Health/National Institute of Biomedical Imaging and Bioengineering NIH R01EB011639.

REFERENCES

Validation of Laboratory Animal and Surrogate Human Models in Primary Blast Injury Studies

Namas Chandra, PhD, PE*; Aravind Sundaramurthy, PhD*; Raj K. Gupta, PhD†

ABSTRACT  Blast-induced neurotrauma has affected more than 300,000 service members. It is important to understand the effect of single and repeated shock-blast wave exposures on the neuropsychological behavior of soldiers, to offer them better protection, diagnostics, and treatment. Preclinical animal models and helmet design studies on human surrogate models have relied on the use of compression gas–driven shock tubes. Traditional shock tubes are so simple that if not carefully designed and operated, the test results can easily introduce detrimental artifacts clouding the conclusions. In this work, we present live-fire test results of an instrumented human surrogate head–neck model and compare with the data obtained in a carefully designed shock tube. We present various features incorporated in the shock tube design that led to better fidelity between live-fire and laboratory shock-blast conditions. The effect of specimen placement, choice of driver gas, pressure and volume of driver, end-plate conditions, and measurement techniques all determine the successful replication of live-fire loading conditions. These parameters become more important when conducting animal testing as the totality of loading will dictate the injury severity and type which ultimately will determine the mechanisms of blast-induced neurotrauma and hence their prevention and treatment strategies.

INTRODUCTION
Blast-related mild traumatic brain injury has been identified as the signature injury of the recent conflicts in Iraq and Afghanistan; an estimated 320,000 or 19.5% of all U.S. service members deployed there have symptoms related to blast-induced neurotrauma (BINT) which accounts for over 92% of all battlefield injuries. Although invisible to the naked eye, BINT is reported to cause debilitating changes in mood, thought, and behavior; medical symptoms associated include migraine, headaches, insomnia, blurred vision, dizziness, vertigo, tinnitus, nausea, and vomiting with exertion. Other manifestation of BINT include memory and concentration problems, verbal and written language problems, emotional liability and depression, fatigue, light and noise intolerance, anxiety, and irritability. It is postulated that all these medical outcomes can arise from a single exposure or multiple low-level exposures; simply feeling the blast wave is sufficient to cause injury.

During an explosion, the high energy release rate compresses the surrounding air and the outer dome of air expands with an instantaneous jump in velocity (exceeding acoustic velocity of the air), density, and pressure, creating a shock front-blast wave pulse. This shock front-blast wave pulse is strictly characterized by a sharp rising blast overpressure (BOP) with positive duration followed by underpressure (less than atmospheric pressure) with negative time duration. In its simplest form, it is represented by a planar Friedlander-type pressure–time profile. A typical explosion can cause primary (pure shock), secondary (penetration by shrapnel), tertiary (inertial acceleration–deceleration), and quaternary (fire, toxic gases) when the subject is near the epicenter. When the subject is at a greater standoff distance, then blast wave can cause either primary or/and tertiary injuries. However, since tertiary injuries are clearly visible, researchers mostly focus on the primary injury where the service member is subjected to blast wave without any gross motion.

Recently, an extensive research effort has been initiated toward the understanding of the mechanisms of primary blast injury from blast exposures. Current postulated mechanisms of BINT are direct transmission, skull flexure, thoracic surge, acceleration–deceleration, and cavitation. Experimental techniques using animal models and human surrogates and numerical simulations are being conducted to understand the origin of BINT. The outstanding questions are the following:

(1) How effective are the personal protective system (body armor, helmets, eye wear, and hearing protection) in combating BINT?
(2) Are the injury mechanisms in blast TBI same or different from that of blunt (tertiary) and ballistic (secondary) TBI?
(3) What are the acute (primary mechanical injury) and chronic (secondary neurobiochemical cascades) induced injury? How are these related to the cognitive, somatosensory, motor, and behavior outcomes observed in service members?
(4) What is the effect of repeated low-level blast exposures experienced in training and during combat? Is there a similarity between these repeated exposures and that observed as chronic traumatic encephalopathy in athletes?
To answer any or all of these questions, we can use actual clinical data and longitudinal medical outcome of service members; unfortunately, the history of exposure is not known and further, pure blast cases are far fewer. This leads us to conduct preclinical studies using animal models to examine biological outcomes, and human head–neck surrogates for helmet protection studies. Though actual exposure to live-fire conditions using known strengths of explosives is possible, it is neither economical nor practical. This leaves researchers to turn to laboratory tests; the use of compressed gas–driven shock tubes has emerged as the clear choice among various research groups. Shock tubes have been used since the 1890s for a variety of other purposes (combustion, aerodynamics, detonation). They comprise a compressed gas section (breech or driver) separated by membranes to a long driven section. When the gas pressure is increased, the membrane bursts (can be manually punctured as well) and the high-pressure gas compresses the atmospheric gas in the driven section, which generates a blast wave. A recent review article shows that there are numerous research groups investigating BINT using a variety of shock tubes with many variations in the design of the tube, animal location, holders and pressure–impulse–pulse generation, and measurements. Since the method for imparting the load on the animal varies with different laboratories, the biomechanical loading experienced by the animal varies. Therefore, it is difficult to compare the results leave alone correlating with the live-fire conditions. The purpose of this article is to measure pressure–duration–impulse conditions of a human head–neck surrogate under live-fire conditions and examine under what experimental conditions those live-fire data can be replicated in the laboratories. Further, what are the key parameters in the design and operation of the shock tube that are needed to be monitored to assure the fidelity?

ConWep (Conventional Weapon Effects) is a computer simulation program that was developed by U.S. Army Corps of Engineers which can accurately numerically simulate a variety of conventional weapons effect calculations including an assortment of air blast routines, fragment and projectile penetrations, breach, cratering, and ground shock. The program is capable of calculating the pressure–time–impulse profile experienced by a subject who is at a standoff distance, when a spherical charge of a given explosive (e.g., C4 or TNT) produces the shock blast when the charge is suspended in the air or on the ground or buried below the ground. Some of the commonly used improvised explosive devices and their strengths are shown in our previous work. Figure 1 shows the schematic of an explosion in the range of 2 to 100 kg and BOP of 70 to 350 kPa. The figure shows that for a given explosive strength, BOP decreases and duration increases with increase in distance.

**FIGURE 1.** BOP-time variation as a function of charge strength and standoff distance. Beyond 5 m distance, the wave can be idealized as planar compared to a human head. For mild to moderate ranges based on 2 to 100 kg of C4 and 1 to 100 m distance, BOP of 30 to 450 kpa and duration of 3 to 7 milliseconds can be considered as the right range for laboratory replication. bmTBI, blast related mild traumatic brain injury.
from the explosion. Compared to the radius of hemisphere of 5 to 10 m, the human head occupies less than 1% of the circumference ensuring that the shock wave in the laboratory needs to be planar. Thus, the shock tube design should be able to generate a Friedlander-type blast wave with 70 to 350 kPa BOP and duration of 3 to 7 milliseconds and that wave shall remain planar. Further, there should be no other artifacts arising from wall reflections (e.g., specimen itself, instruments, end conditions-open or closed) which will drastically change the biomechanical loadings on the animal and hence the results.

In this study, the pressure-acceleration response of a head–neck human surrogate realistic explosion-resistant dummy (RED) head was evaluated in two different conditions; (1) in the field (live-fire) and (2) in the laboratory shock tube with the same set of 13 sensors, 11 pressure gauges, and 2 accelerometers. The RED head represents the anthropometry of 50th percentile male soldier in the tri-services. In the field, four pounds of C4 was used as the explosive to produce blast waves that interacted with the instrumented RED head facing the epicenter of the blast; the same surrogate was then tested inside the shock tube. In both cases, 11 different pressure measurements on the surface of the RED head and linear acceleration at the center of gravity (CG) of the surrogate were measured. Finally, the pressure and acceleration data obtained from field experiments were compared with pressure and acceleration data obtained from the shock tube experiments to determine the efficacy of the shock tube when simulating primary blast injury conditions in a laboratory.

**METHOD**

**Field Experiments and RED Head Preparations**

To produce air blast (i.e., without any ground reflection) four pounds of C4 explosive charge was placed at 2,438 mm from the ground, which produces an incident pressure of 25 psi which was measured 280 mm in front of the head (Fig. 2A). The goal was to expose the RED head to a simple Friedlander-type blast wave. The RED head was instrumented with 11 sensors (102B06) (PCB Piezotronics, Inc, Depew, New York) and two linear accelerometers (Endevco 7270A) (Meggitt Sensing Systems, Irvine, California) mounted to a tri-axial block fixed to the CG of the RED head, which measures the accelerations along the x and z directions (Fig. 3). Out of the 11 pressure sensors, five sensors were instrumented along the mid-sagittal plane along with the remaining six (three on each side) on the RED head (Fig. 3A). All the pressure sensors used in the experiments were calibrated using Rankine–Hugoniot equations of state ahead of time in separate tests. In the live-fire experiments, the RED head was positioned at a radial distance of 2,794 mm from the charge with its anterior side facing the epicenter of the blast (Fig. 2A). Accelerometers were mounted at the CG of the RED head. The head was then attached to Hybrid III neck and mounted on base. The checkered board behind the head form was also placed at 2,794 mm to observe the propagation of the shock front in the high-speed video. The field experiment was repeated 10 times.

**Shock Tube Experiments**

The shock tube shown in Figure 2B comprises of driver, transition, and driven sections as well as a catch tank. The driver section has a diameter of 203 mm and its length can be varied between 67 and 1,210 mm. Transition section is an adapter that connects the circular driver section to the square driven section and has a length of 1,829 mm. The driven section has a 711.2 mm×711.2 mm (28 in × 28 in) square cross-section and a length of 8,661 mm. Finally, catch tank absorbs the kinetic energy of the exiting jet from the shock tube and does not play any role in modifying the profile. The driver and driven sections are separated by frangible mylar membranes. When the high-pressure
gas (helium or nitrogen) is pumped into the driver section, the membranes separating the driver and the driven section rupture resulting in the release of the high-pressure gas. This high-pressure gas expands into the driven section compressing the atmospheric air resulting in the formation of the blast wave. A detailed description of the design and operation of the shock tube can be seen in our previous work.11

The instrumented RED head was fixed inside the driven section at 2,096 mm downstream from the end of the driver section facing the blast wave. For this experiment, helium was used as the driver gas with a driver length of 143 mm. Fourteen mylar membranes with each having 0.254-mm thickness were used between driver and driven sections. This shock tube configuration resulted in an incident pressure of 28 psi, which was measured 137 mm in front of the head. These experiments were repeated 11 times.

RESULTS

RED Head Blast Wave Interaction in the Shock Tube and the Field Blast

The mechanics of the blast wave RED head interactions can be studied by monitoring the pressure on midsagittal plane of the head. Figure 4 shows the experimentally measured peak pressures on five locations along the midsagittal plane of the RED head for both shock tube and field tests. Peak overpressure corresponding to sensor 1 (between eyebrows) is 80 and 70 psi with pressure amplification from incident pressure of 2.85 and 2.8 for shock tube and field tests. The pressure starts to decrease from sensor 1 to sensor 4 with least pressure recorded in sensor 4 for both cases. At sensor 4 (just behind the crown of the head), the peak pressure is reduced to 17.5 and 20 psi for shock tube and field tests. However, at the location of sensor 5 (rear side of the head), the observed pressure is higher than sensor 4 for both shock tube and field experiments due to shock travel along the side compare to the top.

Comparison of Pressure Between Free Field and Shock Tube

Though peak pressures as shown in Figure 4 can be of interest, the actual profile of pressure-time is more illustrative of the flow field simulation. We observed significant correlation between the overall pressure time profiles on all the 11 sensors (Fig. 5A shows representative comparisons of sensors 1
FIGURE 5.  (A) Comparison of pressure profile between live-fire and shock tube experiments. First column shows live-fire and laboratory data on front sensor 1; second column shows data on the side of the dummy head for Sensor 9. See the variations in the profile of the pulse for the same shock-blast loading conditions. (B) Comparison of the normalized overpressure between live-fire and shock tube experimental data. The peak pressure (and the pulse) varies as the blast passes from the front to the back.
and 9). Since there is a difference in the exact magnitude of peak overpressure between field and shock tube experiments due to the difference in incident pressures, we normalized the peak overpressure ($P_{o}$) for all the 11 sensors from both field and shock tube experiments with their corresponding incident pressures ($P_{in}$) (Fig. 5B).

Comparison of Linear Acceleration Between Live Fire and Shock Tube

If we are concerned only with the primary shock loading, then the specimen can be fixed to the structure allowing no motion. If the RED head is allowed to rotate, then the linear and rotational acceleration will be dictated by the applied force as well as the stiffness of the Hybrid III neck as attached to the fixture. Figure 6 shows the comparison of the maximum positive and maximum negative acceleration along x and z directions between shock tube and field experiments. From the data, we observe a maximum positive acceleration of 150 and 145 g for shock tube and field experiments along x direction and 153 and 125 g for shock tube and field experiments along z direction. We observe a maximum negative acceleration of −205 and −110 g for shock tube and field experiments along x direction and −150 and −100 g for shock tube and field experiments along z direction. It is to be noted that although the maximum positive acceleration matches within experimental standard deviation, the maximum negative acceleration is higher in the laboratory compared to the field.

DISCUSSION

Mild to moderate primary blast TBI is believed to be caused by shock-blast wave pulse that interacts with the service members during improvised explosive device explosions. The pulse needs to be of Friedlander type with BOP ranging from 70 to 350 kPa and duration from 3 to 7 milliseconds with no other jet winds or reflection/expansion artifacts. The specimen (animal models or human surrogate) needs to be fixed and not allowed to toss around or ejected which are indicative of tertiary loading. These requirements arise from observations from live-fire experiments and should be replicated in the laboratories. In this work, we studied pressure at different points and acceleration at the CG of the instrumented head in both field experiment and shock tube.

Overall, there is a good match in BOP peak values for all the 11 sensors as shown in Figure 5B. Figure 5A shows the pressure–time–pulse at two sensor locations (sensors 1 and 9) with remarkable similarity. Sensor 1 is located on the forehead and faces the incident shock-blast pressure head on. The sensor data show that the Friedlander shape is maintained here with higher values of peak pressure due to reflection effects. Once the shock front interacts with the anterior part of the surrogate, it diffracts and envelopes the side and top of the head. During this traverse, the pressure starts to decrease between sensors 1 and 4. For the first 3 sensors, the decrease is because of the change in the incidence angle with sensor 1 being perpendicular to the flow of the blast wave resulting in the maximum pressure, whereas with sensors 2 and 3, the angle of incidence starts to change with sensor 3 being parallel to the flow of blast wave. Furthermore, a significant flow separation is observed (sensors 4 and 7); which results in the reduction of overpressure below the incident pressure. The blast wave traversing the head and the blast wave traversing the neck reunite at the back of the head. This reunion causes an increase in pressure on the backside of the head (sensor 5). These observations are similar to those reported by Ganpule et al.12 There was a good agreement between the normalized pressure on all the sensor locations for both field and shock tube experiments (Fig. 5B). Therefore, it can be concluded that the physics of

FIGURE 6. Comparisons of linear acceleration for bare head in free field and shock tube.
the blast-shock wave is captured in the shock tube as the blast wave traverses the dummy head.

We observed good agreement between the maximum positive accelerations recorded along both x and z directions; however, we observed a substantial difference in the maximum negative values of acceleration. The maximum positive acceleration is caused because of the loading of the oncoming blast wave. However, in the case of shock tube, there is more displacement of the head due to the buildup of air in front of the face (due to constraining from shock tube walls) and because the head is mounted on a H3 neck (neck acts like a spring), it springs back causing an increased negative acceleration. This may not have significant effect as the main loading occurs because of the direction of pressure transmission as well as the initial motion. In case of animal models, since the animals are constrained to simulate primary loading, tertiary blast injury will have less effect.

As the planar shock wave approaches a human head specimen, different types of shock-structure interaction takes place depending on the orientation of the surface with respect to the shock flow. For example, in sensor 1 mounted directly on the forehead, the shape of the shock pulse is maintained similar to that of Friedlander wave; however, the magnitude of peak pressure increases due to high reflection ratio. Conversely, we observe oscillations in the waveform recorded in sensor 9, we hypothesize that this is due to the combination of direct interaction of the undisturbed blast wave from the front as well as the diffracted wave from the facial features. The pulse shape changes completely and is captured in the laboratory as well. Figure 5B shows that the pressure profile varies from $\frac{P_0}{P_{in}}$ of around 3 in the front to about 0.8 at the rear sensor 11. Here, $P_0$ represents the peak overpressure measured by the sensor, whereas $P_{in}$ is the...
Validation of Laboratory Animal and Surrogate Human Models

Incident pressure measured before the shock interacts with the RED head. As the laboratory data replicates accurately all the 11 sensors (Fig. 3), we can say with confidence that the laboratory shock tube conditions are identical to that of the live-fire conditions. The fact that the shape of the profile measured in the live-fire experiments matches the laboratory data ensures that the physics of the shock flow and its interaction with the structure are accurately captured in the laboratory. A number of experiments (not reported here) shows that such replication will not occur if the specimen is moved near the exit or kept outside. Similarly, if the size of the specimen (size of the RED head is 9 in length and 7 in width) is more than about 10% of the blast tube cross-sectional area, the secondary reflection from the specimen itself will alter the loading profile and hence the result. Furthermore, area of blockage also increases the dynamic pressure on the specimen. An analytical calculation presented by Needham and his colleagues showed that for cases with blast wavelength (defined approximately as the product of blast duration and speed of sound) longer than the characteristic length of the target species (diameter of the sphere), blockages of 10, 20, and 30% increased the dynamic pressure by 27, 64, and 113%. However, according to Wortman and Lottero, in cases where blast wavelength is approximately same as the characteristic length of the target species, a 20% blockage is acceptable because of the rapid decay of dynamic pressure. This is the very reason that we conduct a 20% blockage is acceptable because of the rapid decay of dynamic pressure. However, according to Wortman and Lottero, in cases where blast wavelength is approximately same as the characteristic length of the target species, a 20% blockage is acceptable because of the rapid decay of dynamic pressure. This is the very reason that we conduct a 20% blockage is acceptable because of the rapid decay of dynamic pressure. This is the very reason that we conduct small animal (rodent) experiment in the smaller (9 in × 9 in) shock tube while conducting human surrogate, swine, and postmortem human specimen (human cadavers) experiments in the larger tube. Numerical simulations of the shock-structure interactions on the RED head conducted by our group also validate this conclusion.

Computational models on shock-head interaction has shown that if \( P_{in} \) is the incident BOP, then on the forehead of a front facing head the peak pressure is about 2 to 3 \( P_{in} \), on the top and side (parallel to flow) the value is about \( P_{in} \), and less than \( P_{in} \) in the rear. Similar ratios are also observed for impulse values (integral of shock pressure and duration).

There are a number of design and operation variables that are carefully controlled for the faithful replication of shock profile within the BOP duration range as shown in Figure 1. Though we do not claim that these features are the only way to achieve the fidelity, these will be very helpful in understanding the effect of certain features on the profile. In any case, any deviation from these basic principles will likely to cause minor to major variations in the shock profile and hence the experimental data. Some of these key features are shown in Figure 7. The picture depicts the smaller shock tube (9 in × 9 in × 20 ft) specifically meant for live rodent testing as the specimen will occupy less than 10% of the opening.

In our previous work, we have performed analysis on the different features of our shock tube and its influence on the blast wave it produces, results of which is not shown for the sake of brevity. The ability to shape the pulse is determined by a number of factors, one of which is the choice of driver gas (Fig. 7B), helium, nitrogen, or compressed air. It should be recognized that the medium that loads the specimen is not these species but the atmospheric air that is present in the tube when the gas bursts. The bursting gas only acts as a fast-moving acoustic piston pushing the air and increasing air velocity to shock wave speed. Because of very low molar mass (0.004 kg/mol) and high acoustic speed (956 m/s), helium produces a much higher BOP compared to nitrogen (0.028 kg/mol and 352 m/s velocity). Thus in our studies, we used helium for achieving higher BOP and lower duration and nitrogen for lower BOP and higher duration.

The other variable that can be controlled is the driver (breech) length. Longer driver length results in longer positive pulse duration. When membranes separating the compressed gas and atmospheric air burst, two waves are generated: a compressive wave travel toward exit pressurizing the air and a tensile expansion wave travel in the opposite direction toward the closed end of the driver. When the latter reaches the closed end, the expansion wave is reflected and travels back into the driver section as a tensile wave (faster due to compressibility effects) and catches up with the first shock wave. Thus, this reflected tensile wave erodes the shock front producing the decreasing part of the Friedlander wave. Thus, if we delay this tensile wave by making the breech length longer, then we can produce longer duration in the shock pulse. The lengths are altered in practice through cylindrical inserts.

Burst pressure \( (P_B) \) is the pressure when the membranes burst (Fig. 7D) and directly dictates the magnitude of BOP. Higher \( P_B \) is obtained by stacking multiple membranes of varying thickness ranging from 0.001 to 0.01 in. Though the sheets are traditionally obtained from the same vendor using the same manufacturing process, still batch to batch variations are noticed. This has necessitated calibration of \( P_B \) vs. membrane thickness when starting a new lot. Other designs have used metallic membranes with fracture initiated manually through mechanical means to control \( P_B \).

The end of the shock tube can be fully open, fully closed, or partially open as shown in Figure 7F. As the shock exits the tube, if the end is open, it will produce a release (tensile) wave that will now travel backward toward the driver and produce a sharp negative pressure (tensile) component. When the end is fully closed, the oncoming shock pulse gets fully reflected as an additional pressure pulse redoubling the BOP depending on the location of measurement. A partially open exit combines the effect of negative pressure due to opening with that of the positive pressure due to reflection. Depending on the desired pulse shape, we alter the exit opening.

The test section shown in Figure 7G can be moved anywhere along the driver length as the units are made in multiple sections with flanged joints and bolted assembly. The square section with shock-proof glass window permits the video capture of shock-structure interactions. With adequate lighting
provided from the front and top, we can capture up to 10,000 frames per second enough for capturing the motion of the specimen. It should be noted that while securing the specimen, the supporting structure should be aerodynamically smooth to avoid any reflection artifacts from the specimen or the structure.

Through these features, we have been able to not only recreate the Friedlander pressure–time–impulse profile similar to the field, but also vary the BOP-duration in the desirable range of 70 to 350 kPa and 3 to 7 milliseconds to simulate mild and moderate blast TBI conditions.

CONCLUSION

In this article, we have outlined the need to replicate the live-fire blast pressure–duration pulse in laboratory test conditions and delineated the BOP and duration ranges. We have shown that the pulse has to be planar and be of Friedlander type without any artifacts that can alter the test conditions. We have also described the key parameters involved in replicating live-fire loading conditions in a shock tube. By careful planning and controlling the key parameters in shock tube studies, we believe that the preclinical test results for animal studies as well as personal protection design studies can be conducted with translatable results under field environments.

ACKNOWLEDGMENTS

NC acknowledges the financial support provided by the U.S. Army Research office with contract number W911NF-08-1-0483 and laboratory members for some of the laboratory shock field data.

REFERENCES

ABSTRACT  Objectives: To examine the incidence and the etiology of corneal and corneoscleral injuries in the setting of combat ocular trauma, and to determine what effect these injuries have on overall visual impairment from combat ocular trauma. Methods: Retrospective, noncomparative, interventional case series, analyzing U.S. service members who were evacuated to the former Walter Reed Army Medical Center (WRAMC). Primary outcome measures were types of corneal injuries, length of follow-up at WRAMC, globe survival, and anatomical causes of blindness. Secondary outcome measures included surgical procedures performed, use of eye protection, source of injury, and visual outcomes. Results: Between 2001 and 2011, there were 184 eyes of 134 patients with corneal or corneoscleral injuries. The average age was 26 years (range, 18–50); 99.3% were male, 31.9% had documented use of eye protection. The average follow-up was 428.2 days (3–2,421). There were 98 right-eye and 86 left-eye injuries. There were 169 open-globe and 15 closed-globe injuries with corneal lacerations occurring in 73 eyes with injuries to Zone I. Most injuries were attributable to an intraocular foreign body (IOFB; 48%), followed by penetrating (19.6%) and perforating (16.3%) injuries. The most common presenting visual acuity was hand motion/light perception (45.7%), yet, at the end of the study, visual acuity improved to 20/40 or better (40.8%). The majority of injuries in eyes with visual acuity worse than 20/200 involved the cornea and retina (58%). Injuries solely to the cornea accounted for only 19% of all injuries sustained. Conclusions: Ocular injuries in military combat have led to significant damage to ocular structures with a wide range of visual outcomes. The authors describe corneal and corneoscleral injuries in combat ocular trauma by classifying injuries by the anatomical site involved and identifying the main source of decreased visual acuity. In combat ocular trauma, corneal or corneoscleral injuries are not the sole etiology for poor vision. A cohesive approach among multiple ophthalmic subspecialties is needed when treating combat ocular trauma.

INTRODUCTION

Combat ocular trauma has been a topic of interest for the Department of Defense, given the relatively high incidence of ocular trauma experienced by recently deployed service members. It has been shown that roughly 13% of evacuated U.S. soldiers suffered a significant ocular injury between 2002 and 2007 in either Operation Iraqi Freedom (OIF) or Operation Enduring Freedom (OEF).1–3 Military service members suffering from ocular trauma may experience significant challenges performing daily activities.4 The authors have previously reported on the overall causes of visual impairment in service members involved in OIF/OEF.1 Additionally, there have been multiple studies documenting the severity and types of injuries, as well as the primary reasons for poor visual acuity in the patients evacuated to the former Walter Reed Army Medical Center (WRAMC).1,3,5,6 This is the first study to examine corneal and/or corneoscleral injuries in combat ocular trauma in order to discern what role these injuries play in long-term visual prognosis.

METHODS

This study was a retrospective, noncomparative, case series analyzing corneal or corneoscleral injury data from the Walter Reed Ocular Trauma Database (WROTD) as part of an ongoing Institutional Review Board–approved protocol of service members evacuated to WRAMC, Washington, DC, from 2001 through 2011. Initially, ophthalmologists at combat support hospitals (CSHs) in Iraq and Afghanistan evaluated and, when necessary, surgically treated eye injuries. After CSH evaluation/treatment, patients were routinely taken to Germany for further stabilization and finally to WRAMC. Some soldiers were taken directly to Germany, bypassing the local CSH, and on rare occasions, soldiers were taken directly from a local CSH to WRAMC. Once at WRAMC, patients were followed during their convalescence until their return to duty, permanent change of station, or discharge into the Veteran’s Affairs medical system. As part
of the protocol, investigators have maintained an eye trauma database, the WROTD, since 2007. The database includes all battle-related ocular trauma patients seen at WRAMC between 2001 and 2011, when WRAMC closed. The WROTD includes data gathered from different sources to include inpatient records, outpatient records, transfer summaries, and operative reports. A standardized data collection sheet was used for data entry into a structured database program for analysis (SPSS software version 21.0; SPSS, Inc., Chicago, Illinois). Data collection included variables from the time of injury to the last follow-up at WRAMC. Primary outcome measures include types of corneal injuries, length of follow-up at WRAMC, globe survival, and anatomical causes of blindness. Secondary outcome measures include surgical procedures performed, use of eye protection, source of injury, and visual outcomes. 

Inclusion criteria for this study were U.S. military personnel injured during OIF/OEF in Iraq or Afghanistan who suffered corneal or corneoscleral injuries. Exclusion criteria were nonbattle injuries, injuries sustained outside a combat zone, and ocular trauma in any non-U.S. soldier or civilian. The Birmingham Eye Trauma Terminology as well as the guidelines established by the Ocular Trauma Classification Group were used to classify the injuries. Injury to Zone I is defined as an injury to cornea or limbus; Zone II is an injury to the anterior 5 mm of sclera; Zone III is an injury more than 5 mm posterior to limbus. The initial and final best-corrected visual acuities (BCVA) were converted from Snellen acuity to vision grades for statistical analysis. Grade 1 was visual acuity of 20/40 or better. Grade 2 was 20/50 to 20/200. Grade 3 was BCVA of 19/200 to 1/200 (count fingers). Grade 4 was hand motion (HM)/light perception (LP), and finally grade 5 was no light perception (NLP).

The initial visual acuity from the CSH was used, when data were available. In patients who were intubated or not responsive at the initial encounter, visual acuity was measured once the soldier was able to communicate. Often, visual acuities were not recorded before initial globe repair. In these cases, visual acuity upon arrival to WRAMC was recorded. When possible, soldiers were evaluated in the ophthalmology clinic. When soldiers were immobile, the initial examination was conducted bedside in the surgical intensive care unit. Visual acuity was measured at the bedside using a Rosenbaum pocket screener at 14 inches. Visual acuity in the clinic was measured using a standard Snellen eye chart using best spectacle acuity when possible. Complete data were not available for all patients (i.e., documented use of eye protection, initial visual acuity, etc.). In some instances, data were not recorded at the time of initial evaluation or before eye surgery, and in others, the patient’s level of consciousness prevented obtaining this information.

**RESULTS**

Between 2001 and 2011, there were 650 soldiers evacuated from Iraq and Afghanistan to WRAMC with 890 eyes sustaining injuries. Of those, there were 184 eyes (20.7%) of 134 patients suffering a corneal or corneoscleral injury. Patients were examined while receiving their care at WRAMC and the duration of follow-up ranged from 3 to 2,421 days. Patient demographics and injury statistics are described in Table I. Table II shows the visual acuity outcomes in the study, the initial median BCVA was grade 4 (HM/LP), whereas the final median BCVA was grade 1 (≥20/40). Among the total cohort population, there were 169 open-globe injuries and 15 closed-globe injuries, as shown in Figure 1. Table III summarizes the location and types of ocular injuries sustained by our study population.

Among the studied population, a majority of eyes underwent zero or one surgical procedure (n = 137, 74.5%),

---

**TABLE I.** Overall Summary of the Study Cohort

| Age (years) | 26 (18–50) |
| Gender | 99.3% male/0.7% female |
| Days of Follow-up | 428.2 (3–2,421) |
| Eye Protection | 44 (31.9%) |

---

**TABLE II.** Initial and Final Visual Acuity during the Study for the Entire Cohort

<table>
<thead>
<tr>
<th>Grade</th>
<th>Initial Visual Acuity</th>
<th>Final Visual Acuity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grade 1 (≥20/40)</td>
<td>19 (10.5%)</td>
<td>75 (40.8%)</td>
</tr>
<tr>
<td>Grade 2 (20/50–20/200)</td>
<td>12 (6.5%)</td>
<td>27 (14.7%)</td>
</tr>
<tr>
<td>Grade 3 (19/200–1/200)</td>
<td>23 (12.5%)</td>
<td>14 (7.6%)</td>
</tr>
<tr>
<td>Grade 4 (HM/LP)</td>
<td>84 (45.7%)</td>
<td>30 (17.3%)</td>
</tr>
<tr>
<td>Grade 5 (NLP)</td>
<td>35 (19%)</td>
<td>36 (19.6%)</td>
</tr>
</tbody>
</table>
28 eyes (15.2%) underwent two surgical procedures, and 13 eyes (7.1%) had three or more surgical procedures while at WRAMC. Figure 2 shows final visual acuity grades based on the total number of surgeries a patient underwent. The most common surgical procedure was pars plana vitrectomy \((n = 102, 55.4\%\); 68 eyes required an intraocular tamponade, as shown in Table I. To achieve retinal reattachment, silicone oil was used in 28/184 eyes (15.2%), octafluoropropane \((C3F8)\) in 26/184 eyes (14.1%), and sulfur hexafluoride \((SF6)\) in 14/184 eyes (7.6%). Focusing on visual acuities in eyes that underwent retinal detachment surgery, no eyes had a final BCVA of grade 1, 10/46 (15.2%) were classified as grade 2, 6/46 (13%) were grade 3, 16/46 (32.1%) were grade 4, and 17/46 (37%) were grade 5. There were 26 penetrating keratoplasties (PKPs) performed in the study cohort. All were performed in patients with open-globe injures, and all were performed after theater evacuation. If there was a globe injury while in theater, the ophthalmologist would perform an exploration and, if needed, an initial open-globe repair. Five of the 26 PKPs were performed at the time of definitive repair at WRAMC, whereas 12 PKPs were delayed and performed later for multiple reasons (i.e., decrease in initial inflammation to decrease likelihood of graft failure because of severe inflammation). Nine of the PKPs failed leading to a repeat PKP; four of 26 patients with PKPs have missing data about graft survival. In 21 eyes (11.4%), corneal glue was used to repair a laceration. There were a total of 28 eyes (15.2%) that underwent enucleation, either primarily or secondarily. Primary enucleation was performed on nonsalvageable eyes at the time of initial repair. Secondary enucleation was performed within the first 14 days of injury on blind and painful eyes or on eyes that were primarily repaired in theater, yet remained NLP, to reduce the risk of sympathetic ophthalmia.

Visual outcomes for the entire study cohort are shown in Table II. When comparing final BCVA of eyes with corneoscleral versus purely corneal laceration, 14/74 (18.9%) eyes were classified as grade 1 in eyes with corneoscleral laceration versus 64/127 (50.4%) eyes with purely corneal laceration. Seven of 74 (9.5%) eyes were classified as grade 2 in the corneoscleral laceration group, whereas 19/127 eyes (15%) were in the corneal laceration group. There were 6/74 (8%) with grade 3 BCVA with corneoscleral lacerations, and 9/127 (7%) with purely corneal lacerations. Sixteen of 74 eyes (21.6%) were classified as grade 4 in corneoscleral laceration group and 21/127 eyes (16.5%) in the corneal laceration group. In the corneoscleral laceration group, there were 31/74 (41.9%) eyes classified as grade 5 versus 14/127 eyes (11%) classified in the corneal laceration group. Almost 41% of the eyes in the study had final BCVA of 20/40 or better regardless of the type of injury the eye sustained.

Of the 184 eyes, 80 eyes of 78 patients had a final BCVA of worse than 20/200. Of these, 78 resulted from open-globe injuries and the other two from closed-globe injuries. Table IV further classifies the injuries by zones involved in this subgroup. IOFB \((n = 26/80, 32.5\%\), perforating injury \((n = 22/80, 27.5\%\), globe rupture \((n = 15/80, 18.8\%\), and penetrating injury \((n = 14/80, 17.5\%\) were the primary causes of

### Table III. Corneal and Corneoscleral Injuries Sustained in the Studied Cohort

<table>
<thead>
<tr>
<th></th>
<th>Corneal Abrasion</th>
<th>Corneal Laceration: Full Thickness</th>
<th>Full Thickness Laceration Length (mm)</th>
<th>Corneal Laceration: Partial Thickness</th>
<th>Glued Cornea</th>
<th>Corneal Foreign Body</th>
<th>Corneoscleral Laceration</th>
<th>Scleral Laceration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open Zone 1</td>
<td>5</td>
<td>73</td>
<td>3.19 ± 2.89</td>
<td>36</td>
<td>10</td>
<td>20</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Zone 2</td>
<td>4</td>
<td>9</td>
<td>6.83 ± 5.2</td>
<td>13</td>
<td>2</td>
<td>12</td>
<td>31</td>
<td>5</td>
</tr>
<tr>
<td>Zone 3</td>
<td>2</td>
<td>31</td>
<td>12.23 ± 10.9</td>
<td>12</td>
<td>9</td>
<td>10</td>
<td>41</td>
<td>19</td>
</tr>
<tr>
<td>Total</td>
<td>11</td>
<td>113</td>
<td></td>
<td>61</td>
<td>21</td>
<td>42</td>
<td>76</td>
<td>26</td>
</tr>
<tr>
<td>Closed Zone 1</td>
<td>34</td>
<td>N/A</td>
<td>N/A</td>
<td>1</td>
<td>0</td>
<td>66</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Zone 2</td>
<td>3</td>
<td>N/A</td>
<td>N/A</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Zone 3</td>
<td>11</td>
<td>N/A</td>
<td>N/A</td>
<td>1</td>
<td>0</td>
<td>23</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>48</td>
<td>N/A</td>
<td>N/A</td>
<td>2</td>
<td>0</td>
<td>95</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

Numbers in the table represent all the injuries eyes sustained. Multiple injuries may be present in a single eye.
injury resulting in final BCVA of worse than 20/200. The use of eye protection in this cohort was 27.5%; however, data were missing for 33.8% of patients. Figure 3 shows the etiologies of the poor visual outcome in these 80 patients.

There were 15/184 (8%) eyes with closed-globe injury. In this subgroup, initial BCVA was as follows: 20/40 or better in 2/15 (13.3%) eyes, 4/15 (26.7%) with BCVA 20/50 to 20/200, 3/15 (20%) with BCVA 19/200 to 1/200, 3/15 (20%) with BCVA HM/LP, and 1/15 (6.7%) with BCVA NLP. Final BCVA in the closed-globe cohort, 11/15 (73.3%) eyes had final BCVA of 20/40 or better, 1/15 eye (6.7%) each with final BCVA of 20/50 to 20/200, 19/200 to 1/200, and HM. The majority of eyes in the closed-globe cohort recovered with the exception of 2/15 eyes (13.3%) with final vision worse than 20/200. Associated findings in the closed-globe eyes included vitreous hemorrhage (8/15 [53%] eyes), commotion retinai within the macula (3/15 [20%] eyes), chorioretinal rupture (3/15 [20%] eyes), intraretinal/subretinal hemorrhage within the macula (2/15 [13.3%] eyes), and 2/15 (13.3%) eyes with a macular hole.

**TABLE IV.** Injury Classification Within the Cohort of Eyes With Final Vision of Worse Than 20/200

<table>
<thead>
<tr>
<th>Zone</th>
<th>Open-Globe</th>
<th>Closed-Globe</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zone 1</td>
<td>14 (17.5%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Zone 2</td>
<td>14 (21.3%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Zone 3</td>
<td>47 (58.8%)</td>
<td>2 (2.5%)</td>
</tr>
<tr>
<td>Total</td>
<td>78 (97.5%)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>2 (2.5%)</td>
</tr>
</tbody>
</table>

<sup>a</sup>There is data missing for 5 eyes in the Open-Globe category, with regard to which zone an injury was sustained in.

**DISCUSSION**

Combat ocular trauma has been a significant cause of morbidity for military service members. Corneal and corneoscleral injuries have been implicated as a major cause of decreased vision and subsequent decrease in quality of life for service members. This subset analysis examines military service members evacuated from Iraq or Afghanistan with ocular trauma that involved corneal and corneoscleral injuries.

Among ocular injuries sustained by U.S. service members during OIF/OEF, roughly 21% involved the cornea and/or sclera. A majority were attributable to a projectile leading to an IOFB, which required a pars plana vitrectomy, the most common surgical procedure performed, to remove the foreign body. Given the studied population was patients with corneal and corneoscleral injuries, most of the injuries were classified as Zone I. The majority of scleral injuries extended to Zone III; however, injury classification was difficult, given the number and complexity of the wounds presented in Table III.

Visual acuity improved over the course of the study, with a majority of eyes having final vision of 20/40 or better. Eyes that had an open-globe injury had a higher likelihood of having visual acuity worse than 20/200 (odds ratio = 3.55),...
as well as having an injury to Zone III (odds ratio = 11.67). Initial visual acuity measurement was not always accurate, given the setting of severe trauma and the use of narcotic medications. Nevertheless, isolating patients with final BCVA worse than 20/200, showed that a majority of the patients sustained multifactorial injuries (injuries to the cornea and the retina or choroid), as seen in Figure 3. A vast majority of injuries that led to poor vision, defined as worse than 20/200, were open-globe Zone III injuries (n = 47 [58.8%]). A majority of the interventions performed were procedures within Zone III (vitreo-retinal surgeries). Eyes that had more than one surgery had worse visual outcomes, as seen in Figure 2. Although, eyes with multiple surgeries were a minority within the study, requirement for multiple surgeries would imply significant ocular damage, which would result in decreased visual acuity.

Of the 80 eyes with poor visual outcomes, 19% were caused by purely corneal injuries. Specifically examining corneal interventions, nine of 26 PKPs failed. Corneal graft rejection is as a result of a complex immune response when the hosts’ immune system recognizes the transplanted corneal antigen as foreign, leading to the initiation of the immune response cascade. In the setting of trauma, a host’s immune system is overactive. Another risk for PKP rejection is the increased incidence of corneal neovascularization, which brings host immune cells in contact with the graft, increasing the rejection rate. In the studied population, the major cause of PKP failure was stromal and endothelial graft rejection. One lesson learned during the course of the study was that there was a higher chance of a PKP failure, if a PKP is performed at the time of multiple other surgeries (i.e., pars plana vitrectomy, lensectomy). PKP survival rate increased when PKP was performed later, giving sufficient time to reduce inflammation generated by other surgeries.

Eyes remain vulnerable to severe injury despite the use of eye protection. The documented use was 31.9%, whereas the unrecorded rate of eye protection use in this cohort was 18.8%. As discussed in previous studies, recording the use of eye protection was not the first priority of the first responder or the squad mate at the time of injury and it was difficult for the injured service member to recall whether eye protection was worn at the time of injury. Although the use of eye protection affords a level of protection in preventing ocular injury, additional work is needed to promote its use and optimize the armor.

This study has numerous limitations, including the loss of patients to follow-up and incomplete records. The WROTD was incomplete as many patients sustained injuries before the implementation of electronic health records. Furthermore, limited ophthalmic examinations were conducted in many of these cases because of the dynamic and highly pressured environment of multisystem injury or mass casualty situations. Because many patients were lost to follow-up, it is possible additional PKPs eventually failed. Despite the numerous limitations, the study demonstrates the limitations of our current surgical capabilities given the severity of combat ocular trauma.

CONCLUSION

Corneal or corneoscleral injuries are not the sole etiology for poor vision (Fig. 3). The analyzed data showed a return to useful vision in most patients with purely corneal injuries; however, when there was Zone III involvement, the prognosis for the return to useful vision was guarded. As has been shown before, preventing ocular trauma is more critical than attempting to restore vision after suffering an injury. Trauma in a combat environment may be inevitable, and for those who sustain an injury, multisurgeon collaboration is needed between vitreo-retinal, corneal and neuro-ophthalmic subspecialties in the hope of improving vision.14–16

REFERENCES

Frequency Responses to Visual Tracking Stimuli May Be Affected by Concussion

Jun Maruta, PhD*; Eugene Jaw, BA*; Peter Modera, BA*; Umesh Rajashekar, PhD*; Lisa A. Spielman, PhD*; Jamshid Ghajar, MD, PhD†

ABSTRACT Human visual tracking performance is known to be reduced with an increase of the target’s speed and oscillation frequency, but changes in brain states following a concussion may alter these frequency responses. The goal of this study was to characterize and compare frequency-dependent smooth pursuit velocity degradation in normal subjects and patients who had chronic postconcussion symptoms, and also examine cases of acutely concussed patients. Eye movements were recorded while subjects tracked a target that moved along a circular trajectory of 10° radius at 0.33, 0.40, or 0.67 Hz. Performance was characterized by the gain of smooth pursuit velocity, with reduced gain indicating reduced performance. The difference between normal and chronic patient groups in the pattern of decrease in the gain of horizontal smooth pursuit velocity as a function of the stimulus frequency reflected patients performing more poorly than normal subjects at 0.4 Hz while both groups performing similarly at 0.33 or 0.67 Hz. The performance of acute patients may represent yet another type of frequency response. The findings suggest that there may be ranges of stimulus frequencies that differentiate the effects of concussion from normal individuals.

INTRODUCTION

We have proposed the utility of a circular visual tracking paradigm in assessing attention impairments associated with concussion.1–3 The circular stimulus target motion is continuous and predictable, and thus is suitable for examining how well a patient can sustain visuomotor synchronization, an attention-dependent feat.4 Although it is known that human visual tracking performance is reduced with an increase of the target’s speed and oscillation frequency,5–7 it is not clear whether this effect is impacted after a concussion. Here, we focus on smooth pursuit characteristics using parameters commonly examined in the literature.5–8 The primary outcome of interest was the smooth pursuit velocity gain, which is the ratio of the smooth pursuit eye velocity to the target velocity. The goal of this study was to characterize and compare frequency-dependent performance degradation in normal subjects and patients who had chronic post-concussion symptoms. We also present cases of acutely concussed patients for comparisons.

METHODS

The sample consisted of 140 normal control subjects (66 male), 43 patients with chronic symptoms that persisted after a concussion (21 male), and 5 patients with acute postconcussion symptoms who were recruited within 2 weeks after the injury (2 male). Recruitment was limited to individuals 18 to 55 years of age and with at least 12 years of education. To be considered for the group of patients with chronic postconcussion symptoms, a subject must have had persistent problems believed to have resulted from a concussive head injury that occurred between 90 days to 5 years before the date of testing, had post-traumatic amnesia at the time of injury, and had a loss of consciousness not exceeding 24 hours in the period following the injury. Of the 43 patients with persistent postconcussion symptoms, 11 had a history of multiple concussions. None of the subjects were pregnant and had a history of drug or alcohol abuse, neurological or psychiatric illness, or seizure. The experimental protocol was approved by the institutional review board of Weill Cornell Medical College, and all subjects provided informed consent before testing.

The eye movement recording and analysis procedures are detailed in another publication.4 Briefly, the test stimulus was a target that moved clockwise in a circular trajectory and was implemented on an integrated stimulus presentation-eye tracking apparatus (EyeLink CL; SR Research, Ontario, Canada; Fig. 1). We used target rotation frequencies of 0.33, 0.4, and 0.67 Hz while the radius of the target trajectory was kept at a 10° visual angle. The corresponding target speeds were 21, 25, and 42°/s, respectively. The frequencies were within the range in which progressive degradation of performance typically occurs in normal individuals, whereby an increase in the stimulus frequency is accompanied by a reduced gain and an increased phase lag.5–7 The order in which the three stimulus frequencies were presented was randomized for each subject. For each stimulus frequency, the semiautomated testing sequence that included recorded instructions and calibration lasted approximately 5 minutes.

The target and eye data were stored for offline analyses. Eye velocities were obtained by two-point differentiating eye position data. Saccades were identified and isolated from smooth pursuit eye movement. Desaccaded eye velocity traces were fit with sine curves of the frequency of the circular movement of the target using fast Fourier transformation, and
gain and phase of smooth pursuit velocity were calculated relative to the target velocity.

The effect of stimulus frequency on smooth pursuit characteristics of normal subjects and patients with persistent postconcussion symptoms was examined with a repeated measures analysis of variance (ANOVA). To account for the known increase in interindividual variability in smooth pursuit gain with increased stimulus frequency, the mean and the standard deviation of the combined sample of control subjects and patients with persistent postconcussion symptoms were used to derive a z score for each subject. Repeated measures ANOVAs were applied to both raw performance scores and z scores with the α level set at 0.05. The Geisser–Greenhouse correction was applied when departure from sphericity was found. An interaction between group and frequency was not found in this analysis (horizontal: $F_{1,181} = 4.44, p = 0.037$), but not for vertical ($F_{1,181} = 1.65, p = 0.20$). An interaction between group and frequency was not found. Although the ANOVA did not detect a change in the relationship between the two groups across frequencies, the main effect of group in horizontal gain largely reflected patients with chronic symptoms performing more poorly than normal subjects at 0.4 Hz ($t_{54.05} = 2.46, p = 0.017$) rather than at 0.33 ($t_{181} = 1.36, p = 0.18$) or 0.67 Hz ($t_{181} = 1.64, p = 0.10$).

The difference in the effects of stimulus frequency on smooth pursuit velocity gain was further characterized by converting individual gain values to z scores calculated for each stimulus frequency (Fig. 3). With this procedure, we considered the relative positions of individual performance within the combined group so as to discount the change in interindividual variability related to stimulus frequency (Fig. 2). By converting individual gain values to z scores, the effect of frequency was eliminated by the nature of the ANOVA model. For the z score of horizontal gain, the overall group difference was significant ($F_{1,181} = 4.76, p = 0.030$), and there was an interaction between unequal variances with pooled standard deviations and degrees of freedom was applied.

**RESULTS**

Even within the small frequency range tested, changes in the gain of horizontal and vertical smooth pursuit velocity were significant (horizontal: $F_{1,38,250.44} = 177.24, p < 0.001$; vertical: $F_{1,64,296.00} = 414.59, p < 0.001$), the gain decreasing with increasing stimulus frequency (Figs. 2A and 2B). The overall group difference was significant for horizontal gain ($F_{1,181} = 4.44, p = 0.037$), but not for vertical ($F_{1,181} = 1.65, p = 0.20$).

**FIGURE 2.** Smooth pursuit velocity frequency response. Small circles and crosses represent individual control subjects and patients with persistent postconcussion symptoms, respectively. Means are indicated by large circles and crosses. (A, B) Smooth pursuit velocity gain against stimulus frequency. (C, D) Smooth pursuit velocity phase against stimulus frequency.
group and stimulus frequency ($F_{2,362} = 3.12, p = 0.045$), consistent with the finding that patients with chronic symptoms performed more poorly than normal subjects at 0.4 Hz rather than at 0.33 or 0.67 Hz. For the $z$ score of vertical gain, neither an overall group difference ($F_{1,181} = 1.75, p = 0.187$) nor an interaction between group and stimulus frequency ($F_{1.90,343.65} = 2.07, p = 0.13$) was found.

The change with stimulus frequency in the phase of horizontal smooth pursuit velocity modulation was small but significant ($F_{1.59,286.81} = 5.38, p = 0.009$), with the modulation of smooth pursuit velocity tending to lag that of the target velocity as the stimulus frequency increased (Fig. 2C). Neither an overall group difference ($F_{1,181} = 1.75, p = 0.19$) nor an interaction between group and stimulus frequency ($F_{1.59,286.81} = 1.24, p = 0.29$) was found. For the phase of vertical smooth pursuit velocity modulation, there was not a significant effect of stimulus frequency ($F_{1.60,289.74} = 0.19, p = 0.78$), group ($F_{1,181} = 0.33, p = 0.57$), or interaction ($F_{1.60,289.74} = 0.58, p = 0.52$). For both control subjects and patients, the modulation of vertical smooth pursuit velocity consistently led that of the target velocity (Fig. 2D), indicating that horizontal and vertical smooth pursuit velocities were differently controlled.

The performance of five patients with acute postconcussion symptoms was compared on a case-by-case basis with those of control and chronic patient groups. However, some trends were noted. Four of the acute patients had a reduced vertical smooth pursuit velocity gain at all three stimulus frequencies, and the same patients had a reduced horizontal gain at 0.67 Hz (Fig. 4). The phase characteristics of smooth pursuit velocity response of all acute patients were similar to those of control and chronic patient groups (not shown).

**DISCUSSION**

We found that frequency response characteristics of smooth pursuit performance were altered in patients with chronic postconcussion symptoms. Specifically, horizontal smooth pursuit velocity gain was reduced in patients for the middle stimulus frequency, but not for the lowest or the highest frequency. The difference in the frequency response characteristics may be explained as a ceiling effect at the lowest stimulus frequency for both patients and normal individuals, but that high gain bandwidths are expanded in normal individuals, albeit only to a point.

The ceiling effect at the lowest stimulus frequency was apparent from both the high average gains and reduced inter-individual variability, suggesting that at this frequency, the task posed few challenges for most of the subjects in either group. However, given the small frequency range tested in the present study, the ceiling effect was very likely incomplete. In
comparison, a study that utilized a horizontal pursuit paradigm with a wider frequency range found that normal subjects as well as patients with various cerebral lesions invariably attained a near-unity smooth pursuit velocity gain at 0.2 Hz. Their study furthermore suggested that some individuals were more resilient than others to an increase in the stimulus frequency, although nearing and exceeding 1 Hz, all individuals showed rapid falls in gains and phases. Our results were similar in that the highest frequency was challenging for both groups. A future study of concussion patients may merit examining smooth pursuit frequency response characteristics using a wider stimulus frequency range than used presently. Nevertheless, our data indicated that the horizontal smooth pursuit velocity gains of patients with chronic postconcussion symptoms started to fall off at a frequency at which most of their normal peers were able to maintain a high gain, yielding a statistical group difference at the middle frequency rather than at the highest frequency. Neural mechanisms that support high gain bandwidths need to be investigated.

An analogous alteration in the vertical smooth pursuit velocity response was not found. It has been suggested that there are separate mechanisms of control for horizontal and vertical tracking, and visual tracking is typically more accurate in the horizontal than vertical direction. We found that the modulation of vertical smooth pursuit velocity consistently led the modulation of the target velocity within the frequency range tested, whereas the modulation of horizontal smooth pursuit velocity showed a tendency to lag the modulation of the target velocity with an increase in stimulus frequency. This finding further highlights the difference in the implementation of horizontal and vertical visual tracking, and may explain why we found in patients with chronic postconcussion symptoms an alteration in the frequency response of horizontal smooth pursuit velocity but not in that of vertical smooth pursuit velocity.

The performance of a small sample of patients with acute postconcussion symptoms was examined on a case-by-case basis, but some trends emerged, which should be tested with a larger sample in future investigations. Specifically, patients with acute postconcussion symptoms indicated a faster falloff of smooth pursuit velocity gain with increased stimulus frequency. In addition, acute patients tended to show a reduced vertical gain relative to the horizontal gain at a low stimulus frequency, indicating that the interaction between the mechanisms for horizontal and vertical smooth pursuit may be altered in a manner that depends on the stimulus frequency. Thus, the performance of acute patients may represent yet another type of frequency response. If so, visual tracking performance degradation in patients with acute and chronic postconcussion symptoms reflect different alterations in attention from normal.

CONCLUSION
Changes in brain states following a concussion may induce different frequency responses to visual tracking stimuli. Using a circular visual tracking test paradigm, we found a group difference in the pattern of decrease in the gain of horizontal smooth pursuit velocity between normal individuals and patients with chronic postconcussion symptoms. Case analyses of patients with acute postconcussion symptoms also hinted an alteration in the frequency response. In assessing attention impairments, there may be ranges of stimulus frequencies that differentiate the effects of concussion from normal individuals.

ACKNOWLEDGMENTS
The authors thank Brain Trauma Foundation’s clinical research staff for assistance with data collection and maintenance. Funding was provided by CDMRP W81XWH-08-1-0646, and James S. McDonnell Foundation.

REFERENCES
Streamlining Participant Recruitment for TBI and PTSD Research Studies

Shristi Joshi, BA†; Kerri Dunbar, MA†; Patricia Taylor, MBA†; Kelli L. Sullivan, BA†; Mariam M. Afzal, BA†; Christina Song, BA†; Maulik Purohit, MD, MPH§; COL Michael J. Roy, MC USA (Ret.)†‡

ABSTRACT

Objectives: Recruitment of participants for traumatic brain injury (TBI) and post-traumatic stress disorder (PTSD) studies is a major challenge, causing delays in study timelines and even study failures. To address this challenge, the Center for Neuroscience and Regenerative Medicine (CNRM) Recruitment Core developed procedures for identification, screening, and referral of participants from screening studies to a broad range of TBI and PTSD studies. Methods: Participants were recruited from civilian hospitals, Military Treatment Facilities, and through various events and presentations. Enrolled participants were referred to other studies during initial enrollment, follow-up visits, or as new CNRM studies became active. A centralized online database was used to streamline the eligibility and referral process. Results: As of October 25, 2016, 1,040 enrolled participants from the two screening studies have been assessed for eligibility for active CNRM studies. Referrals have led to 197 total enrollments into other CNRM studies. Common reasons for exclusion from studies included age, date of injury, injury severity, contraindication to Magnetic Resonance Imaging, state of residence, and military status. Conclusion: Collaborative work with multiple disciplines and institutions, and the use of diverse media, was critical to augmenting participant enrollment, and significantly diversified the demographics of the participant population. Streamlining the referral process helps studies meet their timelines and target enrollment.

INTRODUCTION

Traumatic brain injury (TBI) and post-traumatic stress disorder (PTSD) are signature injuries of the wars that have been fought in Iraq and Afghanistan over the past decade. There is still much to learn about the prevention, diagnosis, treatment, and short- and long-term prognosis associated with each of these conditions. To improve the diagnosis and treatment of TBI and PTSD in the military, it is necessary to study individuals with a full range of symptom severity. Recruitment of participants for TBI and PTSD studies is a significant challenge that can cause delays in study timelines and even lead to study failures. Thus, success in clinical research is founded not only on the quality and innovation of the intervention, or the accuracy of measurements, but also by the ability to recruit and retain a sufficient number of research participants that characterize the target population.1

Participant recruitment has previously been identified as one of the most difficult aspects of the research process.2,3 Researchers frequently recruit participants solely from a single institution. Although this may be convenient, it can be difficult to find enough interested participants who meet the study criteria, and the resultant sample may not be representative of the intended study population. Unfortunately, low participant recruitment reduces the probability that the study will detect treatment differences between control and experimental groups, and attrition can pose a significant threat to internal and external validity.4 Diversifying recruitment efforts, however, can be time-consuming, tedious, and logistically difficult. Despite its importance to successful research efforts, there is a paucity of literature on the recruitment and retention process.4

The Center for Neuroscience and Regenerative Medicine (CNRM) is a federal intramural program that studies the full spectrum of TBI and PTSD, with a special focus on military-related injuries. To address the challenge of recruiting participants, the CNRM Recruitment Core developed screening studies to recruit civilian and military subjects, and to serve as an entry point for participation in other research protocols. Recruitment Core staff facilitate referrals for studies seeking to enroll participants more than 30 days from injury. The Recruitment Core also works in collaboration with the CNRM Acute Core, which facilitates referrals for participants within 30 days of injury. The Recruitment Core developed procedures and tracking tools for initial identification, screening, and referral of participants from these screening studies to a wide array of CNRM research studies. The centralized recruitment and referral process is designed to effectively recruit participants for studies, minimize burden on participants and study investigators, and identify common reasons for participants being excluded from studies.
METHODS
The CNRM screening studies, which use a multisite, multidisciplinary approach, recruit participants from civilian hospitals and Military Treatment Facilities (MTFs). Recruitment sites include Suburban Hospital, Washington Hospital Center, Virginia Commonwealth University Medical Center, the National Institutes of Health (NIH), Walter Reed National Military Medical Center, and Fort Belvoir Community Hospital. Recruitment is accomplished via fliers at outpatient clinics, physician referrals, direct recruitment at the Emergency Department, and recruitment tables. Varied recruitment methods at multiple sites provide a diverse sample of individuals, including healthy subjects and participants with both acute and chronic injuries. In addition to recruiting participants from these study sites, the Recruitment Core also promotes awareness of active studies via the CNRM website, recruitment booths at various events and conferences, and presentations to targeted audiences. CNRM also maintains an active social media presence, with daily postings on Facebook and Twitter to raise awareness of TBI and PTSD. Social media posts also include recent publications and news articles, with frequent suggestions to contact us if interested in participating in research. Interested participants can enroll by telephone or in person at the NIH Clinical Center or MTFs.

Written informed consent is obtained for all screening study participants. Participants who enroll by telephone are required to mail in a copy of the signed consent form. On initial enrollment, participants provide comprehensive information about their medical history, injury history, current medications, and mental health symptoms. Validated questionnaires collected include the PTSD Checklist, Civilian Version (PCL-C), Neurobehavioral Symptom Inventory, Ohio State University TBI Identification Method Short Form, Glasgow Outcome Scale, Patient Health Questionnaire, Short Form 36, Epworth Sleepiness Scale, Combat Exposure Scale, Satisfaction With Life Scale, and a review of the International Classification of Diseases 10th Edition (ICD-10) criteria for Postconcussive Syndrome. This information is updated during a 6-month follow-up interview, and during annual follow-up interviews thereafter. Participants enrolled through in-person visits at the hospital sites, NIH or MTFs, may complete additional procedures: structural and resting state functional magnetic resonance imaging, diffusion tensor imaging, a medical history and physical examination by a physician, a blood draw to store samples for biomarkers, and the NIH Toolbox neurocognitive assessment battery.

Independent of participant enrollment in this protocol, the CNRM Recruitment Core maintains a running list of active CNRM-funded and collaborative protocols and their respective inclusion and exclusion criteria. The Core is able to assess each participant’s eligibility for active studies by matching their individual characteristics (e.g., age, TBI history) against each set of inclusion and exclusion criteria.

Referral of participants to other CNRM studies involves three critical steps. The Core staff assess each participant’s eligibility for other studies during initial enrollment, follow-up, and ad hoc as new CNRM studies become available. After eligibility has been assessed, participants are contacted to gauge their interest in the studies for which they are eligible. Recruitment Core staff describe the relevant studies, their required procedures, and the time commitment to participants. Referrals are made only after participants actively express interest in being referred to additional studies. Referrals are not made if the Core is unable to get in touch with the participant or the participant declines to be referred to other studies. However, as long as a participant remains in the screening study, the Recruitment Core reassesses study eligibility based on new information obtained (e.g., new injuries, changes in medication), and attempts to gauge participant interest during follow-up visits.

The Recruitment Core has developed organizational tools to track the eligibility and referrals of participants. The

![FIGURE 1](image)

FIGURE 1. Eligibility and referral tracking log. This log is completed for each participant in the screening studies.
Recruitment Core’s tracking log (Fig. 1) has been instrumental in managing referral of participants and tracking enrollment outcomes for the referred studies. This log includes a list of all CNRM studies, participant eligibility status, reasons for participant exclusion from particular studies, and reasons that participants decline the study. The log also includes information about completed referrals and whether participants were enrolled in the referred study. This information is critical to understanding the challenges to each study’s recruitment, and it helps the Recruitment Core reassess eligibility during future visits. Participants may decline studies during one visit but express rekindled interest during subsequent assessments. For example, participants may decline a study if they are busy with work at initial contact, but during follow-up visits they may be more available and willing to participate in research. Detailed documentation allows Core staff to reassess eligibility and re-offer the study during follow-up visits. A centralized online database has been utilized to streamline the eligibility and referral process.

To minimize participant burden, and to maintain data integrity, participants are referred to only one study at a time. If participants are interested in multiple studies, referral to a second study will occur on either completion of, or exclusion from, the first study. Participants are referred to CNRM studies based on the following criteria: existence of an active CNRM approved protocol, participant preference, eligibility timeline, study priority, and professional judgment.

RESULTS

As of October 25, 2016, 1,167 participants have been enrolled in the two screening studies, of which 1,040 (89%) have been assessed for eligibility for active CNRM studies. Participants who were unable to be contacted successfully, or those who chose to decline all available studies, were not referred. Common reasons for declining participation in a study included lack of interest in the type of study being offered, number of study visits required, travel distance, insufficient compensation, and time commitment. Of the participants who were deemed eligible and expressed interest in being referred to other studies, 334 participants have been referred to a single study, 149 to two studies, and 80 to three or more studies. Thus far, referrals have led to 197 total enrollments in other CNRM studies: 140 participants in a single study, 24 in two studies, and 3 in three studies.

Referred participants were contacted by study teams to discuss next steps in the enrollment process. Most research teams requested additional information from participants prior to enrollment (e.g. previous medical records), and some also required an in-person study-specific evaluation to further determine eligibility. Participants who remained eligible following this additional screening were then successfully enrolled into the referred studies (Fig. 2).

To further understand the challenges to enrollment, as the percentage enrolled is relatively low, the Recruitment Core has identified some of the common reasons participants were being excluded from studies after being referred. Some participants were not eligible after review of the medical records or after in-person evaluation by the study teams. Reasons for exclusion could include current use of prohibited medication, medical history, injury history, and risk factors as determined by individual study teams. Core staff members iteratively review such reasons and revise inclusion and exclusion criteria accordingly in order to improve the accuracy of referrals.

A number of participants from the two screening studies were excluded from some or all of the active CNRM studies and therefore could not be referred. Some common reasons participants were excluded from different studies were age, date of injury, injury severity, contraindication to imaging, travel distance, and military status. However,
these participants remain in the Recruitment Core database and may be eligible for future CNRM studies as they become available.

**DISCUSSION**

Clinical trials currently spend considerable time trying to recruit participants, and frequently fall short of their enrollment goals. Recruitment challenges may lead investigators to extend the duration of studies, and often their power to draw evidence-based conclusions is diminished. The establishment of standardized recruitment procedures, to enable collection of a wealth of data from interested participants, could expedite recruitment of participants for studies by providing them with a list of eligible individuals. The implementation of standardized procedures for the early identification, screening, and referral of participants to studies focusing on TBI and PTSD, and tracking enrollment outcomes for those referred, have proven to be critical steps in facilitating the flow of participants from one study to the next. Although tracking tools have been effective in referring and successfully enrolling participants in other CNRM studies, the Recruitment Core will continue to try to improve the enrollment rate of those referred. On the basis of lessons learned from implementation of the centralized referral process, participants are more likely to be successfully contacted by study teams if there is a shorter time period between when the Core gauges the participant’s interest in the study and when the study teams contact the participants. The Recruitment Core’s goal is to refer participants as soon as they express interest, in order to increase the probability that they will enroll in the study. Maintaining contact with participants continues to be a challenge. TBI and PTSD participants, and particularly military service members, may relocate or change their contact information several times during the course of their participation in the screening studies. Thus, it is essential to make immediate referrals so that participants are not lost to follow-up.

Although the Recruitment Core screening studies have simplified the recruitment process, enrollment into active research studies continues to be a challenge. Eligibility assessment by study teams can be time consuming, with requests for supplemental documents to confirm exposures or other eligibility criteria often delaying the enrollment process, sometimes for as much as several months. Such obstacles have contributed to the relatively low enrollment rate of those referred. The Recruitment Core has made iterative improvements in its eligibility assessment process, as well as the referral process, and will continue to do so going forward. While prioritization of participant preference before referral to other studies is important, it is also essential to take into consideration study timelines and use professional judgment when making referrals. For example, referral to a study that requires only one-time participation first, followed by referral to a study that requires a 3-month duration of participation, would generally be preferable to the reverse order, especially when combined with improved tracking of referrals to ensure that target study staff exercise rapidity in both contacting the participant initially, and in notifying the Core on completion of their participation in order to facilitate referral to a second study.

**CONCLUSION**

The development of new studies that could accommodate participants who are more geographically remote, have chronic injuries, have contraindications to Magnetic Resonance Imaging, or are of older age, would most effectively maximize the participation of previously recruited individuals. Studies that use measures that are administered over the telephone or computer, e.g., would provide investigators with valuable data, which may otherwise not be captured, on participants who are unable to travel due to distance, physical disabilities, or lack of transportation. Availability of different types of studies, including observational, clinical, and rehabilitation studies will be important in providing a range of options for participants to choose from, and will be another crucial factor in the retention of participants in the screening protocols.

Collaborative work with multiple disciplines and institutions, combined with the use of modern recruitment methods, has proven critical to the augmentation of participant enrollment, and has significantly diversified the demographics of the participant population. Streamlining the referral process will help studies meet their timeline and enrollment targets. It will also allow investigators to focus primarily on science instead of requiring inordinate investment in participant recruitment. Thus, military and civilian research programs could benefit from a centralized recruitment process to simplify and expedite one of the greatest challenges in clinical research.

**REFERENCES**

Expanding Clinical Assessment for Traumatic Brain Injury and Comorbid Post-Traumatic Stress Disorder: A Retrospective Analysis of Virtual Environment Tasks in the Computer-Assisted Rehabilitation Environment

Marie M. Onakomaiya, PhD*§; Sarah E. Kruger, MS*; Krista B. Highland, PhD†‡; Paula N. Kodosky, DPT*; Marcy M. Pape, MPT, CLT, CBIS*; COL Michael J. Roy, MC USA (Ret.)†‡

ABSTRACT The objective of this study was to determine whether physical performance during virtual environment (VE) tasks in the Computer-Assisted Rehabilitation Environment (CAREN) could differentiate between service members (SMs) with a history of traumatic brain injury (TBI) with and without comorbid post-traumatic stress disorder (PTSD). Data were obtained by independent review of clinical notes, objective outcomes, and validated questionnaires from 214 SMs (208 males) with a history of TBI assessed in the CAREN from 2010 to 2015. Three preliminary VEs acclimatized patients to the CAREN: Balance Balls, weight shifting on a static platform (timed); Balance Cubes, step shifting with and without platform motion (timed); and Continuous Road, flat ambulation (self-selected speed). Multiple regression analyses revealed that patients with comorbid TBI-PTSD were significantly slower in completing the VE tasks than patients without PTSD. Logistic regression showed that the Balance Cubes VE without platform motion significantly predicted diagnostic category (i.e., no PTSD vs. comorbid PTSD). In conclusion, in SMs with a history of TBI, physical performance on the CAREN effectively distinguished those with comorbid PTSD, as their performance was significantly slower than SMs without PTSD. These results portray the potential of the CAREN as a novel assessment tool in SMs with a history of TBI.

INTRODUCTION

Mild traumatic brain injury (mTBI) and post-traumatic stress disorder (PTSD) are common and frequently comorbid in service members (SMs) returning from the conflicts in Afghanistan (Operation Enduring Freedom [OEF]) and Iraq (Operation Iraqi Freedom [OIF]), and are associated with functional impairment. Since 2000, over 330,000 SMs have been diagnosed with TBI, 82.4% of whom had mTBI.1 An estimated 15 to 44% of mTBI patients experience persistent (i.e., greater than three months postinjury) postconcussive symptoms such as memory impairment, headaches, and vestibular disturbances2–4 and one report found that up to 44% of returning SMs with mTBI met criteria for probable PTSD.5 There is evidence to suggest that the traumatic nature of the circumstances that lead to a TBI, as well as the structural and physiological changes that result from damage to the brain, contribute to the etiology of psychiatric comorbidities linked to TBI such as PTSD and depression.6–8 However, the relationship between chronic postconcussive symptoms and psychological symptoms is complex.

PTSD can contribute to the persistence of postconcussive symptoms,9 and PTSD morbidity in the context of a TBI may differ from that of either condition alone.10 In addition, a history of mTBI may interfere with recovery from PTSD; e.g., a study in a cohort of Vietnam-era veterans found that mTBI was associated with a current diagnosis of PTSD, and markedly impaired long-term recovery.11 Also, mTBI increased the risk of developing PTSD at 6-months posttrauma in a civilian population, such that 26% of patients with mTBI met criteria for probable PTSD, whereas only 15% of patients who did not have a mTBI met criteria for probable PTSD.12 The striking overlap between postconcussive and PTSD symptoms13,14 presents a significant challenge to diagnostic accuracy, and standard neuropsychological evaluations and neuroimaging techniques have so far shown limited efficacy in discriminating between the two. The diagnosis of a TBI is established with a clinical interview to determine factors such as the duration of loss of consciousness and post-traumatic amnesia; a careful neurological examination; and neuroimaging, such as computed tomography or magnetic resonance imaging, which is typically negative in...
mTBI. The diagnosis of PTSD is made through the assessment of the presence and severity criteria of symptoms in each of these categories: (1) re-experiencing; (2) hyper-arousal and reactivity; (3) avoidance; and (4) negative mood and cognition, as defined in the Diagnostic and Statistics Manual, 5th edition (DSM5). Symptom severity is often obtained through the administration of self-report scales, which are subject to recall bias. Even the gold standard Clinician-Administered PTSD Scale (CAPS), which mitigates some of the limitations of self-reported scales, is subjective and can be influenced by recall bias. TBI is characteristically related to a traumatic event, and approximately 10 to 15% of those who are diagnosed with a TBI will have persistent postconcussive symptoms, but it can be incredibly difficult to distinguish whether the symptoms are more likely attributable to the physical or the psychological impact of the trauma. With many clinical studies using incomplete medical records and subjective self-report questionnaires to categorize patients, it is perhaps not surprising that TBI clinical drug trials to date have continued to show no benefit despite promising preclinical and even early phase results. This has prompted efforts to develop and refine clinical assessments that ensure SMs receive accurate diagnoses to facilitate appropriate and effective treatment. In particular, novel technology with integrative capabilities may afford the specificity and sensitivity needed to adequately differentiate between the impact of TBI and psychological health comorbidities.

The Computer-Assisted Rehabilitation Environment (CAREN; Motekforce Link, the Netherlands; Figure 1) is an immersive, interactive technology that is suited to the assessment of TBI and psychological health-related symptomatology. It is a dynamic system that incorporates multi-planar motion, on a platform with an embedded treadmill, with interactive virtual environments (VEs) projected onto a curved panoramic screen. Overhead infrared cameras, which capture real-time movement from reflective markers placed on participants, enable this interaction, and an overhead surround sound system provides a realistic auditory experience. For safety, participants are outfitted with a full-body harness and secured to the safety stand of the platform. The immersive and adaptive CAREN-based VEs allow clinicians to integrate advanced technology in the assessment and rehabilitation of injured SMs including assessments of gait, motion sensitivity, and balance while also incorporating cognitive and visual scanning tasks. To date, the majority of the CAREN literature on U.S. SMs have focused on its utility in rehabilitation; in particular, for gait and balance deficits in SMs with lower extremity amputations and vestibular dysfunction in SMs with TBI. In addition, our recent work reported that SMs experience high levels of presence (i.e., a feeling of being immersed in a VE) during CAREN tasks regardless of TBI or PTSD severity. Moreover, interaction with the CAREN VEs may better reflect real-world movement when compared to standard treadmill or standing balance assessments.

At the National Intrepid Center of Excellence (NICoE) (Walter Reed National Military Medical Center [WRNMMC], Bethesda, Maryland), SMs are referred to the CAREN largely because of the multimodal capabilities of the system, which aids the physical therapists in addressing certain deficits (e.g., visual over-reliance) that may be more challenging.

FIGURE 1. The NICoE CAREN.
to stimulate in the conventional physical therapy gym (Kruger et al under review). Although comorbid TBI and psychological health symptoms have not previously been assessed with CAREN VE tasks, VEs have had demonstrable efficacy in both the assessment and treatment of SMs with PTSD, as well as for those with a history of TBI. Therefore, by facilitating multisensory and motor integration, the CAREN represents a novel assessment tool with the potential to distinguish between PTSD and TBI.

Thus, the purpose of this study was to determine whether physical performance in CAREN VE tasks could effectively differentiate SMs with and without PTSD in a cohort of patients with a history of TBI. We hypothesized that in SMs with a history of TBI, those with comorbid PTSD would perform differently in CAREN VE tasks than those without PTSD.

**METHODS**

**Participants**

**Demographics**

All SMs admitted into the NICoE 4-week program were referred for evaluation because of a history of TBI with associated psychological health complaints, most commonly anxiety (44.9%), depression (53.7%), and PTSD (70.1%). During their stay at the NICoE, SMs were referred to the CAREN by their physical therapists after the interdisciplinary team, which includes primary care, neurology, radiology, behavioral health, rehabilitation, sleep, and integrative health professionals, completed their evaluations. A total of 214 SMs (6 female) with TBI who were referred to the CAREN from 2010 to 2015 consented to the collection of their clinical data in a database for research purposes. All branches of service were represented: Air Force (8%), Army (37%), Marines (30%), and Navy (24%). Ages ranged from 20 to 63 years (mean: 35 years), 82% were Caucasian, whereas 19.2% had moderate and 6.1% had severe TBI; severity was unknown for three SMs (1.4%).

**Data Integrity**

Data were obtained by independent, retrospective review of self-report questionnaires at NICoE intake, clinical notes from evaluations by neuropsychologists, psychiatrists, and neuropsychologists, and objective outcomes from the CAREN session notes. The NICoE Informatics team compiled the diagnosis and self-report data from consented SMs’ medical records into a deidentified database, which was a source of data for our study. Two blinded, independent reviewers reviewed the diagnosis and CAREN dataset for data integrity and quality assurance after data were manually extracted and compiled from the session notes. In all cases, if diagnoses differed between data sources, the diagnoses from the NICoE clinical interviews were used because they were the most recent and thorough notes available.

**Patient Protection and Informed Consent**

All appropriate measures were taken to protect patient privacy and information in accordance with procedures approved by the WRNMMC institutional review board. SMs provided written informed consent to participate in the NICoE database study before any data were collected. Each SM was assigned a unique study identification number and data were encrypted using a one-way hash with time-stamped audit trails (to track researcher access), and stored in a password-protected database on a secure server.

**TBI and Psychological Health Diagnosis**

All SMs in this study had a history of TBI. TBI severity (mild, moderate, and severe) was derived from diagnoses in SMs’ medical history and clinical interviews at NICoE; for all SMs, the most severe injury on record was reported. All SMs in this study had one or more psychological health comorbidities, most frequently PTSD, depression, and anxiety as noted earlier. The PTSD diagnosis was derived by reconciling information in the SM’s medical history with NICoE clinical interviews, and PTSD Checklist—Military Version total scores (when available; 70% of SMs). Similarly, depression and anxiety diagnoses were derived by reconciling diagnoses in SMs’ medical history, with NICoE clinical interviews, and relevant self-report questionnaires. In all cases, if diagnoses differed between sources, the diagnoses from the NICoE clinical interviews were used.

**Self-Report Questionnaires**

**Beck Anxiety Inventory**

A validated 21-item self-report questionnaire for anxiety in which the severity of generalized anxiety symptoms over the past week is rated on a 4-point Likert scale; scores range from 0 to 63. The questionnaire includes items about both emotional and somatic symptoms of anxiety and a total score ≥20 indicates probable generalized anxiety disorder (GAD).

**Beck Depression Inventory—II**

A validated self-reported 21-item questionnaire for depression in which the severity of symptoms over the past week is rated on a 4-point Likert scale. Scores range from 0 to 63; total scores ≥20 for mTBI and >36 for moderate to severe TBI are considered probable major depressive disorder.

**Generalized Anxiety Disorder 7-item**

A validated 7-item self-report screening instrument for generalized anxiety in which the frequency of anxiety symptoms experienced in the preceding 2 weeks are rated on a 4-point
Likert scale. Scores range from 0 to 21; scores ≥10 are indicative of probable GAD.39

**Patient Health Questionnaire-9**
A validated 9-item self-report screening instrument for major depression in which symptoms experienced in the preceding 2 weeks are rated on a 4-point Likert scale.40 Scores range from 0 to 27; a cutoff score of ≥10 had high specificity (88%) and sensitivity (88%) for probable major depression.

**PTSD Checklist—Military Version**
A validated self-reported 17-item screen for PTSD. The degrees to which participants are bothered by symptoms are rated on a 5-point Likert scale41; scores range from 17 to 85. A total score ≥50 indicates probable PTSD.42

**CAREN Preliminary VEs**
Three preliminary VE tasks were used to acclimatize patients to navigation in the CAREN (Figure 2).

1. **Balance Balls VE**: Acclimation to weight shifting on a static platform. Two reflective markers were placed immediately above the SM’s pelvis, on the back of the safety harness, and synced to a yellow ball in the center of the screen. Red balls appeared sequentially at random positions on the screen, requiring SMs to weight shift (feet planted) to hit each target, moving the yellow ball back to center before the next red ball appeared. The main outcome measure was time in seconds taken to complete the task.

2. **Balance Cubes VE**: Acclimation to step shifting with and without platform motion. Patients were synced to a white diamond with the same marker placement as described earlier for the Balance Balls VE. Different colored cubes appeared at random positions on the screen, requiring SMs to step shift (feet planted) to hit each target. SMs were required to move from one target cube to the next, as the cubes appeared sequentially. In the first run of the task, the platform was flat and static (Balance Cubes—Static); at the physical therapist’s discretion, a second run of the task included platform motion with subtle fore/aft and medial/lateral rotations (Balance Cubes—PM). The main outcome measure was time in seconds taken to complete each run of the task.

3. **Continuous Road VE**: Acclimation to flat ambulation on the CAREN treadmill. The platform remained static and treadmill speed was synced to the optical flow of the VE path to maintain congruence. The system operator gradually adjusted the treadmill speed to the SM’s comfortable/preferred walking pace. The main outcome measure was the final self-selected speed in m/s at the end of the task.

**Statistical Analyses**
Statistical analyses were conducted in SPSS 19 (IBM Corp., Armonk, New York). Pearson’s correlation determined significant correlations across the VE tasks, TBI severity, PTSD, depression, anxiety, history of substance abuse, and demographics (age, gender, ethnicity, handedness, branch of service, and marital status). Then, stepwise multiple regression analysis was used to determine which psychological health comorbidities were most associated with VE task performance. Significant correlations with demographic variables and TBI severity were controlled for in the regression analysis. Finally, a binary logistic regression examined whether VE task performance could predict diagnostic category, and a receiver operating characteristic (ROC) curve was used to determine VE task performance sensitivity and specificity.

**RESULTS**
First, to determine whether there were significant correlations across VE tasks, Pearson’s correlations were completed and identified that performance was significantly correlated across the preliminary VEs. We found that greater time spent on the Balance Balls task correlated with greater time spent on the Balance Cubes—Static ($r = 0.57; p < 0.001$) and Balance Cubes—PM ($r = 0.62; p < 0.001$) tasks, and time spent was also significantly correlated between both Balance Cubes tasks ($r = 0.81; p < 0.001$; Table I). In contrast, slower self-selected speed on the Continuous Road task was negatively correlated with time spent on the Balance Balls task ($r = -0.24; p = 0.002$), and the

---

**FIGURE 2.** Screenshots of the 3 preliminary CAREN VEs: Balance Balls, Balance Cubes, and Continuous Road.
Balance Cubes—Static \((r = -0.25; p = 0.001)\) and Balance Cubes—PM \((r = -0.38; p < 0.001)\) tasks (Table I). Pearson’s correlation was also used to determine whether there were significant correlations between VE task performance, TBI severity, PTSD, depression, anxiety, history of substance abuse, and demographic variables; only PTSD diagnosis was significantly correlated with VE task performance. PTSD was also significantly correlated with depression \((r = 0.30; p < 0.001)\) and anxiety \((r = 0.14; p = 0.044)\) indicating multicollinearity between these comorbid diagnoses in our cohort.

We then determined which psychological health comorbidities were most associated with physical performance on the CAREN VE tasks. Using stepwise multiple regression with PTSD, depression, and anxiety as independent variables in the regression model, we found that only PTSD was significantly associated with performance on each of the four tasks. We found that SMs with comorbid PTSD spent significantly more time on the Balance Balls \((\beta = 14.21; p = 0.001)\), Balance Cubes—Static \((\beta = 10.36; p = 0.002)\), and Balance Cubes—PM \((\beta = 14.47; p < 0.001)\) tasks (Figure 3) than SMs without PTSD. On the Continuous Road task \((\beta = -0.231; p = 0.018)\), SMs with comorbid PTSD selected significantly slower walking speeds than SMs who were not diagnosed with PTSD (Figure 3). Neither depression nor anxiety was significantly associated with VE task performance.

### TABLE I. Correlation Between CAREN VE Tasks

<table>
<thead>
<tr>
<th>CAREN Task</th>
<th>Balance Balls (Seconds)</th>
<th>Balance Cubes—Static (Seconds)</th>
<th>Balance Cubes—PM (Seconds)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Balance Cubes—Static (Seconds)</td>
<td>0.568***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Balance Cubes—PM (Seconds)</td>
<td>0.624***</td>
<td>0.809***</td>
<td>-0.239***</td>
</tr>
<tr>
<td>Continuous Road (m/s)</td>
<td>-0.246***</td>
<td>-0.383***</td>
<td></td>
</tr>
</tbody>
</table>

\*\*\*p ≤ 0.01

![FIGURE 3](image-url). Performance on each of the CAREN tasks was significantly poorer in SMs with comorbid PTSD compared to those without PTSD. For Balance Balls, Balance Cubes—Static, and Balance Cubes—PM VE tasks, mean time (seconds) to complete each task is illustrated ± standard error of the mean (SEM) error bars. For the Continuous Road VE task, mean self-selected speed (m/s) is illustrated ± SEM. The means ± SD values are provided in Table II. *\(p < 0.05\); **\(p < 0.01\); ***\(p ≤ 0.001\).
performance, thus, they were excluded from the regression model. Mean (and standard deviation) task performance scores for each VE task are provided in Table II.

Finally, to determine whether CAREN task performance could predict diagnostic category (i.e., No PTSD vs. Comorbid PTSD), a binary logistic regression was used; time spent on the Balance Cubes—Static VE task significantly predicted PTSD status ($\beta = 0.031; p = 0.003; \text{OR} = 1.03$ [95\% CI = 1.01, 1.05]), such that a 1-second increase in time spent on the task was associated with a 3\% increased odds of being in the comorbid PTSD category. In the ROC analysis, area under the curve (AUC) for the Balance Cubes—Static task was modest (AUC = 0.664), but statistically significant ($p < 0.001$; Figure 4). Similarly, the other VE

<table>
<thead>
<tr>
<th>CAREN VE Task</th>
<th>PTSD Diagnosis</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Balance Balls (Seconds)</td>
<td>No PTSD Comorbid PTSD</td>
<td>60.8</td>
<td>17.1</td>
<td>63</td>
</tr>
<tr>
<td>Balance Cubes—Static (Seconds)</td>
<td>No PTSD Comorbid PTSD</td>
<td>75.4</td>
<td>34.4</td>
<td>147</td>
</tr>
<tr>
<td>Balance Cubes—PM (Seconds)</td>
<td>No PTSD Comorbid PTSD</td>
<td>58.5</td>
<td>19.5</td>
<td>62</td>
</tr>
<tr>
<td>Continuous Road (m/s)</td>
<td>No PTSD Comorbid PTSD</td>
<td>68.8</td>
<td>22.9</td>
<td>147</td>
</tr>
</tbody>
</table>

FIGURE 4. ROC curves for the CAREN VE tasks showed modest but statistically significant AUC. Balance Balls AUC = 0.618 ($p = 0.007$), Balance Cubes—Static AUC = 0.664 ($p < 0.001$), Continuous Road AUC = 0.610 ($p = 0.021$), and Balance Cubes—PM AUC = 0.688 ($p < 0.001$).
DISCUSSION

Our retrospective study of performance on CAREN VE tasks in a cohort of SMs with a history of TBI identified significant differences between those with and without comorbid PTSD. We found that demographic factors, TBI severity, substance abuse, and other comorbid psychiatric diagnoses (i.e., depression and anxiety) exerted no significant influence on performance. Our findings suggest that PTSD symptoms may inhibit performance on physical tasks such as step and weight shifting and self-selected walking speeds. We expand on previous literature showing that such physiological changes that lead to balance and vestibular deficits among those with a history of TBI, with increased complications from the comorbid history of behavioral and psychological health disorders. There is also intriguing evidence that PTSD may exacerbate the somatic manifestations of postconcussive symptoms, suggesting that PTSD treatment may be effective in reducing postconcussive symptoms. However, the etiological mechanisms that enhance PTSD vulnerability after TBI have yet to be elucidated, and various therapeutic strategies aimed at ameliorating PTSD symptoms may inhibit performance on physical tasks select slower walking speeds and take longer to complete tasks in novel CAREN VEs compared to SMs without PTSD. Hypervigilance is a defining feature of PTSD; in some settings, it can be protective by helping SMs to avoid danger, but it can also be maladaptive. Thus the slower performance observed on CAREN tasks may not necessarily represent greater functional impairment, as it could alternatively be interpreted as the exhibition of precaution in the face of the unexpected, which is a potentially reasonable behavioral adaptation in this population. This warrants future study, perhaps assessing performance in CAREN-based VEs that connote varying levels of threat. Our findings also suggest the need to incorporate measurement of the severity of PTSD symptoms in those with TBI, as greater time may need to be allotted for those with greater symptom severity to complete rehabilitative tasks, for example.

Moreover, our results suggest that CAREN tasks can help to distinguish those with comorbid TBI-PTSD from TBI patients without PTSD. We found that performance on a CAREN weight-shifting task significantly predicted diagnostic category, suggesting a potential utility in the diagnostic process for SMs who have experienced trauma. Indeed, objective functional performance in the CAREN would be a significant advancement in TBI diagnosis and care because there are currently few established objective diagnostics for TBI, particularly mTBI with co-occurring psychiatric conditions. There is also often a demonstrable benefit to providing objective measures to patients, especially in relation to “invisible wounds” for which there is still a measure of misunderstanding and even stigma.

Finally, this is the first report directly using CAREN task performance in the assessment of comorbid TBI and behavioral health conditions (particularly PTSD), which serves as an indication of the predictive potential of this innovative technology. Furthermore, the multimodal sensorimotor integration, and incorporation of locomotion with virtual reality, makes the CAREN suited for the assessment of complex cases, ranging from physical trauma, as it has been traditionally used, to TBI and psychological trauma, as we report here. The CAREN is efficient, objective, and immersive, places a low burden on patients, and may afford a measure of “real-world” functional status.

LIMITATIONS

Our work has several potential limitations. First, we were not able to study the performance of SMs without TBI, or with PTSD alone. Second, all data were collected at a single, tertiary care site, so there may be elements of referral or selection bias, and the results may not be generalizable to all SMs, much less to civilian populations. Third, most (97.2%) of the SMs in our cohort were reported to be taking at least one, often two or more medications at the time of evaluation. Since SMs were primarily seen for clinical care and not research, the degree of polypharmacy was inconsistent between patients and was thus difficult to control for statistically. Furthermore, the side effects of some of these medications could impact balance and reaction times. However, although this could be addressed in prospective studies with inclusion/exclusion criteria, in clinical settings, patients with comorbid TBI and psychological health comorbidities are likely to have been prescribed these medications, making this cohort reasonably representative. Fourth, the use of self-report measures introduces the possibility of recall bias. We attempted to mitigate the impact by using well-validated questionnaires and the diagnostic consensus of the multidisciplinary team, incorporating clinical assessments by neurologists, neuropsychologists, and psychiatrists into the diagnosis of TBI and psychological health comorbidities. Finally, this is a retrospective analysis, thus some data are missing; however, considering the constraints of the clinical data, we were able to establish a data set with high internal consistency through careful review, reconciliation of multiple data sources, and quality assurance checks by two independent reviewers. Well-controlled covariates also ensured that we only reported the strongest correlations and associations. The PTSD prevalence in this patient sample is concordant with rates reported in the literature, which adds some validity to the data, but future prospective studies are important to validate the use of the CAREN in this manner.

Virtual Environment Assessment for Traumatic Brain Injury and Comorbid PTSD

Virtual Environment Assessment for Traumatic Brain Injury and Comorbid PTSD
CONCLUSIONS

In a population of SMs with a history of TBI, performance on the CAREN VE tasks was statistically different between those with and without comorbid PTSD. This is the first report of performance on CAREN VE tasks in comorbid TBI and PTSD, delineating the potential for this advanced technology to serve as a novel assessment tool in the differentiation of TBI patients with and without comorbid PTSD. Utilizing advanced technology platforms such as the CAREN, which enable the objective evaluation of physical performance during engaging VE tasks, could significantly enhance our understanding of the nuanced differences in those with dual diagnoses. Moreover, in a population where symptom overlap is often a major barrier to appropriate diagnosis and intervention, these simple tasks provide a valuable assessment of functional status as it may relate to performing mission essential tasks that require SMs to move and multi-task. Our report also provides a foundation for future examinations of performance on more complex CAREN tasks and in larger, prospective cohort studies to validate and further define the utility of CAREN tasks in the evaluation of patients with complex comorbidities.

ACKNOWLEDGMENTS

We thank Paula Bellini and Mary Coughlin for their quality assurance reviews of the CAREN data and clinical diagnoses. We also thank the clinicians in the NICoE Rehabilitation Department (other than MMP and PNK) who referred patients for assessment in the CAREN. We thank the NICoE Informatics Department for compiling the diagnostic clinical notes and questionnaires. This study was supported by CDMRP–MRMC Contract no. W81XWH-14-C-0139 - NICoE Network Research Program (NNRP) Infrastructure.

REFERENCES

Virtual Environment Assessment for Traumatic Brain Injury and Comorbid PTSD


Long-Term Outcomes and Needs of Military Service Members After Noncombat-Related Traumatic Brain Injury

Kelly J. Miller, MPH*; Jan E. Kennedy, PhD†; Karen A. Schwab, PhD‡

ABSTRACT  Objectives: Assess the prevalence of self-identified unmet service needs in a military sample an average of 5 years following noncombat traumatic brain injury (TBI). Examine relationships between unmet needs and background, injury-related and outcome variables. Methods: The study sample consisted of 89 veterans and service members who sustained non-combat TBI between 1999 and 2003, selected from enrollees in the Defense and Veterans Brain Injury Center TBI registry. Semistructured telephone interview was used to collect information about participants’ self-reported unmet service needs, symptoms, and functional status. Results: Most participants (65%) reported having at least one unmet service need. The most prevalent needs were “getting information about available post-TBI services” (47%) and “improving memory and attention” (45%). Unmet needs were associated with cognitive difficulties, physical and emotional symptoms, mental health diagnosis/treatment, and poorer functional status. Conclusions: Needs for services following TBI are associated with poor symptomatic and functional outcomes and may persist for years after injury in military service members and veterans. The study suggests service members’ needs post TBI for improved cognition, support for emotional issues, and resources for vocational skills. Information about available services should be made accessible to those recovering from TBI to reduce the incidence of long-term unmet needs.

INTRODUCTION
Traumatic brain injury (TBI), defined as “an alteration in brain function, or other evidence of brain pathology, caused by an external force,”1 is known to impact physical, social, and emotional functioning for a period of weeks, months, or years following injury.2–9 Although many of the 5.3 million individuals with TBI-related disability10 adapt to postinjury life without assistance,11–14 others may require longer-term, specialized services in areas such as activities of daily living (ADL), education, vocation, and community integration to cope with functional limitations.15,16 “Unmet” needs for services generally refer to expressed needs for assistance that are not satisfied by current service provision.17,18 Population-based studies of service needs following TBI have found unmet service needs reported by 35 to 70% of individuals with TBI of varying severity (mild through severe).18–20 Research to date points to consistent patterns of unmet cognitive, emotional, and vocational needs that persist well beyond the acute TBI recovery period.17–19,21 even when physical impairments are minimal20 or personal independence has increased over the postinjury period.21

Existing studies of long-term outcome and needs after TBI have predominantly focused on hospitalized, civilian populations. Few studies have examined the outcomes and unmet needs for services in military populations. This may be due in part to the difficulty of locating individuals for follow-up: military personnel relocate every 2 to 3 years on average.22–24 Nonetheless, there is an acknowledged need for data on TBI-related disabilities, impairments, and persistent symptoms in military personnel.25,26

Currently, service members receive routine postdeployment screenings for TBI27 that are followed by clinical evaluation to determine the need for follow-up and care. Service members returning from deployments are required to complete the Department of Defense (DoD) Form 2796, Post Deployment Health Assessment,28 which since 2007 has incorporated questions regarding potential exposures to TBI. Veterans Administration/DoD clinical practice guidelines specify steps for assessment and treatment of military personnel and veterans who present with a potential TBI and for those who have symptoms lasting more than 7 days.29

Most TBI sustained by military service members is mild TBI (mTBI),30 which is often found associated with complete recovery within days or weeks.31–33 This typically rapid recovery trajectory is similar to that seen in sports concussion in the civilian world.34 However, investigations have documented persistent postconcussive-type symptoms and poor functional outcomes in some service members with mTBI, even years after injury.35,36 In service members, mTBI may be comorbid with post-traumatic stress and symptoms that appear to have a relationship to the level of combat stress experienced.37 Comorbidities, preinjury, and noninjury-related factors may all impact TBI recovery in military personnel.38 A systematic review of mTBI prognosis in adults found evidence that complete recovery may take 6 months to 1 year.39

The time course of recovery from moderate and severe TBI is variable but there is some evidence of more
accelerated recovery within the first 5 to 6 months,\textsuperscript{40} and of moderation of the recovery trajectory by age.\textsuperscript{41} Those recovering from moderate to severe TBI may have significant neurological impairments, including motor deficits and movement and communication disorders, in addition to impairments in attention, memory, executive control, mood disorders, and seizure disorders that can result in significant reductions of functional independence and quality of life,\textsuperscript{42} including the likelihood of loss of productivity.\textsuperscript{13} The cognitive and behavioral changes that may result from TBI across the spectrum of severity may warrant the support of specially trained persons to provide assistance for routine activities and tasks.\textsuperscript{18} Unmet needs may indicate barriers to maximal recovery from TBI for those experiencing postinjury sequelae.

The present investigation of a cohort of military service members and veterans who sustained non-combat-related TBI an average of 5 years earlier was undertaken in an effort to address gaps in knowledge regarding long-term outcomes and postinjury needs after TBI in military and former military personnel. Specifically, the primary goals of this study were (1) to identify the self-identified unmet service needs in this sample, (2) to characterize self-reported symptoms and functional status, and (3) to investigate the association between the number of endorsed unmet service needs and symptoms, functional status, and TBI-related and demographic variables. A secondary objective of this study was to investigate the feasibility of using a registry database of individuals in the military who received inpatient and outpatient care for TBI to retrospectively locate military service members and veterans.

METHODS

Sample Selection and Recruitment

The study was approved by the institutional review boards of the 3 centers of Defense and Veterans Brain Injury Center (DVBIC) participating in the study. The study sample was drawn from a DVBIC TBI registry database used for evaluation, surveillance, and patient recruitment. TBI diagnosis and information regarding the circumstances of injury (cause of injury, whether or not sustained in battle) in the registry were derived from the treating clinician’s assessment in the patient’s medical record. Determination of injury severity was consistent with the Veterans Administration/DoD Clinical Practice Guideline for the Management of Concussion/mTBI.\textsuperscript{29} TBI with duration of loss of consciousness (LOC) or alteration of consciousness of no more than 30 minutes and posttraumatic amnesia (PTA) no more than 24 hours was identified as mTBI. Those associated with more than 30 minutes of LOC or more than 24 hours of PTA were defined as moderate to severe TBI.

Inclusion criteria for this study included (1) TBI sustained between March 1, 1999, and February 28, 2003, for which the patient was treated in an emergency department, or overnight at one of the participating DVBIC centers; (2) clinical diagnosis of a nonpenetrating TBI that was not sustained in combat; (3) 12 months or more postinjury; (4) military service member (including reserve status) or veteran at the time of injury; (4) age 18 to 75 years; and (5) informed consent and authorization signed by the patient. The sample was not restricted by prior neurological or psychiatric conditions, previous TBI, or substance abuse. Potential participants were excluded if they were unwilling to provide informed consent and authorization or if they demonstrated confusion during a telephone conversation with a study coordinator and subsequently failed a brief assessment of comprehension (Table I) of the study purpose and informed consent process.

Investigators used a variety of resources, including the Defense Manpower Data Center, and the public databases USSearch (https://www.ussearch.com) and Intelius (www.intelius.com), in an exhaustive attempt to update registry contact information for potential participants. Although 500 individuals met TBI and demographic criteria for inclusion, valid contact information could not be located for 258 individuals. The remaining 242 potential participants were successfully reached by mail (Fig. 1).

Ninety-three individuals refused to participate by returning postcards provided for this purpose, and 17 more refused participation verbally when contacted by telephone (six of whom denied having TBI). Study investigators contacted the remaining 138 individuals who expressed willingness to participate in the study. Telephone screening with these individuals was conducted to provide information about the study’s purpose and informed consent process, to verify study eligibility criteria, and to conduct a five-question standardized assessment of comprehension to ensure cognitive

<table>
<thead>
<tr>
<th>TABLE I.</th>
<th>Assessment of Comprehension Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Interviewer to potential participant:</strong> “I would like to ask you some questions about this study before we continue.”</td>
<td></td>
</tr>
<tr>
<td>In this study, we want to evaluate problems in patients with head injuries. (T)</td>
<td></td>
</tr>
<tr>
<td>You will still receive your usual medical care, even if you decide not to do this interview. (T)</td>
<td></td>
</tr>
<tr>
<td>Once you begin this interview, you may stop it at any time. (T)</td>
<td></td>
</tr>
<tr>
<td>By taking part in this study, you will help us learn more about spinal cord injuries. (F)</td>
<td></td>
</tr>
<tr>
<td>We will keep your answers private, and only people involved with the study will be able to view them. (T)</td>
<td></td>
</tr>
</tbody>
</table>

The study investigator read the questions to the potential participant during the telephone screening, and recorded their responses. Failure to answer 3 or more questions correctly resulted in the assessment that the individual had insufficient understanding of the study purpose and informed consent process to participate.
competency (Table I). Individuals who answered three or more questions incorrectly on the comprehension screening were deemed to have insufficient understanding of the study to complete the interview.

Telephonic eligibility screening revealed that 10 of the remaining 132 individuals were deceased, and 11 individuals reported that their TBIs had been sustained outside the March 1999–February 2003 date range (TBI date was recorded inaccurately in the registry). Of the remaining 111 individuals, an additional 13 failed the assessment of comprehension. Participants whose eligibility was confirmed and who demonstrated satisfactory comprehension of the study purpose and procedures were sent informed consent and Health Insurance Portability and Accountability Act documents. All consenting participants (n = 89) completed telephonic interviews. Participants were not compensated for their involvement in the study.

Data Collection and Analysis
Primary outcome variables in this investigation were obtained from self-report during telephonic interview and included unmet service needs, symptoms, and functional status. Telephone interviews were conducted between September 2004 and October 2007 by trained study coordinators using an interview script, and took from 1 to 2 hours to complete. The time interval between injury and interview ranged from 2.5 to 8.2 (mean = 5.0) years. The interview was based on an instrument developed for the Centers for Disease Control and Prevention TBI prevalence study conducted in South Carolina.44 The questionnaire was modified by DVIC and Centers for Disease Control and Prevention investigators, with the permission of the South Carolina investigators, to add items related to the participants’ military status and military insurance coverage (DVIC: A Prevalence Study of Chronic Problems and Sequelae after TBI in the Military and Veterans. Unpublished Questionnaire, 2003). The questions assessed self-reported injury and health status (TBI history, pre-existing conditions, psychiatric problems, general health, disability, ADL), postinjury symptoms (physical, psychological and behavioral, and cognitive), psychosocial function (life satisfaction, living situation, alcohol use, and social integration and support), socioeconomic and occupational status (education, income, and employment), and service needs across several domains.

Unmet service needs were those service needs that the participants stated they currently needed, in any of the following areas: controlling alcohol, finding paid employment, improving job skills, finding places and/or opportunities to socialize, obtaining help from a personal assistant, increasing personal independence, getting information about available post-TBI services, getting and/or managing services, improving mood and/or managing stress, traveling in the community, improving memory and/or attention, or “other” need (open ended) that they felt was important. If a participant reported a need for services in a particular area at the time of interview, it was counted as an unmet service need, whether or not services or assistance in that area were received in the past or at the time of the interview. Similar to the approach used previously by Pickelsimer et al.,18 participants were grouped into 3 categories for analysis according to the number of unmet needs they reported: 0, 1 or 2, and 3 or more.

Self-reported symptoms were categorized into three domains: physical, affective, and cognitive. A participant was counted as endorsing physical symptoms if he or she reported that one or more of the following symptoms caused either “a little” or “a lot” of limitation in daily life within the 4 weeks preceding the interview: vision problems, headache/migraine, hearing problems, balance problems, dizziness, gait problems, weakness in limbs, seizures, and stroke. A participant was counted as endorsing affective symptoms if he or she reported that one or more of the following symptoms caused either “a little” or “a lot” of daily limitation in the previous 4 weeks: depressed/sad, irritability, sleep problems, fatigue, tense, keyed up, temper problems, and problems getting along with others. For the cognitive domain, a participant was counted as
endorsing cognitive symptoms if he or she reported any of the following problems: slowed reaction, forgetfulness, poor concentration, decision/planning difficulty, and attention difficulty.

Counting physical and affective symptoms as endorsed only if the participant reported that they caused some level of daily limitation established a threshold of functional limitation, rather than merely affirmative endorsement of physical or affective symptoms. Because of the way the questionnaire was constructed, cognitive symptoms were not rated by degree of daily functional limitation in the same way that physical and affective symptoms were. For the purposes of this study, any of the cognitive symptoms listed above that the participant reported was counted as a “positive” symptom.

Functional status outcomes included requiring supervision or assistance with at least one ADL such as bathing, dressing, eating, getting in/out of chairs/beds, walking or using the toilet, or at least one instrumental ADL (IADL) such as preparing meals, shopping, or managing money. Other functional status outcomes included self-perceived disability, self-perceived general health, life satisfaction, and mental health diagnosis or treatment within the past year. The symptoms and functional status outcomes assessed in the interview were those endorsed by participants as being present at, or proximate to, the time of the interview.

Data were analyzed using SPSS Statistics Version 23 for Windows (IBM, Armonk, New York). χ² tests compared dichotomous demographic, injury-related, symptomatic, and functional outcome variables by unmet needs groups (0, 1–2, and 3+ needs) in 2 × 3 tables. Variables found to be significantly associated with unmet needs in the 2 × 3 analyses were further examined in pairwise analyses, comparing the 3+ needs group with both the 0 and the 1 to 2 needs group, to determine whether those with 3+ needs differed significantly from the other unmet needs groups. To offset the potential for spurious significant findings due to multiple comparisons, we established the 0.01 level (as opposed to 0.05) as our p-value cutoff for statistical significance.

RESULTS

Self-identified Current Unmet Needs

The total number of unmet needs endorsed by participants ranged from 0 to 9, out of 12 possible needs. Slightly over one-third of participants (34.8%; n = 31) reported having no unmet needs, one-fourth (23.6%; n = 21) reported one or two unmet needs, and almost half (41.6%; n = 37) reported that they had three or more unmet needs at the time of interview.

Table II lists the service needs in the order that participants were asked about them during the interview, along with the proportion of the sample (n = 89) endorsing each need. The need for information about available post-TBI services was the most prevalent need, followed by needs for improving memory and attention, mood and/or managing stress, job skills and finding paid employment, and assistance getting or managing services. Lower-frequency needs included finding social opportunities, “other” needs, increasing personal independence, obtaining help from a paid personal assistant, controlling alcohol use, and transportation within the community.

Service Provision and Service Barriers

Endorsement of an unmet need indicated the existence of a need at the time of interview; it did not necessarily signify that participants received no services or assistance for that need, or other needs. Fifty-eight per cent of the sample reported having received services to address one or more of their needs (excluding transportation, which nearly all participants reported that they had received) during the year preceding the interview. For participants lacking services, the most common service barrier was...

<table>
<thead>
<tr>
<th>TABLE II. Self-identified Current Unmet Needs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study Sample (N = 89)</td>
</tr>
<tr>
<td>n (%)</td>
</tr>
<tr>
<td>Controlling Alcohol</td>
</tr>
<tr>
<td>Improving Mood/Managing Stress</td>
</tr>
<tr>
<td>Improving Memory/Attention</td>
</tr>
<tr>
<td>Finding Paid Employment</td>
</tr>
<tr>
<td>Getting/Managing Services</td>
</tr>
<tr>
<td>Obtaining Help From a Paid Personal Assistant</td>
</tr>
<tr>
<td>Improving Job Skills</td>
</tr>
<tr>
<td>Getting Information About Available post-TBI Services</td>
</tr>
<tr>
<td>Finding Places/Opportunities to Socialize</td>
</tr>
<tr>
<td>Increasing Personal Independence</td>
</tr>
<tr>
<td>&quot;Traveling Within the Community&quot;</td>
</tr>
<tr>
<td>Other Needs a</td>
</tr>
</tbody>
</table>

"Other needs reported include financial planning, pain/symptom relief, patient information regarding TBI, marital counseling, specific head injury treatments, and support group."
Chi-square tests were used to test the significance of association between unmet needs groups and dichotomous categorical variables (2 × 3 tables). Differences between unmet needs groups were tested using one-way ANOVA. Differences between unmet needs groups on demographic or TBI-related factors. However, current income showed a trend toward significance, with 70% of those with 3+ needs earning less than $35,000 annually at the time of interview, followed by 43% of those endorsing 1 to 2 needs, and 36% of participants reporting no needs \( \chi^2 (2, N = 89) = 9.03, p = 0.011 \).

Demographic and TBI-related characteristics were analyzed by TBI severity groups (mild vs. moderate to severe), to investigate whether injury severity was associated with these variables. Differences between mild and moderate-to-severe participants were observed for 3 factors (not shown in tables): active duty status: 44% of mTBI group vs. 13% moderate to severe \( \chi^2 (1, N = 89) = 11.07, p = 0.001 \), hospitalized for TBI: 15% of mTBI group vs. 65% moderate to severe \( \chi^2 (1, N = 89) = 22.71, p < 0.001 \), and inpatient rehabilitation post-TBI: 9% of mTBI group vs. 67% of moderate to severe \( \chi^2 (1, N = 89) = 29.73, p < 0.001 \).

**TABLE III.** Demographic and TBI-Related Characteristics, by Unmet Needs Groups

<table>
<thead>
<tr>
<th>Statistical Test</th>
<th>One-way ANOVA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Sample</td>
<td></td>
</tr>
<tr>
<td>( (N = 89) )</td>
<td>( F )</td>
</tr>
<tr>
<td>Mean Age (SD)</td>
<td>40.18 (12.14)</td>
</tr>
<tr>
<td>Mean Years Since Injury (SD)</td>
<td>5.31 (1.32)</td>
</tr>
<tr>
<td>Number of Unmet Needs</td>
<td></td>
</tr>
<tr>
<td>0 ( (N = 31) )</td>
<td>40.68 (13.08)</td>
</tr>
<tr>
<td>1–2 ( (N = 21) )</td>
<td>37.57 (12.79)</td>
</tr>
<tr>
<td>3+ ( (N = 37) )</td>
<td>41.2 (11.04)</td>
</tr>
<tr>
<td>( \chi^2 )</td>
<td>2.86</td>
</tr>
<tr>
<td>( p )</td>
<td>0.647</td>
</tr>
<tr>
<td>df</td>
<td>2</td>
</tr>
<tr>
<td>Main Person Family/Household Relies on for Income</td>
<td>54 (62.1)</td>
</tr>
<tr>
<td>Visit(s) to Health Care Provider(s) in Past Year</td>
<td>80 (89.9)</td>
</tr>
<tr>
<td>Health Care Coverage (Pays All or Part)</td>
<td>83 (93.3)</td>
</tr>
<tr>
<td>TBI Severity Moderate/Severe</td>
<td>48 (53.9)</td>
</tr>
<tr>
<td>Hospitalized for TBI (vs. Treatment in ER/Clinic)</td>
<td>37 (41.6)</td>
</tr>
<tr>
<td>Vehicular Injury( ^a )</td>
<td>42 (47.2)</td>
</tr>
<tr>
<td>Inpatient Rehabilitation post-TBI</td>
<td>36 (40.4)</td>
</tr>
<tr>
<td>Family/Friend Made Work Sacrifice to Provide Post-TBI Care</td>
<td>23 (25.8)</td>
</tr>
<tr>
<td>Additional TBI (Before or Since Index TBI)</td>
<td>41 (46.1)</td>
</tr>
<tr>
<td>Male</td>
<td>74 (83.1)</td>
</tr>
<tr>
<td>AD</td>
<td>24 (27.0)</td>
</tr>
<tr>
<td>Married</td>
<td>43 (48.3)</td>
</tr>
<tr>
<td>One Year or More of College</td>
<td>61 (70.1)</td>
</tr>
<tr>
<td>Current Income &lt;$35,000/Year</td>
<td>46 (51.7)</td>
</tr>
<tr>
<td>Main Person Family/Household Relies on for Income</td>
<td>54 (62.1)</td>
</tr>
<tr>
<td>Visit(s) to Health Care Provider(s) in Past Year</td>
<td>80 (89.9)</td>
</tr>
<tr>
<td>Health Care Coverage (Pays All or Part)</td>
<td>83 (93.3)</td>
</tr>
<tr>
<td>TBI Severity Moderate/Severe</td>
<td>48 (53.9)</td>
</tr>
<tr>
<td>Hospitalized for TBI (vs. Treatment in ER/Clinic)</td>
<td>37 (41.6)</td>
</tr>
<tr>
<td>Vehicular Injury( ^a )</td>
<td>42 (47.2)</td>
</tr>
<tr>
<td>Inpatient Rehabilitation post-TBI</td>
<td>36 (40.4)</td>
</tr>
<tr>
<td>Family/Friend Made Work Sacrifice to Provide Post-TBI Care</td>
<td>23 (25.8)</td>
</tr>
<tr>
<td>Additional TBI (Before or Since Index TBI)</td>
<td>41 (46.1)</td>
</tr>
</tbody>
</table>

\( ^a \)Chi-square tests were used to test the significance of association between unmet needs groups and dichotomous categorical variables (2 × 3 tables). \( ^b \)Valid text statistic could not be computed due to insufficient cell counts. \( ^c \)Other injury agents include blunt objects, sports/recreation, training accidents, and unspecified causes.

AD, active duty; ANOVA, analysis of variance; ER, emergency room. Differences between unmet needs groups were tested using one-way ANOVA. Chi-square tests were used to test the significance of association between unmet needs groups and dichotomous categorical variables (2 × 3 tables). Valid text statistic could not be computed due to insufficient cell counts. Other injury agents include blunt objects, sports/recreation, training accidents, and unspecified causes.

MILITARY MEDICINE, Vol. 182, March/April Supplement 2017 141
Long-term Outcomes and Needs After Noncombat-Related TBI

**TABLE IV.** Self-reported Symptoms and Functional Status, by Unmet Needs Groups

<table>
<thead>
<tr>
<th>Number of Unmet Needs</th>
<th>Total Sample (n = 89)</th>
<th>0 (n = 31)</th>
<th>1-2 (n = 21)</th>
<th>3+ (n = 37)</th>
<th>Chi-square</th>
<th>df</th>
<th>( \chi^2 )</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1+ Physical Symptom</strong></td>
<td>51 (57.3)</td>
<td>10 (32.3)</td>
<td>9 (42.9)</td>
<td>32 (86.5)</td>
<td>2</td>
<td>22.62</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td><strong>1+ Affective Symptom</strong></td>
<td>57 (64.0)</td>
<td>12 (38.7)</td>
<td>10 (47.6)</td>
<td>35 (94.6)</td>
<td>2</td>
<td>26.10</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td><strong>1+ Cognitive Symptom</strong></td>
<td>63 (70.8)</td>
<td>14 (45.2)</td>
<td>12 (57.1)</td>
<td>37 (100.0)</td>
<td>2</td>
<td>27.00</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Mental Health Diagnosis or Treatment in Past Year</td>
<td>29 (32.6)</td>
<td>2 (6.5)</td>
<td>5 (23.8)</td>
<td>22 (59.5)</td>
<td>2</td>
<td>22.54</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Dissatisfied with Life, or Neither Satisfied nor Dissatisfied</td>
<td>34 (38.2)</td>
<td>3 (9.7)</td>
<td>5 (23.8)</td>
<td>26 (70.3)</td>
<td>2</td>
<td>28.64</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Self-assessed General Health: Fair to Poor</td>
<td>23 (25.8)</td>
<td>4 (12.9)</td>
<td>2 (9.5)</td>
<td>17 (45.9)</td>
<td>2</td>
<td>13.43</td>
<td>0.001</td>
<td></td>
</tr>
<tr>
<td>Considers Him/Herself to Have a Disability</td>
<td>45 (50.6)</td>
<td>8 (25.8)</td>
<td>10 (47.6)</td>
<td>27 (73.0)</td>
<td>2</td>
<td>15.12</td>
<td>0.001</td>
<td></td>
</tr>
<tr>
<td>Requires Assistance With At Least 1 ADL</td>
<td>7 (7.9)</td>
<td>1 (3.2)</td>
<td>1 (4.8)</td>
<td>5 (13.5)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Requires Supervision or Assistance With At Least 1 IADL</td>
<td>15 (16.9)</td>
<td>0 (0.0)</td>
<td>1 (4.8)</td>
<td>14 (37.8)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Chi-square tests were used to test the significance of association between unmet needs groups and dichotomous categorical variables. *One or more of the following symptoms causing some level of daily limitation: vision problems, headache (migraines and other), hearing problems, balance problems, dizziness, gait/walking problems, weakness in arms/legs, seizures, and stroke history. **One or more of the following symptoms causing some level of daily limitation: feeling depressed/sad, irritability, sleep difficulty, fatigue, tense/keyed up, temper problems, and problems getting along with others. ***One or more of the following symptoms causing some level of daily limitation: slowed reaction, forgetfulness, poor concentration, decision-making difficulty, and attention difficulty. **Bathing, dressing, eating, getting in/out of chairs/beds, walking, using toilet. ***Preparing meals, shopping, managing money. ****Valid test statistic could not be computed due to insufficient cell counts.

**Self-reported Symptoms**

Seventy-one per cent of the sample reported one or more cognitive problem, 64% reported one or more functionally limiting affective symptom, and 57% reported one or more functionally limiting physical symptom (Table IV). Symptoms differed significantly across unmet needs groups for all three symptomatic domains. In each domain, subjects with 3+ unmet needs differed from subjects with fewer unmet needs. In the cognitive domain, 100% of 3+ needs participants reported one or more cognitive symptom, compared to 57% of 1 to 2 needs participants [\( \chi^2 (1, N = 58) = 18.77, p < 0.001 \)] and 45% of participants endorsing no needs [\( \chi^2 (1, N = 68) = 27.05, p < 0.001 \)]. In the affective symptom domain, 95% of 3+ needs participants reported one or more functionally limiting symptom as compared to 48% of the 1 to 2 needs participants [\( \chi^2 (1, N = 58) = 17.00, p < 0.001 \)] and 39% of participants endorsing no needs [\( \chi^2 (1, N = 68) = 24.68, p < 0.001 \)]. In the physical symptom domain, 87% of 3+ needs participants reported 1 or more functionally limiting symptom as compared to 43% of the 1 to 2 needs participants [\( \chi^2 (1, N = 58) = 12.31, p < 0.001 \)] and 33% of participants endorsing no needs [\( \chi^2 (1, N = 68) = 21.00, p < 0.001 \)].

**Self-reported Functional Status**

Mental health diagnosis and/or treatment and lower life satisfaction (defined as being dissatisfied with life, or neither satisfied or dissatisfied) were reported by approximately one-third of participants. Each of these outcomes showed statistically significantly higher proportions among the 3+ needs group than in the other two needs groups (Table IV). Sixty per cent of participants with 3+ needs reported mental health diagnosis and/or treatment, compared to 24% of participants with 1 to 2 needs [\( \chi^2 (1, N = 58) = 6.84, p = 0.009 \)] and 7% of those reporting no needs [\( \chi^2 (1, N = 68) = 20.75, p < 0.001 \)]. Lower life satisfaction, reported by 70% of participants with 3+ needs, was a statistically significantly higher proportion than the 24% of those with 1 to 2 needs [\( \chi^2 (1, N = 58) = 11.62, p = 0.001 \)] and the 10% of those reporting no needs [\( \chi^2 (1, N = 68) = 25.32, p < 0.001 \)]. Slightly more than one-quarter of participants (26%) rated their health as fair to poor, and approximately half of study participants (51%) considered themselves to have a disability. Those with 3+ unmet needs were significantly more likely than those with 0 or 1 to 2 needs to report lower self-assessed general health outcomes: 46% of 3+ needs group with lower self-assessed general health vs. 10% of the 1 to 2 needs group [\( \chi^2 (1, N = 58) = 8.07, p = 0.005 \)] and 13% of the 0 needs group [\( \chi^2 (1, N = 68) = 8.63, p = 0.003 \)]. We note that 73% of the 3+ needs group participants assessed themselves as having a disability, a statistically significantly higher proportion than the 26% of participants reporting no needs [\( \chi^2 (1, N = 68) = 15.12, p < 0.001 \)]. However, the comparison of the 3+ needs group with the 1 to 2 needs group proportion reporting self-assessed disability (48%) did not reach statistical significance [\( \chi^2 (1, N = 58) = 3.73, p = 0.053 \)]. Few participants in the sample (17%) reported a need for support with IADLs, such as shopping, household tasks, and managing money, and only 7% reported needing assistance with ADLs involving basic personal care. Thus, there was insufficient power to enable detection of any statistically significant differences between unmet needs groups for these aspects of functional abilities.

As with demographics and TBI-related characteristics, analyses were conducted to ascertain differences in self-reported symptoms and functional status variables by TBI severity. Equivalent proportions of mTBI vs. moderate-to-severe
TBI groups endorsed one or more physical, affective, or cognitive symptoms; mental health diagnosis or treatment; lower life satisfaction; and self-assessed health or disability. As noted above, too few participants overall endorsed needs for assistance with ADLs or IADLs to permit statistical comparisons.

Nonparticipant Comparisons
A limited amount of demographic information was available from the registry for individuals who did not participate in the study (N = 411). This included injury year, gender, categorized age, whether or not the patient experienced any duration of LOC, and whether or not the patient experienced any duration of PTA. With the caveat that this information could not be verified via interview as it was for participants, we compared nonparticipant data to that of participants to provide some context for the representativeness of the study sample. Categorized age was the only factor for which a statistically significant difference between groups was observed, with significantly more participants (36%) than nonparticipants (22%) over 40 years of age [χ² (1, N = 500) = 7.78, p = 0.005 (data not shown in tables)]. There were no statistically significant differences observed in injury year distribution [χ² (4, N = 500) = 8.02, p = 0.091], gender [χ² (1, N = 500) = 2.07, p = 0.150], LOC [χ² (1, N = 500) = 0.123, p = 0.726], or PTA [χ² (1, N = 500) = 0.059, p = 0.808].

DISCUSSION
The present study examined unmet service needs that were reported by noncombatant military service members who sustained a traumatic brain injury an average of 5 years before interview. Approximately 65% of this sample endorsed at least one unmet need and 42% endorsed three or more unmet needs.

Present results are consistent with findings from some, but not all prior research. In a population-based study in Norway, 70% of participants reported at least one perceived need at 5 years following hospitalization for TBI. The 65% prevalence of at least one self-identified service need in this sample an average of 5 years postinjury in the present investigation is also similar to the 59% prevalence of at least one unmet need during the first year postinjury in a 2004 civilian study. However, in that study, a significantly lower percentage of participants had a persisting need after 1 year (40%).

It was suggested that the 1-year assessment prompted some participants to address their reported needs during the ensuing period. If this is the case, it indicates that follow-up of individuals with TBI is beneficial. DVBIC endorses the value of continued clinical follow-up for service members through the development of a network of nationwide Regional Support Specialists who maintain telephonic contact with service members diagnosed with TBI, to provide assessment and offer referrals for needed services after the service member is discharged from acute care at the military treatment facility.

The civilian registry study conducted in South Carolina, using the interview questionnaire adapted for use in the present study, found only 35% of community-dwelling participants with TBI reported at least one service need 1 year after discharge from acute care facilities. A higher proportion endorsed one or two needs (26%) than three or more needs (9%). However, the investigators determined that over half (51%) of the sample had at least 1 unmet need that the participants did not recognize or self-identify. This estimate emphasizes the fact that endorsed needs are influenced by the individual’s level of awareness of needs. Information from knowledgeable others (i.e., caregiver, spouse, family member) is of value in assessing the status and needs of individuals following TBI, and inclusion of these sources of information allows a more comprehensive and potentially more accurate estimate of actual needs. When feasible, supplemental information from individuals familiar with the patient’s daily functioning should be included in future investigations and clinical assessments.

The most prevalent needs in the present investigation were obtaining information about available post-TBI services, improving memory and attention, improving mood and/or managing stress, and improving job skills. These needs reflect the chronic problems with postconcussive cognitive and affective symptoms endorsed by these individuals, and are consistent with the areas of postinjury need found to be most prevalent in other studies. Ensuring that information about available services is provided to individuals recovering from TBI at the outset has the potential to reduce the incidence of long-term unmet needs.

Implications for the Military Health System
In this investigation, participants with a high number of unmet needs (3+) reported more postconcussive symptoms across cognitive, affective, and physical realms than those with fewer unmet needs. Those with high unmet needs also were more likely to have received mental health diagnoses and treatment and to report lower life satisfaction, worse general health, and presence of disability than those with fewer unmet needs. These results suggest that service members and veterans who report multiple unmet needs are in more general distress and are experiencing poorer global outcomes across a range of important domains of function.

The overall prevalence of access to health care coverage (93%) and receipt of health care services within the past year (90%) was high in our sample. Although these statistics are encouraging, the presence of unmet needs an average of 5 years after injury demonstrate that some individuals with TBI experience difficulty resolving their persisting problems despite access to the health care system. As the Military Health System (MHS) looks ahead to supporting a new generation of service members and veterans with TBI, the
findings in this study regarding prevalent needs after TBI can inform the MHS regarding the anticipated need for specific services, and guide resources to where they are likely to be most needed to improve long-term outcomes and lower the economic, personal, and social costs for the population of military service members with TBI, specifically to the areas of information, cognition, affective symptoms, and job skills. It is noteworthy that the landscape of TBI awareness, diagnosis, and treatment has changed since the participants in the present study sustained their injuries (1999–2003). In general, awareness of TBI and its sequelae in the military has increased in the wake of OEF/OIF casualties, and the subsequent recognition of TBI as the “signature wound” of recent conflicts.45 The findings of this study provide a pre–OEF/OIF-era point of reference against which to compare outcomes in future studies aimed at evaluating the impact of changes in the MHS for individuals with TBI.

**Methodological Considerations**

From a methodological standpoint, this investigation demonstrates the use of a registry database (as opposed to hospital records) for recruitment in a mobile military population years after treatment for TBI. The benefit of this method is that it allowed inclusion of individuals with less severe injuries (48% had mTBI), and who were not hospitalized for TBI (58% of the sample). Most long-term outcome studies of postinjury needs have been hospital based18–21 and thus not representative of individuals who have sustained less severe TBI. However, the registry data were limiting in this investigation, in that a majority (52%) of the individuals who met study criteria were not locatable due to the unavailability of current contact information an average of 5 years after injury. Existing literature on recruitment selection bias does not appear to address the consequences of noncurrent contact information for potential study participants, but it is worth exploring possible implications to this investigation. It is possible that service members who were not able to be located and invited to participate in the study may have been more mobile, relocating since their TBI due to new duty assignments or other work or personal opportunities. The fact that there were significantly more participants who were over 40 years of age compared to nonparticipants may be an indication that the increased geographic mobility of younger individuals contributed to their being less likely to be included. Since it is plausible that younger, more geographically mobile persons may be generally healthier than those who are older and less likely to relocate for work or personal reasons, there may have been a selection bias toward those with more long-term postinjury problems. This caveat regarding generalizability to the entire military population with TBI should be considered. It should also be noted that inaccuracies were discovered when potential participants were contacted to verify eligibility (11 individuals with inaccurate injury dates; six who denied having TBI, and therefore may have been misdiagnosed in the registry). As data repositories invariably contain inaccuracies, verification of registry information is necessary to ensure the quality of data used for research.

A 2012 hospital-based study of veterans’ caregivers described a similar phenomenon with respect to unreachable potential study participants, but to a lesser degree. In that sample of veterans who received inpatient rehabilitation for polytraumatic injuries an average of 4 years earlier, 20% could not be located despite exhaustive efforts similar to those employed in the present study.46 The difference in nonlocatable proportions between that study and the present investigation may be due to greater severity of the polytrauma patients’ injuries, which may have limited their postinjury geographical mobility to a greater extent than the individuals in the present study, resulting in a lower proportion of patients relocating since their last contact with the rehabilitation facility.

Hospital-based patients may thus be somewhat easier to locate than those treated for TBI on an outpatient basis. However, information on outcomes of individuals with TBI who were not hospitalized were underrepresented in the body of current research, and registry-based tools such as the one used in the present investigation—despite their limitations—offer an opportunity to include individuals with TBI who were treated in clinic or emergency room settings as well. The recruitment experience of this investigation offers a point of reference for conducting such a study in a military population. Future registry-based study designs should consider oversampling individuals who are younger and/or more geographically mobile in order to provide the power needed to analyze this subgroup, and consider periodic contact with registry subjects in order to reduce loss to follow-up for outcome studies.

**Limitations**

This study employed a standardized modified interview that had been used in a large-scale civilian outcome study, facilitating comparison to prior results. However, as with other studies,17–20 the lack of a non-TBI-injured control group is a limitation of this investigation. Also, the small sample size reduced the ability to detect potentially statistically significant differences in some analyses. This investigation relied on self-reported data from participants, collected at a single time point post injury, and without reference to preinjury or previous postinjury reports or assessments. Longitudinal data on the participants’ trajectory of postinjury problems and needs were not available.

As discussed, difficulties locating individuals may result in a selection bias toward those whose geographical mobility may be more restricted due to postinjury limitations as compared to those who could not be located. Findings of this study might have limited generalizability to the general military population with TBI, being more representative of
older individuals and those more likely to experience persisting difficulties than those who had uneventful recoveries from TBI. On the other hand, the findings may also underrepresent postinjury difficulties and service needs to a certain extent, because participants had to be cognitively competent enough to understand the purpose and procedures of the study and to provide informed consent in order to be interviewed. In this investigation, 22 of 149 (15%) individuals contacted by telephone were screened for inclusion in the study either failed the assessment of comprehension or were unable to provide informed consent. It is plausible, given the association between cognitive difficulties and increased service needs found in this study, that the individuals who were too impaired to participate experienced postinjury difficulties and service needs at a higher level than those who were interviewed. Finally, caution should be exercised in attributing unmet needs and postinjury difficulties solely to TBI, particularly in an investigation such as this, in which participants were interviewed an average of 5 years post injury. Comorbid health and psychiatric problems, whether related to the TBI or not and life events experienced since sustaining TBI also may have influenced the unmet service needs reported by participants.

CONCLUSION

Sixty-five per cent of this predominantly non-hospitalized, non–active duty, and non–combat-injured military sample with mixed TBI severity had at least one unmet need for assistance or services an average of 5 years post injury. Service needs in this sample were generally endorsed at higher rates than in civilian hospitalized samples. Participants reporting unmet needs were disproportionately symptomatic, had mental health diagnoses or treatments, needed some level of assistance with daily activities, perceived themselves as being disabled, were more likely to identify with fair to poor health, and had lower satisfaction with life. These results suggest that postinjury problems and perceived needs for services following TBI across severity spectrum may persist for years after injury in military and former military service members. These results support the need for the military programs aimed at following service members with military service members. These results support the need for the military programs aimed at following service members with TBI to detect ongoing needs for services and treatment, and military programs aimed at following service members with TBI to detect ongoing needs for services and treatment, and improving cognition, TBI to detect ongoing needs for services and treatment, and improving cognition, to improve the quality of life for service members postinjury. These results support the need for the military programs aimed at following service members with TBI to detect ongoing needs for services and treatment, and improving cognition, to improve the quality of life for service members postinjury.

ACKNOWLEDGMENT

This study was funded by the Defense Centers of Excellence, Defense and Veterans Brain Injury Center.

REFERENCES

Development of a Portable Tool for Screening Neuromotor Sequelae From Repetitive Low-Level Blast Exposure

Christopher K. Rhea, PhD*; Nikita A. Kuznetsov, PhD*; Scott E. Ross, PhD, LAT, ATC, FNATA*; Benjamin Long, MS*; Jason T. Jakiela, MS*; Jason M. Bailie, PhD†‡; Matthew A. Yanagi, MS§; LT F. Jay Haran, MSC, USN||; W. Geoffrey Wright, PhD¶; Rebecca K. Robins, PhD¶; CDR Paul D. Sargent, MC USN**; CDR Joshua L. Duckworth, MC USN††

ABSTRACT Blast exposure is a prevalent cause of mild traumatic brain injury (mTBI) in military personnel in combat. However, it is more common for a service member to be exposed to a low-level blast (LLB) that does not result in a clinically diagnosable mTBI. Recent research suggests that repetitive LLB exposure can result in symptomology similar to symptoms observed after mTBI. This manuscript reports on the use of an Android-based smartphone application (AccWalker app) to capture changes in neuromotor functioning after blast exposure. Active duty U.S. Navy personnel (N = 59) performed a stepping-in-place task before repetitive LLB exposure (heavy weapons training), and again immediately after, 24 hours after, and 72 to 96 hours after the completion of the training. The AccWalker app revealed that there are changes in neuromotor functioning after LLB exposure (slower self-selected movement pace and increased stride time variability) in participants who experienced neurocognitive decline. These data suggest that neurocognitive and neuromotor decline can occur after repeated LLB exposure.

INTRODUCTION

Traumatic brain injury (TBI) has received considerable attention within the military, as the number of injuries sustained has increased during Operations Iraqi Freedom and Enduring Freedom. The majority of research on TBI has focused on mild traumatic brain injury (mTBI), synonymous with concussion, to severe TBI. Two primary goals within the military TBI literature are to provide information to enhance a clinician’s ability to 1) identify functional deficits to aid in the clinical diagnosis of a mTBI and 2) make return-to-duty decisions. Much has been learned in this area and military specific guidelines have been developed to provide guidance and recommendations to clinical personnel in regard to mTBI screening, diagnosis, and injury management. These guidelines require mandatory neurological and functional evaluations, which focus on changes sequela following a clinically diagnosed mTBI, including subjective symptomology, neurocognitive functioning, and neuromotor functioning (i.e., assessment of gait and/or balance).

The aforementioned guidelines focus on the identification of mTBI following a head trauma event and very little is known regarding the effects of subclinical head perturbations. For example, do repetitive subclinical head perturbations result in cumulative functional deficits over time similar to those observed in a diagnosable mTBI? The majority of subclinical head perturbations experienced by military members are a result of mechanisms similar to those seen in the sport domain (i.e., blunt-force trauma). However, a significant number of military personnel are exposed to subclinical head perturbations as a result of blast exposure associated with training. Blast exposure presents a potential risk for periventricular injury, rather than the diffuse axonal injury from direct impacts that result from blunt-force trauma, and which can cause a different neural cascade and may potentially result in a different presentation of sequelae. Further, recent research has shown that blast exposure can lead to scarring across multiple interfaces in the brain, including the subpial glial plate, penetrating cortical blood vessels, gray-white matter junctions, and in the structures lining the ventricles.

It is more than feasible that the aforementioned injury mechanisms can also result in subclinical head perturbations that do not lead to a diagnosable mTBI. What is currently unclear due to limited research, is whether the effects of repetitive blast exposure, especially low-level blast (LLB) exposure, experienced by military members during their training and operational careers has a cumulative effect similar to those associated with repeated subclinical head
impacts in the sports domain. It has been suggested that even mild brain perturbations can impair performance, and the potential sequelae of repetitive exposures require further attention. As such, personnel at military breaching training schools have been studied in recent years to examine if there are any cumulative effects of repetitive LLB. Breacher training environments present a unique opportunity for research, as the environment isolates overpressure exposure as the primary mechanism of subclinical insults to the brain because it is unlikely that blunt-force or penetrating injury will occur.

In the seminal studies on the effects of repetitive LLB in breachers, neurocognitive deficits and changes in functional neuroimaging during a memory task were reported in only the training staff personnel with a chronic exposure history (i.e., number of blasts). Stone et al reported differences in neuroimaging between breacher training course instructors with 7 to 15 years of experience when compared with personnel attending the 2-week course. Carr et al reported that breachers with chronic blast exposure history self-reported more mTBI-related symptoms. The authors also reported more severe symptoms when compared to a nonbreacher cohort, and a history of shoulder-fired weapons was found to be the most reliable predictor of symptomology outside of blast exposure history. The results of these studies suggest that chronic LLB exposure may result in changes in symptomology, neuroimaging outcomes, and neurocognitive functioning.

To date, only one study has focused on the effects of repetitive short-term LLB exposure using a within-subjects design. Tate et al reported that military personnel attending 2-week breaching training course had changes in symptomology, biomarker loading, as well as neurocognitive deficits following course completion compared to baseline measurements. This result is considerably important, as most repetitive LLB exposures occur in operational environments. It is critical that military personnel who may have suffered any severity of brain injury are properly identified near the time of injury so they can receive appropriate care as early as possible, and so that critical return-to-duty decisions can be made with appropriate clinical guidance. This highlights the need for an objective means to diagnose and triage blast-exposed personnel in a field-based setting to not only screen for a possible mTBI, but also to identify any subclinical effects of the blast.

Much attention and funding has been devoted to developing field-deployable neurocognitive assessment tools (i.e., Defense Automated Neurobehavioral Assessment). However, the ability to measure neuromotor function in a similar manner is currently lacking. The latest concussion consensus statement urged for inclusion of a gait and balance assessment to aid in clinical decision-making about neuromotor function. To meet this challenge, we developed the AccWalker app (an Android-based smartphone application) as an objective, portable, field-based, and cost-effective tool for screening neuromotor sequelae following an insult or perturbation to the brain (i.e., blunt-head trauma or LLB exposure). The AccWalker app captures the acceleration profile of the lower extremity during a stepping-in-place task, which is used to derive several metrics of gait timing as an assessment of neuromotor function. A stepping-in-place task was selected as a dynamic balance activity that is a surrogate of gait.

The purpose of this project was to examine changes in neurocognitive testing and neuromotor functioning of military personnel who have been exposed to multiple subclinical head perturbations as result of their participation in a military heavy weapons (i.e., shoulder-fired weapons) training course to better understand the effects of repetitive short-term LLB exposure. Since neurocognitive testing is a well-accepted method to identify performance decline after head perturbations, our primary question was whether neuromotor performance decline was also present in this population. This was examined by splitting the studied population into two groups (with and without neurocognitive decline) and then determining whether neuromotor decline was also present. The rationale behind this design is that head trauma can lead to neurocognitive and/or neuromotor decline, but it is presently unclear if a decline in one domain (i.e., neurocognitive) generally leads to a decline in another domain (i.e., neuromotor). The reason the literature is unclear is two-fold: 1) most studies only focus on assessment in one domain after head trauma and 2) studies that have examined both domains have primarily focused on populations with blunt-force trauma leading to a concussion, not populations with repeated subclinical head perturbations. If the latter group experiences changes in neuromotor performance along with neurocognitive decline, that is important knowledge to discern, as it suggests that physical performance may suffer in this population in addition to cognitive performance. That knowledge could also be useful for medical professionals to ensure they more fully understand the service member’s performance deficits and engage them in an appropriate treatment plan. Thus, the goal for this article was to determine whether neuromotor performance declines are observed after neurocognitive decline has been identified, not if one assessment method is more useful than the other in screening for mTBI. The neurocognitive data were used to split the studied population into two groups and then the rest of this manuscript focuses on the neuromotor testing. Our hypothesis was that the neuromotor decline would be observed in participants who exhibited neurocognitive decline and vice versa.

**METHODS**

**Participants**

Active duty U.S. Navy personnel \(N = 90\) who participated in a 21-day Desert Warfare Training Program completed the informed consent process as approved by the Institutional Review Board at Naval Medical Center, San Diego. From the total sample, 60 were trainees newly exposed to the heavy weapons training (as described below), 16 were control subjects who participated in the program, but did not
take the heavy weapons training, and 14 were Range Safety Officers. Four control subjects and one trainee were excluded from the analysis because they were only measured at baseline and left the training afterward due to medical or family reasons. The neuromotor data were also missing for one control subject in the baseline testing (BSL) and 5 control subjects in the immediate post-training testing (see Design section for the description of the measurement points). For the purposes of this article, we focus on the performance of trainees only (N = 59) because the number of measurements from the control and Range Safety Officer subjects was too little to break down by neurocognitive status.

Trainees were recruited voluntarily by research staff not associated with the military training program or by ranking members of their service. The mean age was 26.3 ± 3.5 years and all subjects were men. The mean time in military service at the beginning of the program was 4.2 years (SD = 2.9). Fourteen subjects have been previously deployed with a median number of deployments of 2 (maximum 7). The majority of participants were Caucasian (N = 55), whereas other participants were American Indian/Alaska natives (N = 2) and Hispanic (N = 2). The level of education in the sample was split into the following categories: no degree (N = 14), high school diploma (N = 1), some college (N = 10), Associate’s degree (N = 4), Bachelor’s degree (N = 28), and Master’s degree and higher (N = 2). A subset of the participants (N = 28) self-reported having had a concussion in the past. In the majority of cases (all but 1) the concussion occurred more than 6 months before the training. Median number of self-reported concussions was 2.

Design
The Desert Warfare Training Program (heavy weapons training) lasted 21 consecutive days. This training included repetitive LLB exposure as part of the rocket fire training, which consisted of repetitive firing from shoulder-mounted rocket launchers such as M2CG 94mm (Carl Gustaf), M72 LAW 66mm (Light Anti-Tank Weapon), and RPG (Rocket-Propelled Grenades) with varying munitions.

On the second day of training, before any blast exposures, participants completed BSL in a battery of neurocognitive tests that included Hopkins Verbal Learning Test-Revised, components of the Defense Automated Neurobehavioral Assessment (Simple Reaction Time, Procedural Reaction Time, Go/No-Go), Trail Making Tests Parts A and B (TMT A & B), King-Devick Test (KDT), and performed the neuromotor test (stepping-in-place task with the AccWalker app). Approximately 7 days later, participants took part in the shoulder-mounted rocket launcher training. During this training segment, participants were outfitted with pressure Blast Gauge sensors (BlackBox Biometrics, Inc., Rochester, New York) placed on the anterior and posterior of the helmet, on the shoulder, chest, and back. The median number of blasts was 4 (minimum = 2 and maximum = 9). The maximum peak of pressure that occurred during training was 5.5, 7.3, 5.9, 4.8, and 8.5 pounds per inch (PSI) for the anterior head, posterior head, shoulder, chest, and back, respectively. The sum total impulse of the blasts was calculated as the integral of the positive overpressure data from the sensors. The total impulse that occurred during training was 7.5, 19.4, 15.7, 7.8, and 15.6 PSI-ms for the anterior head, posterior head, shoulder, chest, and back, respectively. Within 30 minutes on the completion of their initial rocket launcher training day, participants were tested on the neurocognitive battery of tests and on the neuromotor (i.e., stepping-in-place) test again (POST-1).

Individuals were classified with neurocognitive decline following blast exposure if their performance on neurocognitive testing in POST-1 was indicative of a reliable and significant change based on a priori cut-off scores on four measures of neurocognitive function: Hopkins Verbal Learning Test (Total Recall and Delayed Recall)30 and Trail Making Test Part A and Part B.31 A reliable change index (RCI)32,33 was calculated for each cognitive measure; scores below the lower limit of a 90% confidence interval were indicative of decline. Reliable change indices were derived by estimating measurement error surrounding test–retest difference scores. Specifically, the standard error of the difference (Sdiff) was used to create a confidence interval for the baseline–retest difference score. The formula used for calculating Sdiff employed the standard error of measurement (SEM) for baseline (SEM1) and the associated retest (SEM2): Sdiff = \sqrt{SEM_1^2 + SEM_2^2}. SEM1 and SEM2 were based on published normative data on the psychometric reliability of each measure. Participants classified with neurocognitive decline (29 out of 59 participants) were additionally tested 24 hours post-training (POST-2). All participants were tested again at 72 to 96 hours post-training (POST-3) (Fig. 1). This window was used due to participant availability for POST-3 testing.

Since the assessments for this project detracted for the 21-day Desert Warfare Training Program in which the subjects were participating, it was decided that all subjects would be tested at before LLB exposure (baseline, BSL) and immediately after LLB exposure (POST-1). To not further detract from their training, only participants who showed neurocognitive decline at POST-1 were tested 24 hours later (POST-2) to see if they still had their declined neurocognitive performance. All participants were then recruited back for testing 72 to 96 hours later (POST-3) to provide a third time point of assessment for all participants to measure performance before and after LLB exposure, and a fourth time point for those who exhibited neurocognitive decline to determine whether their neurocognitive performance returned to baseline levels. The specific focus of this article was to assess the neuromotor performance trajectory before and after LLB exposure in participants with and without neurocognitive performance declines.
Materials

Six Android-based phones (Model GT-S7710L; Samsung Galaxy Xcover 2; Samsung, Seoul, South Korea) with the AccWalker app installed were used for the measurement of leg acceleration during the stepping-in-place task. The phone recorded acceleration along the x-, y-, and z-axes of the phone (Fig. 1A). The acceleration data were sampled at 96 Hz.

Procedures

During the neuromotor portion of testing, subjects were instructed to step-in-place at a comfortable walking pace for 120 seconds. The instruction was “Please walk in place” and the experimenter made sure they did not turn during the task and that they moved their arms in a symmetric fashion (Fig. 1A). The phone-based accelerometer was generally placed on the lateral side of the thigh, but in some cases (approximately 12.6% of total), size of the thigh did not permit proper placement of the strap and the accelerometer was placed on the shank. This difference in placement, however, did not affect the calculation of movement timing variables during stepping in place because they both contained relevant timing landmarks (described below).

Dependent Measures

Preferred movement speed and variability of timing are commonly used to examine neuromotor ability in healthy and concussed individuals.34–43 Accordingly, we examined stride time as a measure of neuromotor performance, which is defined as time between two equivalent landmarks during gait cycle of the same limb (i.e., heel strike, peak knee flexion, or any other identifiable landmark). We used the acceleration profile recorded by the AccWalker app to derive stride time mean, stride time SD, and its coefficient of variation (CV) to characterize timing performance during stepping-in-place. The CV (CV = SD/mean) was used because different individuals had different mean stride times in our sample, indicative of individual pacing preferences typically present for locomotor activities. CV provides a relative measure of

FIGURE 1. (A) Stepping-in-place task used for neuromotor testing. The las frame in the movement sequence shows the coordinate system of the phone with the AccWalker app in the sagittal plane. In the majority of trials, phone was placed on the thigh (as depicted in the figure). In a subset of trials (~12.6%), it was placed on the shank midway between the ankle and the knee, but this did not affect stride time estimation (see Methods section); (B) study design (POST-2 was only completed by participants with detected neurocognitive decline, n = 29); (C) relation between the knee angle recorded via motion capture and thigh acceleration/velocity recorded by the AccWalker app; only phone acceleration was recorded by the phone in this study; (D) phone velocity on the leg was derived from recorded acceleration via integration (see Methods section). Stride time was defined based on velocity minima (peak velocity during leg return to stance); and (E) time series of the stride time as a function of step number within a single 120-second trial.
magnitude of variability that is independent from the mean stride time.

To obtain stride time measurements from the phone’s accelerometer recordings, we used the following steps: 1) phone’s acceleration along the z-axis of the phone was filtered with a fourth-order 5 Hz low-pass Butterworth filter, 2) filtered acceleration was integrated to obtain velocity, 3) velocity was high-pass filtered using a 0.25 Hz third-order polynomial adaptive filter to remove the effects of accelerometer drift, and 4) stride time was identified based on timing between the consecutive minima in the velocity time series (Fig. 1D and E). Minimum velocity occurred when the leg was in midway of returning to the stance phase, between the maximum knee flexion and maximum knee extension (Fig. 1C). This landmark was used because it was robustly identifiable across different subjects and phone placements (thigh or shank). Using other landmarks (such as maximum velocity) did not alter the results. All computations were performed in Matlab 2015b (Mathworks, Natick, Massachusetts).

Statistical Analysis
We used a linear mixed model (LMM) to compare stride time measures as a function of testing time: BSL (baseline), POST-1 (immediately after weapons training), and POST-3 (72–96 hours post-training) and as a function of group assignment (neurocognitive decline vs. no-neurocognitive decline). POST-2 data were excluded from the LMM since only neurocognitive decline group data were collected. The $p$ values for the main and the interaction effects were extracted by comparing the goodness of fit of the model with and without the effect of interest. A $p$ value of $<0.05$ was considered as statistically significant. Post hoc tests were conducted using simple contrast comparisons using the Kenward–Roger approximation to the degrees of freedom. All analyses were conducted in R using lme4 and pbkrtest packages. The R code (The R Foundation, Vienna, Austria) for the full model was: lmer(DV~Time*Group+(1|Subject),data,REML=0), where DV is the dependent measure, Time is factor specifying measurement time, and Group specifies neurocognitive status group. Subjects were specified as random effect. The covariance structure was set to be compound symmetric. The model was fit using maximum likelihood estimation.

We also calculated the statistical effect sizes for between- and within-groups across measured time points. The between-group effect size was defined using Cohen’s $d$ with the variance pooled across the two groups using root mean square. The within-group effect size was extracted by dividing the mean difference scores between the sessions by the SD of the difference scores.

RESULTS
Because of time constraints and other training environment factors, some participants missed some testing sessions. These observations were missing at random in the BSL (12 out of 59 missing, $N = 47$ tested), POST-1 (5 out of 59 missing, $N = 54$ tested), and POST-3 (10 out of 59 missing, $N = 49$ tested) sessions. Five observations showed substantial slowing down or speeding up of stride time during trial and these observations were removed from the analysis because drift affects the calculation of stride time-dependent measures, which require stationarity. See Figure 2 for the final number of measurements per group at each time point.

The results of the LMM analysis revealed a main effect of testing for mean stride time ($p < 0.001$). Post hoc $t$ tests showed statistical differences between the BSL and POST-1, $t(131.46) = -6.35$, $p < 0.001$, $d = 0.73$, and the BSL and POST-3, $t(131.46) = -5.41$, $p < 0.001$, $d = 0.57$, indicating that mean stride time was greater at baseline than immediately after and 72 to 96 hours after LLB exposure. There was no main effect of neurocognitive status ($p > 0.05$) and no interaction effect ($p = 0.36$). However, the between-group effect sizes indicated that the difference between the
neurocognitive decline and no-neurocognitive decline groups was greatest in the POST-1 condition (see Fig. 2A), suggesting that the group with the neurocognitive decline showed relatively slower stepping pattern than the group without neurocognitive decline.

The main effect of time for stride time CV was also significant ($p < 0.001$). There were statistical differences between the BSL and POST-1, $t(153.06) = -3.67$, $p < 0.001$, $d = 0.57$ and BSL and POST-3, $t(153.06) = -3.17$, $p = 0.001$, $d = 0.46$ (Fig. 2B). These results indicated that relative variability of movement timing was greater at baseline than immediately after and 72 to 96 hours after LLB exposure. There was also an interaction effect ($p = 0.014$) such that the difference between the groups was greater in POST-1 compared to baseline and POST-2, indicating that the neurocognitive decline group was more variable than the no-neurocognitive decline group. Between-group effect sizes across time points for the neurocognitive decline group vs. no-neurocognitive decline group are reported in Figure 2B.

**DISCUSSION**

The goal of article was to determine whether neuromotor performance declines are observed after neurocognitive decline has been identified after repeated LLB exposure. This goal was accomplished by developing a novel, objective, portable, field-based, and cost-effective tool to measure neuromotor function in a dynamic balance test in the context of military training environment. Our results showed that trainees with identified neurocognitive decline after LLB performed the stepping-in-place task slower and with a higher level of variability in stride time immediately after exposure to LLB compared to trainees without neurocognitive decline. Although both groups became faster and less variable on the stepping-in-place task as a function of repeated neuromotor testing, the relative divergence of performance immediately after LLB exposure suggests that neuromotor function can decline similarly with the neurocognitive performance after repeated subclinical head perturbations.

The overall increase of the stepping pace (lower mean stride time) and a decrease in stepping variability (lower stride time CV) from baseline to immediate testing sessions likely reflects a practice effect with this task in addition to any LLB effects. This interpretation is suggested by our recent follow-up work where we tested young healthy civilians not exposed to blasts or other potentially concussive events. In that civilian cohort, stride time CV decreased from the first to the second measurement to a similar extent (about 1%) as in the current dataset, where stride time CV was about 4.5% in the first test session to 3.5% in the second. Thus, practice effects should be accounted for in future research to more accurately determine the extent to which LLB alters neuromotor function. This could be done by mapping out the practice effect trajectory in the neuromotor task and including practice trials in future assessments of this test to negate the practice effect.

Even with the practice effect present, there was a clear divergence between groups in the mean stride time and stride time CV when they were tested immediately after range training (POST-1), suggesting repeated LLB exposure can lead to concurrent neurocognitive and neuromotor decline. It could be argued that the magnitude of change from baseline to POST-1 in the no-neurocognitive decline group reflects the practice effect from performing the stepping-in-place task a second time. Accordingly, a smaller change in movement pace and variability in the neurocognitive decline group could indicate less adaptive and slower practice effect in this task. As a result, subjects with identified neurocognitive decline were slightly slower (by about 100 milliseconds) and more variable in their performance than the no-neurocognitive decline group during testing immediately following range training (POST-1). Both slower pace of movement and increased movement variability are typically observed gait alterations in individuals with TBI. Collectively, our findings support previous research showing that cognitive ability predicts motor learning in patients with TBI. It should be noted that the group differences in neuromotor performance dissipate by the POST-3 testing, suggesting that the effect of repeated LLB exposure on neuromotor control lasts less than 72 to 96 hours. However, the cumulative effect of chronic exposure to repeated LLB exposure remains an empirical question.

The finding that repeated LLB exposure is also associated with a decline in neuromotor performance is inconsistent with previous research showing that LLB exposure does not affect vestibular function. However, repeated subclinical head perturbations have been shown to chronically affect neurological functioning in the sports domain, so follow-up work with military personnel to track the effects of repetitive LLB exposure on acute and chronic neuromotor performance is warranted. Further, it should be noted that the blast level in this study was very low (5.5 and 7.3 PSI at the anterior and posterior of the head, respectively), which is only slightly above the recommended safety standard of 4 PSI. Thus, blast magnitude and the number of blasts should be factored into future studies to determine whether they lead to similar or divergent neurocognitive and neuromotor performance declines. The orientation of the head with respect to the LLB wave (i.e., consistently on the right or left side of the blast) may also influence the effect of LLB exposure neurocognitive and neuromotor performance.

There are a number of ways the neuromotor test described in this manuscript could be improved to increase its sensitivity to identify changes in neuromotor performance following subclinical head trauma. First, the level of difficulty in the stepping-in-place task could be increased to enhance the test's sensitivity. The task was completed with eyes open, so visual feedback about step timing and orientation in the environment was available. Performing the task with eyes closed would remove visual feedback, which has been shown to be a useful method in discriminating between participants with and
without head trauma.\textsuperscript{49} Moreover, perturbing the vestibular system during the stepping-in-place task could also increase the sensitivity, as vestibular dysfunction affecting balance is common after head trauma.\textsuperscript{50,51} Finally, previous research has suggested that gait velocity and medial–lateral range of motion of the trunk (or center of mass) during gait are strong indicators of neuromotor dysfunction after a concussion.\textsuperscript{52–55}

We measured leg movement timing, whereas trunk control, especially in the medial–lateral direction, may be affected in this subclinical population.

A number of limitations exist in this study. First, the most important limitation is that there is no control group that did not receive any LLB during training. We acknowledge that any change in neuromotor or neurocognitive performance could arise from a host of factors, including practice effects and other physical status changes (arousal, sleep deprivation, dehydra-tion, etc.). There was limited data that could have been used a control group in this dataset, but it was not included in the results due to a low number of subjects in the session immediately following LLB (n = 7) and also because these subjects participated in training with different physical demands from the LLB-exposed trainees. Second, although some of the military personnel in our sample exhibited a significant decline in neurocognitive performance, a medical doctor was not present to evaluate any potential diagnosis of a concussion. Even though all of the blast waves were below known concussive thresholds, it is possible that some participants could have presented with concussion symptoms had they been examined by a medical doctor. Thus, we cannot definitively say that all participants experienced only subclinical perturbations.

In conclusion, we presented data from the first step in developing an objective, portable, field-based, and cost-effective tool to measure neuromotor function in a dynamic balance test. We showed that neuromotor decline accompanied neurocognitive decline in a subset of participants who were repeatedly exposed to LLB from heavy weapons training. This suggests that neurological dysfunction affects multiple domains of performance, which should be taken into account when deciding on appropriate medical care. It is especially important to note that all participants in this study were exposed to subconcussive LLB, adding to a growing body of research showing that repeated subclinical head trauma can affect neurological functioning.\textsuperscript{13–19} Finally, although these early data are encouraging, the neuromotor assessment methods presented here are still in development and ultimately will require validation with medical outcomes data in order to have clinical utility.

\textbf{ACKNOWLEDGMENTS}

The U.S. Army Medical Research Acquisition Activity, 820 Chandler Street, Fort Detrick MD 21702-5014 is the awarding and administering acquisition office. This work was supported by the Office of the Assistant Secretary of Defense for Health Affairs under award no. W81XWH-15-1-0094 to Christopher K. Rhea and HU0001-14-1-0022 to Joshua L. Duckworth.

\textbf{REFERENCES}


A Retrospective Cohort Analysis of Battle Injury Versus Disease, Non-Battle Injury—Two Validating Flight Surgeons’ Experience

Col William P. Butler, USAF MC (Ret.)*‡; Col Lawrence W. Steintraus, USAFR MC†; Brittany L. Fouts, MS*; Jennifer L. Serres, PhD*

ABSTRACT Today, military combat medical care is the best it has ever been. Regulated U.S. Air Force aeromedical evacuation (AE) is one important reason. The Theater Validating Flight Surgeon (TVFS) validates that a patient is ready for flight. Two TVFSs’ experiences, successfully deployed in 2007, are the focus of this study. A unique operational worksheet used to manage the AE queue was used for approximately 5 months. A descriptive analysis of the worksheet’s 1,389 patients found the majority male (94%), median age 30 years, and mostly Army enlisted soldiers (63%). U.S. civilians made up 9%. Battle Injury (55%) surpassed Disease, Non-Battle Injury (45%); most frequently seen were extremity injuries (73%) and cardiac illness (31%), respectively. Common to both Battle Injury and Disease, Nonbattle Injury were several TVFS prescriptions including no “remain overnights” (79%), head of bed elevation (78%), cabin altitude restriction (57%), no stops (44%), Critical Care Air Transport Team (27%), and supplemental oxygen (22%). This study is a first look at the TVFS experience and it offers up an initial accounting of the TVFS clinical and prescriptive practices. It is also a jumping point for future TVFS investigations using the available AE databases.

INTRODUCTION

Today, military combat medical care is the best it has ever been; wound lethality is exceedingly low. In fact, lethality rates have been variously reported from 9.6% to 10.2% (U.S. Casualty Status, published February 12, 2015; Department of Defense; http://www.defense.gov/news/casualty.pdf; accessed February 17, 2015; website no longer operable).1,2 There are a number of reasons for this success. Care is farther forward and delivered at a higher level of skill with a greater degree of technological support than ever before. Additionally, Critical Care Air Transport Teams (CCATT) are moving “stabilized,” albeit very sick, patients in numbers never before seen (particularly during the study time frame). Indeed, en route CCATT mortality has been recorded as low as 0.2%.3 And lastly, aeromedical evacuation (AE) has become more elastic than ever before.

AE, as practiced by the U.S. Air Force, is a regulated, fixed-wing logistics function whereby patients are transported from one medical facility to another, most commonly to a facility that can offer more patient care resources. Two flight surgeons are involved in every patient movement. There is the on-site, or clearing, flight surgeon who examines the patient and determines his/her preflight preparation requirements and in-flight needs; this is often a junior flight surgeon. In contrast, the other flight surgeon is generally a senior, Aerospace Medicine specialty-trained physician; this is the so-called Theater Validating Flight Surgeon (TVFS) who has final oversight and approval authority for a given patient’s transport. TVFSs are regionally based—North and South America, Europe, the Pacific, and Central Command—and oversee patient evacuation for their respective region.

Each patient transport request is entered into the web-based TRANSCOM (Transportation Command) Regulating and Command and Control Evacuation System (TRAC2ES) by the originating medical facility. This entry is often input by the clearing flight surgeon (although it may well be entered by the attending physician or a patient administrative technician) and consists of standard administrative and clinical information forming the foundation for the Patient Movement Request (PMR). Once a completed PMR has been submitted, it undergoes a validation process that has two components. The first is an administrative validation. Here, the nonclinical aspects of the transport are validated as “good to go;” examples include presence of passports, diplomatic clearances for territorial overflight, and destination resources. Second is the clinical validation. Here, the TVFS, geographically separated from the patient, determines that the patient is clinically ready for flight. In order to minimize patient vulnerability at altitude, this validation may include prescriptions that apply to the patient (e.g., supplemental oxygen, head-first loading, and assignment of CCATTs) and those that apply to the aircraft (e.g., long, slow landing; no stops; cabin altitude restriction).4 Without the TVFS’s validation, the logistics requirement for a transporting aircraft cannot be established.

While deployed in 2007, two of the authors (WPB and LWS) served as successive TVFSs over an 8-month period. During that time they oversaw the movement of 8,600+ patients; 1,800+ were Urgent and Priority precedence patients. An Urgent patient requires transport as soon as possible to preserve life, limb, or eyesight, whereas a Priority patient requires transport within 24 hours to preserve life, limb, or eyesight.

*U.S. Air Force School of Aerospace Medicine, 2510 5th Street, Building 840, Wright-Patterson Air Force Base, OH 45433.
†Division of Preventive, Occupational, and Aerospace Medicine, Mayo Clinic, 200 First Street SW, Rochester, MN 55905.
‡Decypher Technologies, 200 Concord Plaza Drive 500; San Antonio, TX 78216.

Presented as a poster at the Military Health System Research Symposium, Fort Lauderdale, FL, August 17–20, 2015.
doi: 10.7205/MILMED-D-16-00166
Commonly, aircraft must be diverted from ongoing missions to move Urgent and Priority patients. Of note, Routine patients can wait until a regularly scheduled AE flight.4

Toward the end of the first deployment, as patient volume surged, it became clear that a streamlined local patient tracking tool was needed. An operational worksheet was created to facilitate managing the AE queue and simplifying the clinical validation process. It helped the TVFS keep track of the many patients and prescriptions concurrently being considered and was used for approximately 5 months. As a result, it did not include all of the patients validated during the two TVFS’s tours; it was particularly void of the less ill/injured Routine precedence patients. Indeed, the bulk of patients on the worksheet were Urgent and Priority precedence due to their clinical acuity and serious need for higher level care. The worksheet proved operationally useful in clinical decision making for both TVFSs.

Interestingly, although much has been written about transporting critically ill patients and the CCATTs, little can be found aimed at the TVFS. In fact, the authors’ are unaware of any study focused on the roles or experiences of the TVFS. As the worksheet offers a glimpse into two TVFSs’ validating practice, it was deemed a useful place to start. This report is a first look at the TVFS’s medical practice.

METHODS
This is a descriptive study of the authors’ experience as Central Command’s TVFS. Following approval by the Air Force Research Laboratory Institutional Review Board, a descriptive analysis of the 1,389 patients entered into the operational worksheet was performed. The patients entered into the worksheet were usually the sicker patients often requiring added laboratory or historical information and they often consisted of concurrent submissions of multiple patients. Since the worksheet was a tool to facilitate clinical validation and decision-making, patients and data fields were entered at the discretion of the TVFS. Consequently, only certain data from the PMR were extracted. This included some demographic characteristics, a few pertinent laboratory and physiologic parameters, and the anatomic system(s) affected by the patient’s clinical issue(s). In addition, there was a brief, less than one sentence, “reminder” clinical vignette and the prescriptions (patient and aircraft) levied by the TVFS. Notably, the worksheet, as opposed to the TRAC2ES database, differentiated between the initial provisional PMR submission and the final validated PMR record in several data fields: Precedence, AE Classification, CCATT assignment, and fraction of inspired oxygen (FiO2). This fact permits an analysis of the agreement between the clearing flight surgeon and the TVFS.

Data were analyzed using the Statistical Package for Social Sciences, version 22 (IBM SPSS Statistics for Windows, Version 22.0, IBM Corp., Armonk, New York). Descriptive statistics were applied to basic demographic characteristics and laboratory/physiologic parameters using number and percent for categorical variables and median with interquartile range (Q1,Q3) for continuous variables. Since Battle Injury (BI) and Disease, Non-Battle Injury (DNBI) were recorded, this natural dichotomy of patients lent itself as a platform for describing the clinical issues faced by the TVFS, their various etiologies, and the various interventions prescribed by the TVFS. Number and percent were recorded as descriptors and, where appropriate, Mann–Whitney U nonparametric and χ2 statistics were calculated (to compare the BI vs. DNBI physiological parameters and the TVFS’s BI vs. DNBI prescriptive practices, respectively). A p value of 0.05 was considered statistically significant. In addition, agreement between the clearing flight surgeon and TVFS was tested using the kappa statistic for categorical variables and the Mann–Whitney U nonparametric statistic for the one continuous variable (FiO2).5

RESULTS
During the two TVFS’s rotations, 8,634 patients were validated for AE (Urgent = 303, Priority = 1,500, Routine = 6,831). When precedence in the overall 8,634 patients was compared to that of the 1,389 patients in the worksheet, a statistically significant difference was detected ($\chi^2 = 1.726$, $p < 0.00001$) suggesting that the worksheet was not fully representative of the entire deployment.

When characterizing the cohort of 1,389 patients recorded in the worksheet, the majority was male (94%) and the median age was 30 years (22, 35; range 1 day to 67 years). Army made up 71% of the patients, much more than the Marines (6%), Air Force (4%), or Navy (2%); most were enlisted (90%). In addition, contractors (8%) and allies (2%) were regularly moved. Patient evacuations were mainly from Operation Iraqi Freedom (77%) and BI (55%) transports surpassed those for DNBI (45%). Priority precedence (63%) dominated, as did AE Classification 2a/2b (inpatients on a litter) (84%). Of note, there were only a few cancelled patients and only a few patients who died before transport. See Table I for further details.

Several key laboratory and physiological parameters were recorded in the worksheet: hemoglobin (Hgb) level (median = 13.0 g/dL; 11.0, 14.7; range 5.0–20.0); Hgb saturation (median = 98%; 98, 99; range 29–100); FiO2 (median = 40.0%; 21, 50; range 21–100); positive end expiratory pressure (median = 5.0 cm H2O; 5, 5; range 0–40); and, arterial oxygen partial pressure (PaO2, median = 128.0 mmHg; 97, 201; range 37–467). When comparing BI with DNBI, significantly lower Hgb, higher FiO2, and higher PaO2 were seen with BI. See Table II for further details.

Extremity issues made up 48% of transports found in the worksheet cohort followed by face–neck (18%), head (15%), chest (13%), abdomen (13%), spine (8%), and pelvis (6%). Six percent were burns, whereas 14% were cardiac. BI predominated in all anatomic systems except the cardiac where DNBI held sway (98%). In fact, Army soldiers (59%) and U.S. civilians (25%) had most of the cardiac issues and, generally, they proved older (Army soldier median = 38 years of age; 30, 45; range 19–58; U.S. civilian median = 50 years
of age; 47, 59; range 28–67). Indeed, the Army cardiac patients proved significantly older than the worksheet Army population (p < 0.00001). Similarly, the U.S. Civilian cardiac patients proved significantly older than the worksheet U.S. Civilian population (p < 0.0001). BI etiologies were relatively limited, the top four being coronary artery syndrome (CAS, 29%), infection (27%) of the worksheet’s patients, most being assigned to BI (BI = 231, DNBI = 149; χ² = 6.54, p = 0.01). Further analysis of the BI/DNBI dichotomy demonstrated significantly different prescriptive patterns. BI prescriptions dealt with the aftermath of trauma and surgery—most notably, spine precautions (p < 0.0001), C-collars (p < 0.0001), neurovascular checks (p < 0.0001), head first loading (p < 0.0001), long landing (p = 0.013), supplemental oxygen (p < 0.0001), transfusion (p < 0.0001), and postsurgical tube maintenance (p ≤ 0.004). On the other hand, DNBI prescriptions focused on infection and cardiac illness as seen with infectious precautions (p < 0.0001), N-95 mask (p < 0.0001), medical attendants (p < 0.0001), and cardiac monitoring (p < 0.0001). See Table IV for further details.

There were four actions (aka prescriptions) where the decision of the clearing flight surgeon and TVFS were recorded: Precedence, Classification, CCATT assignment, and FiO2 level (aka supplemental oxygen). Kappa statistic demonstrated moderate agreement in Precedence (kappa = 0.54, SE = ± 0.02, 95% confidence interval [CI] 0.50–0.58) and substantial agreement in both Classification (kappa = 0.67, SE = ± 0.02, 95% CI 0.63–0.71) and CCATT assignment (kappa = 0.74, SE = ± 0.02, 95% CI 0.70–0.78). For all intents and purposes, there was disagreement on the initial inflight FiO2 prescription. A close examination of FiO2 demonstrated a significant difference between that prescribed by the clearing flight surgeon and the TVFS (p < 0.0001). Indeed, the TVFS prescribed a higher level FiO2 (median difference between the TVFS and clearing flight surgeon = 10%; 9, 15; range –54 to + 64).

**DISCUSSION**

In today’s conflicts AE of very ill patients is commonplace. This is a direct result of an increased emphasis on forward-based care, damage control surgery, and CCATTs, all superimposed on the echelon of care system. To make this paradigm work, a robust AE mission capability is essential. Indeed, it must be prepared to transport “stabilized” or physiologically volatile patients. Thrusting such patients into an aircraft cabin with its various environmental stressors (acceleration, low humidity,
thermal instability, vibration, noise, hypobaria, and hypoxia) is physiologically challenging. Unfortunately, the precise physiological impact on these compromised patients has not been thoroughly studied. Despite this, the flight surgeon must configure the patient-aircraft environment to minimize clinical risk. This is initially addressed on site by the junior clearing flight surgeon, but it is finalized by the more senior TVFS. In fact, it is the TVFS who warrants that a given patient is optimally prepared to withstand these stressors of flight. This study looked specifically at the TVFS.

This study is a first look into the clinical practice of the TVFS. The study’s findings describe some of the experiences of two TVFS’s deployed during 2007, as recorded in an operational worksheet employed during the busiest part of their deployments. The worksheet offers a depiction of a large number of the patients they saw, the spectrum of injury and illness they encountered, and the range of interventions they prescribed to minimize the negative impact of flight. It also provides a first look at the clearing versus validating flight surgeon relationship.

As expected, being 2007, the vast majority of evacuations recorded in the worksheet were young, enlisted Army soldiers from Iraq, reflecting the operational emphasis at the time. Surprisingly, BI (55%) exceeded DNBI, as opposed to Harman’s report where DNBI made up 86% of the evacuations. This observation probably reflects the nature of the worksheet. It was devised to facilitate queue management of the stabilized patient. Most such patients are quite ill requiring evacuation to preserve life, limb, or eyesight (i.e., Urgent or Priority precedence). In fact, just over a quarter of the worksheet’s patients required CCATT team assignment (BI = 61%, DNBI = 22%). With such a high number of CCATT patients, it should not be surprising to see an elevated degree of BI. Indeed, Bridges, looking at CCATT transports from 2001 to 2006, found 64% of those patients suffered BI. Relatively few Routine patients were entered into the worksheet further reflecting the nature and purpose of the worksheet, not the overall patient population itself. In fact, only a very small number of Routine patients during the TVFS’s deployments went into the worksheet (0.5%). This, coupled with the preponderance of BI explains how most patients in the worksheet were AE Classified 2a/2b (inpatients on a litter).

As for the injuries and illnesses suffered by this worksheet cohort, the authors’ found the anatomic system(s) affected to be

---

**TABLE II.** Worksheet Laboratory and Physiologic Parameters as Extracted from the Patient Movement Request (N = 1, 389)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Number (%)</th>
<th>Battle Injury Median (Q1, Q3) (n = 762)</th>
<th>Disease, Non-Battle Injury Median (Q1, Q3) (n = 626)</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hemoglobin, g/dL (n = 1,184)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤8.0</td>
<td>21 (1.8)</td>
<td>12.6 (10.6, 14.0)</td>
<td>14.0 (13.0, 15.0)</td>
<td>&lt;0.0001*a</td>
</tr>
<tr>
<td>8.1–9.0</td>
<td>75 (6.3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9.1–10.0</td>
<td>104 (8.8)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10.1–15.0</td>
<td>856 (72.3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;15.0</td>
<td>128 (10.8)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hemoglobin Saturation, % (n = 1,276)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;90</td>
<td>7 (0.6)</td>
<td>98.0 (98.0, 100.0)</td>
<td>98.0 (97.0, 99.0)</td>
<td>&lt;0.0001 ab</td>
</tr>
<tr>
<td>91–94</td>
<td>33 (2.6)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>95–100</td>
<td>1,236 (98.8)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction of Inspired Oxygen, % (n = 357)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>95 (26.6)</td>
<td>40.0 (28.0, 50.0)</td>
<td>28.0 (21.0, 36.0)</td>
<td>&lt;0.0001 a</td>
</tr>
<tr>
<td>22–30</td>
<td>45 (12.6)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>31–40</td>
<td>108 (30.3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>41–50</td>
<td>55 (15.4)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>51–60</td>
<td>28 (7.9)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>61–99</td>
<td>13 (3.6)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>13 (3.6)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive End Expiratory Pressure, cm H2O (n = 186)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤5</td>
<td>143 (76.9)</td>
<td>5.0 (5.0, 5.0)</td>
<td>5.0 (5.0, 5.0)</td>
<td>0.967</td>
</tr>
<tr>
<td>6–10</td>
<td>36 (19.3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;10</td>
<td>7 (3.8)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arterial Oxygen Partial Pressure, mmHg (n = 227)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤50</td>
<td>3 (1.3)</td>
<td>143.0 (98.3, 204.8)</td>
<td>99.0 (80.0, 128.0)</td>
<td>0.006a</td>
</tr>
<tr>
<td>51–60</td>
<td>3 (1.3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>61–70</td>
<td>8 (3.6)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>71–80</td>
<td>15 (6.6)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>81–90</td>
<td>14 (6.2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>91–100</td>
<td>47 (20.7)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;100</td>
<td>137 (60.3)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*aDenotes statistical significance based on independent-samples Mann–Whitney U analysis. bHemoglobin saturations were significantly different despite identical medians, the difference having more mathematical than clinical relevance.
useful in the validation process. It provided a quick-look gestalt to help rank-order patients in the validation queue. That said, the BI/DNBI dichotomy proved valuable as a way of examining the various maladies encountered. The top etiology for all BIs was the IED. This was predictable as the IED is probably this conflict’s signature ordnance.\textsuperscript{8–11} Equally predictable was the high prevalence of extremity injuries (48\%), at least two-fold higher than any other anatomic system. This observation, to a great extent, replicates those seen in earlier reports (22–43\%).\textsuperscript{8,11} In fact, both Belmont and Mason put the prevalence of extremity injuries closer to 50\%.\textsuperscript{9,10} This almost certainly is a function of improved body armor systems, leaving the extremities as the least protected of all anatomic systems.\textsuperscript{9}

It was not so clear cut for DNBI. Many diagnoses were seen and those appearing with some frequency, excepting CAS, never rose above 8\%. CAS, however, dominated cardiac conditions and cardiac conditions dominated DNBI. Unexpectedly, Army soldiers predominated CAS followed by U.S. civilians. Although almost 70\% of the Army soldiers were under 30, those with CAS were almost a decade older (38 years); this was statistically significant. Similarly, the U.S. Civilians were over a decade older (50 years) than their worksheet peers; this, too, was statistically significant. These results reinforce the generally accepted "age greater than 40" risk factor as well as Bridges observation of a 10 year age difference between those CCATT patients with and without a cardiac diagnosis, confirming the continued need for dedicated predeployment risk assessment.\textsuperscript{8,12}

With this breadth of injury and illness, a standard paradigm for approaching the AE patient was requisite. The authors’ over-riding concern was the physiologic maintenance of tissue oxygen delivery to each anatomic system throughout transport no matter the altitude or condition. The goal was to

<table>
<thead>
<tr>
<th>Anatomic System*</th>
<th>Injury Source*</th>
<th>Battle Injury (n = 762) N (%)</th>
<th>Illness/Injury Source*</th>
<th>Disease, Non-Battle Injury (n = 626) N (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Head (n = 208)</td>
<td></td>
<td>126 (60.6)</td>
<td>82 (39.4)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IED</td>
<td>77 (61.1)</td>
<td>Seizure</td>
<td>20 (24.4)</td>
</tr>
<tr>
<td></td>
<td>GSW</td>
<td>20 (15.9)</td>
<td>Stroke</td>
<td>18 (22.0)</td>
</tr>
<tr>
<td></td>
<td>CCATT Assignment</td>
<td>65 (51.6)</td>
<td></td>
<td>20 (24.4)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spine (n = 112)</td>
<td></td>
<td>88 (78.6)</td>
<td>24 (21.4)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IED</td>
<td>59 (67.1)</td>
<td>MVA</td>
<td>4 (16.7)</td>
</tr>
<tr>
<td></td>
<td>GSW</td>
<td>12 (13.6)</td>
<td>Fall</td>
<td>4 (16.7)</td>
</tr>
<tr>
<td></td>
<td>CCATT Assignment</td>
<td>45 (51.1)</td>
<td></td>
<td>6 (25.0)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Face-Neck (n = 256)</td>
<td></td>
<td>232 (90.6)</td>
<td>24 (9.4)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IED</td>
<td>155 (66.8)</td>
<td>Fall</td>
<td>7 (29.2)</td>
</tr>
<tr>
<td></td>
<td>GSW</td>
<td>28 (12.1)</td>
<td>Fight</td>
<td>2 (8.3)</td>
</tr>
<tr>
<td></td>
<td>CCATT Assignment</td>
<td>95 (41.0)</td>
<td></td>
<td>2 (8.3)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extremity (n = 666)</td>
<td></td>
<td>560 (84.1)</td>
<td>106 (15.9)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IED</td>
<td>315 (56.3)</td>
<td>Fall</td>
<td>18 (17.0)</td>
</tr>
<tr>
<td></td>
<td>GSW</td>
<td>91 (16.3)</td>
<td>Infection</td>
<td>16 (15.1)</td>
</tr>
<tr>
<td></td>
<td>CCATT Assignment</td>
<td>156 (27.9)</td>
<td></td>
<td>9 (8.5)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chest (n = 176)</td>
<td></td>
<td>130 (73.9)</td>
<td>46 (26.1)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IED</td>
<td>59 (45.4)</td>
<td>Infection</td>
<td>14 (30.4)</td>
</tr>
<tr>
<td></td>
<td>GSW</td>
<td>32 (24.6)</td>
<td>Pneumothorax\textsuperscript{d}</td>
<td>5 (10.9)</td>
</tr>
<tr>
<td></td>
<td>CCATT Assignment</td>
<td>67 (51.5)</td>
<td></td>
<td>12 (26.1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abdomen (n = 176)</td>
<td></td>
<td>119 (67.6)</td>
<td>57 (32.4)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IED</td>
<td>58 (48.7)</td>
<td>Cholecystitis</td>
<td>12 (21.1)</td>
</tr>
<tr>
<td></td>
<td>GSW</td>
<td>29 (24.4)</td>
<td>Appendicitis</td>
<td>7 (12.3)</td>
</tr>
<tr>
<td></td>
<td>CCATT Assignment</td>
<td>70 (58.8)</td>
<td></td>
<td>6 (10.5)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pelvis (n = 88)</td>
<td></td>
<td>74 (84.1)</td>
<td>14 (15.9)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IED</td>
<td>36 (48.7)</td>
<td>Genitourinary</td>
<td>4 (28.6)</td>
</tr>
<tr>
<td></td>
<td>GSW</td>
<td>22 (29.7)</td>
<td>Fall</td>
<td>2 (14.3)</td>
</tr>
<tr>
<td></td>
<td>CCATT Assignment</td>
<td>31 (41.2)</td>
<td></td>
<td>3 (21.4)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Burns (n = 79)</td>
<td></td>
<td>65 (82.3)</td>
<td>14 (17.3)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IED</td>
<td>45 (69.2)</td>
<td>Electrical</td>
<td>2 (14.3)</td>
</tr>
<tr>
<td></td>
<td>RPG</td>
<td>3 (4.6)</td>
<td>JP8 Trash Fire</td>
<td>2 (14.3)</td>
</tr>
<tr>
<td></td>
<td>CCATT Assignment</td>
<td>41 (63.1)</td>
<td></td>
<td>6 (42.9)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cardiac (n = 196)</td>
<td></td>
<td>4 (2.0)</td>
<td>192 (98.0)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IED</td>
<td>3 (75.0)</td>
<td>CAS</td>
<td>178 (92.7)</td>
</tr>
<tr>
<td></td>
<td>RPG</td>
<td>1 (25.0)</td>
<td>Peri/Myocarditis</td>
<td>8 (4.2)</td>
</tr>
<tr>
<td></td>
<td>CCATT Assignment</td>
<td>4 (100.0)</td>
<td></td>
<td>75 (39.1)</td>
</tr>
</tbody>
</table>

MVA, motor vehicle accident; JP8, jet plane no. 8 fuel. *More than one system may be involved in any given patient. \textsuperscript{b}Two most commonly recorded etiologies for each anatomic system’s BI patients. \textsuperscript{c}Two most commonly recorded etiologies for each anatomic system’s DNBI patients. \textsuperscript{d}Spontaneous pneumothorax.
TABLE IV.  Worksheet Prescriptions Levied by the Theater Validating Flight Surgeon (N = 1,389)

<table>
<thead>
<tr>
<th>Preventive Actions (n = 975)</th>
<th>Battle Injury (n = 762)</th>
<th>Non-Battle Injury (n = 626)</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spine precautions</td>
<td>76 (7.8)</td>
<td>9 (0.9)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>C-Collar</td>
<td>50 (5.1)</td>
<td>11 (1.1)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Infectious Precautions</td>
<td>—</td>
<td>47 (4.8)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>N-95 Mask</td>
<td>—</td>
<td>14 (1.4)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Seizure Precautions</td>
<td>1 (0.1)</td>
<td>21 (2.2)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Neurovascular Checks</td>
<td>161 (16.5)</td>
<td>36 (3.7)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Medical Attendant</td>
<td>3 (0.3)</td>
<td>49 (5.0)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>CCATT Assignment</td>
<td>231 (23.7)</td>
<td>149 (15.3)</td>
<td>0.091</td>
</tr>
<tr>
<td>Wire-Cutters</td>
<td>16 (1.6)</td>
<td>3 (0.7)</td>
<td>0.018</td>
</tr>
<tr>
<td>Saline Lock</td>
<td>5 (0.5)</td>
<td>67 (6.9)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>No or Gentle</td>
<td>17 (1.7)</td>
<td>9 (0.9)</td>
<td>0.406</td>
</tr>
<tr>
<td>Total</td>
<td>560 (57.4)</td>
<td>415 (42.6)</td>
<td>0.417</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Positioning Actions (n = 1,176)</th>
<th>Battle Injury (n = 762)</th>
<th>Non-Battle Injury (n = 626)</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forward Loading</td>
<td>22 (2.9)</td>
<td>19 (3.1)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Head First Loading</td>
<td>61 (5.2)</td>
<td>15 (1.3)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Head of Bed Elevation</td>
<td>569 (48.1)</td>
<td>507 (42.9)</td>
<td>0.052</td>
</tr>
<tr>
<td>Extremity Elevation</td>
<td>2 (0.2)</td>
<td>0.00002</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>632 (53.7)</td>
<td>544 (46.3)</td>
<td>0.0002</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Medicinals (n = 750)</th>
<th>Battle Injury (n = 762)</th>
<th>Non-Battle Injury (n = 626)</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Supplemental Oxygen</td>
<td>242 (32.3)</td>
<td>68 (9.1)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Transfusion</td>
<td>79 (10.5)</td>
<td>8 (1.1)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Promethazine</td>
<td>134 (17.9)</td>
<td>64 (8.5)</td>
<td>0.493</td>
</tr>
<tr>
<td>Oxymetazoline</td>
<td>29 (3.9)</td>
<td>13 (1.7)</td>
<td>0.920</td>
</tr>
<tr>
<td>Enoxaparin</td>
<td>11 (1.5)</td>
<td>28 (3.7)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Morphine</td>
<td>14 (1.9)</td>
<td>23 (3.1)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Antibiotics</td>
<td>8 (1.1)</td>
<td>10 (2.5)</td>
<td>0.549</td>
</tr>
<tr>
<td>Other</td>
<td>5 (0.7)</td>
<td>19 (2.5)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Total</td>
<td>522 (69.6)</td>
<td>228 (30.4)</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Equipment Actions (n = 417)</th>
<th>Battle Injury (n = 762)</th>
<th>Non-Battle Injury (n = 626)</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chest Tube Suction</td>
<td>77 (18.5)</td>
<td>13 (3.1)</td>
<td>0.004</td>
</tr>
<tr>
<td>Nasogastric Tube Suction</td>
<td>142 (34.1)</td>
<td>27 (6.5)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Ostomy/Drain</td>
<td>64 (15.4)</td>
<td>6 (1.4)</td>
<td>&lt;0.0002</td>
</tr>
<tr>
<td>Maintenance</td>
<td>24 (5.8)</td>
<td>6 (1.4)</td>
<td>0.410</td>
</tr>
<tr>
<td>Overweight Litter</td>
<td>58 (13.9)</td>
<td>58 (13.9)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Total</td>
<td>307 (73.6)</td>
<td>110 (26.4)</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Aircraft Actions (n = 2530)</th>
<th>Battle Injury (n = 762)</th>
<th>Non-Battle Injury (n = 626)</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cabin Altitude Restriction</td>
<td>463 (18.3)</td>
<td>324 (12.8)</td>
<td>0.098</td>
</tr>
<tr>
<td>No “Remain Overnights”</td>
<td>608 (24.0)</td>
<td>494 (19.5)</td>
<td>0.273</td>
</tr>
<tr>
<td>No Stops</td>
<td>329 (13.0)</td>
<td>277 (11.0)</td>
<td>0.229</td>
</tr>
<tr>
<td>Long Landing</td>
<td>27 (1.2)</td>
<td>8 (0.3)</td>
<td>0.013</td>
</tr>
<tr>
<td>Total</td>
<td>1427 (56.4)</td>
<td>1103 (43.6)</td>
<td>0.003</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Miscellaneous Actions (n = 127)</th>
<th>Battle Injury (n = 762)</th>
<th>Non-Battle Injury (n = 626)</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>54 (42.5)</td>
<td>73 (57.5)</td>
<td>0.0002</td>
</tr>
</tbody>
</table>

χ² and Fisher’s exact tests were used when appropriate. *Denotes most prevalent prescription in a given category of intervention for BI or DNBI. †Denotes statistical significance.

The authors manipulated these three factors (cabin altitude, FiO₂, and Hgb) to optimize patient physiology. Supplemental oxygen was frequently used to maintain ground level equivalency (it was the most frequent Medicinal prescribed); transfusion was occasionally used to improve oxygen carrying capacity, and cabin altitude restriction was used to counter hypoxia, low Hgb level, and hypobaria. This approach explains the laboratory and physiologic parameters recorded in the worksheet. With these values a rough tissue oxygen delivery rate could be calculated. If more than the aforementioned 7.3 mL O₂/kg/min, the patient was considered physiologically optimized for flight.

In addition, other aircraft and patient interventions were prescribed. The TVFS frequently imposed on the aircraft no “remain overnights” and no stops during transit; this ensured that the patient arrived at the receiving medical facility using the shortest, most direct routing. Patient prescriptions fell into preventive, positioning, equipment, and medicinal actions. The most common and probably most effective Preventive Action strategy in these very ill patients was the assignment of a CCATT with its intensive care unit capability. Indeed, with BI the CCATT was exceedingly commonplace across all anatomic systems, ranging from a low of 28% with extremity injuries to a high of 100% with cardiac injuries. With DNBI, outside of cardiac illness, CCATT was seldom used. Other less prescribed preventive measures for BI included wire-cutters to prevent aspiration in the vomiting patient with wired jaws as well as spine precautions and C-collar to minimize vertebral injury motion; signifying actions. The most common and probably most effective Preventive Action strategy in these very ill patients was the assignment of a CCATT with its intensive care unit capability. Indeed, with BI the CCATT was exceedingly commonplace across all anatomic systems, ranging from a low of 28% with extremity injuries to a high of 100% with cardiac injuries. With DNBI, outside of cardiac illness, CCATT was seldom used. Other less prescribed preventive measures for BI included wire-cutters to prevent aspiration in the vomiting patient with wired jaws as well as spine precautions and C-collar to minimize vertebral injury motion; significant DNBI prescriptions included medical attendants to add an extra layer of care to a patient (particularly with cardiac patients), infectious precautions (e.g., airborne/contact precautions) and N-95 mask to limit spread of infection, and saline lock to ensure intravenous access should the need arise. Head of bed elevation proved to be the most common Positioning Action for both BI and DNBI. For BI, it was prescribed for comfort and pulmonary function enhancement while for DNBI comfort and cardiac function were favored reasons. The primary Equipment Action prescribed in BI was nasogastric suction and in DNBI cardiac monitor, reflecting the dominance of trauma and cardiac illness, respectively. Among the principal Medicinals prescribed in both BI and DNBI for nausea control and oxymetazoline
to facilitate middle ear equalization. Interestingly, there was significantly more use of enoxaparin and morphine in DNBI patients.

As previously noted, both the clearing flight surgeon and the TVFS evaluate each AE patient for flight. Beyond anecdotes, there has been no indication that any difference in approach existed. For the first time, using the worksheet, assignment of Precedence, Classification, CCATT, and FiO₂ as entered into the PMR (presumably by the clearing flight surgeon), could be compared to the final validated PMR (as determined by the TVFS). Except for FiO₂, there was good agreement between the two flight surgeons. With the FiO₂ agreement was not demonstrated. In fact, the FiO₂ prescribed by the TVFS was statistically different from the FiO₂ assignment by the clearing flight surgeon. Indeed, the TVFS raised the FiO₂ by 10 mmHg. This oxygen supplement was designed to ensure ground level equivalency and retain adequate tissue oxygen delivery in spite of the cabin’s ascent. Added education and/or preflight protocols/guidelines might well eliminate this prescribing difference.

The study had a number of limitations. First and foremost, it was not a systematic collection of data from a well-maintained database. Rather, it was an examination of an extraction of real-time data as used by the TVFS to make decisions. As such, inaccurate data and missing data can be found. This was most likely due to patient volume and time constraints imposed by the operational moment. As the worksheet was fully deidentified, a detailed retrospective quality check of the data could not be performed. In addition, there was a preponderance of Urgent and Priority patients. This was a function of the worksheet’s goal of streamlining the validation queue of high acuity patients. The worksheet came into being as the tempo of casualty evacuation rose and, as such, it was created in the exigency of the moment. As a result, Routine patients were underrepresented and approximately 3 months of early experience was missed. Consequently, the patients recorded in the worksheet probably are not representative of the entire deployment period (patient precedence was statistically different), rather they represent a cohort of high acuity patients encountered during a very busy time. And, lastly, the study concentrated on only two TVFSs and was a reflection of their experience alone; it cannot necessarily be generalized to all TVFSs. However, it is felt that despite these limitations, the worksheet provides a good snapshot of the two TVFSs’ validation activities and offers an initial glimpse into the medical practice of the TVFS.

CONCLUSION
This study is a first look at the clinical practice of the TVFS. It offers a snapshot of 1,389 patients seen by two successive TVFSs, their clinical variety and their clinical severity. In addition, it proffers an initial accounting of the many prescriptions that these TVFSs used to mitigate the physiologic stressors encountered during AE missions paying notice to the different prescriptive patterns associated with BI versus DNBI. Moreover, it tenders an early insight into the clearing versus validating flight surgeon approach to patient preparation for AE. And, lastly, the study provides a jumping point for future TVFS investigations using the available AE databases.

ACKNOWLEDGMENT
This project was supported by USAF School of Aerospace Medicine; Project No. 15-064.

REFERENCES
Navy En Route Care: A 3-Year Review of 428 Navy Air Evacuations

CDR Benjamin Walrath, MC USN*; Alejandra Mora, BS†; Victoria Ganem, RN, BSN‡; MAJ Stephen Harper, MC USA*; LCDR Elliot Ross, MC USN*; Col Chetan Kharod, USAF MC*; CDR Gerard Demers, MC USN§; Vikhyat S. Bebarta, MD¶

ABSTRACT Background: Navy medical personnel have been recording en route care (ERC) missions through Search and Rescue (SAR) reports since the 1970’s. Our objective was to report clinical ERC cases treated by Navy operational assets from January 2012 to January 2015. Methods: The Search and Rescue Model Manager office collects SAR reports for all patient transports involving Navy personnel and equipment. From these reports, descriptive statistics to include total number of patients transported, percentages of Advanced Life Support versus Basic Life Support transports, time of transport, and type of ERC provider for the transport were collected. Data reported as median (interquartile range) or percentages. Results: During a 3-year period, 428 patients were transported. Transport time was 54 (30–78) minutes. Missions were staffed by more than one provider 22% of the time. Individual providers included 76% Search and Rescue Medical Technicians, 25% Flight Surgeons, and 21% Other. Patients required ALS transport 54% of the time. Less than half (48%) of the patients were trauma related. Conclusion: In our review of 428 SAR reports from Navy ERC (2012–2015), we found that 76% of the missions were performed by Search and Rescue Medical Technicians and 54% met Advanced Life Support transport criteria.

INTRODUCTION

The U.S. Navy provides medical support to contingency and humanitarian operations across the globe, frequently requiring patient movement. En Route Care (ERC), as defined by the July 2013 Navy Tactics, Techniques, and Procedures (NTTP) Manual 4-02.2M and Marine Corps Reference Publication 4-11.1G on Patient Movement, is “transitory medical care, patient holding, and staging capabilities during transport of stabilized and/or stabilizing patients through successive medical care capabilities.” It further states “en route care requires appropriately trained medical attendants and the use of standardized, interchangeable, critical care equipment to ensure the evacuation system can successfully transport a patient to a higher capability of care.” ERC has not been designated a program of record in the Navy, thus there are no dedicated resources or requirements to man, train, and equip for the mission. Instead, ERC is usually conducted as an ad hoc mission with available equipment and personnel. The NTTP 3-50.1, Search and Rescue (SAR) Operations, governs all aspects of SAR and includes a framework to document patient transports and the ERC provided. According to NTTP 3-50.1, a rescue report is required in all cases where Navy assets are dispatched for a SAR or patient transport mission, to include humanitarian assistance, disaster relief, and medical evacuations. The rescue report includes both a section to describe the mission parameters as well as a section dedicated to the assessment and care of the transported patient. These SAR reports are collected by the SAR Model Manager (SARMM) for quality assurance purposes. Over 100 reports are filed annually and have been collected since the late 1960s. While this NTTP provides a standardized reporting mechanism for ERC missions, anecdotally many ERC providers outside of the SAR community are not familiar with the instruction. Alternatively, providers use standard message traffic to send situation reports, track patient status on locally generated patient care forms, or simply not document care in the prehospital setting. These methods of documentation are not captured by the SARMM.

Despite the consistent need for resources and assets for patient movement from the sea base, there is some debate whether patient transport is a requirement for the Navy. Common interpretation of current policies indicates intratheater patient transport is an Army mission (DUSTOFF), and intertheater transport is the responsibility of the Air Force (Strategic Aeromedical Evacuation). The confusion concerning the division of labor in ERC may be related to the task assignments described in Joint Publication 4-02 Health Service Support. In Appendix B, although it states each service
component is responsible for patient evacuation, it also specifically assigns the U.S. Army responsibility for ship-to-shore and shore-to-ship patient transports. This does not appear to address the need to transport patients from ship-to-ship, either in a “lily pad” pattern until ultimate delivery to a shore-based medical treatment facility or to receive higher level of medical care from a larger ship with more robust medical resources. The delineation of responsibility for patient transport is further emphasized in Department of Defense (DoD) Directive 5100.01 from December 2010, where the U.S. Army is specifically tasked with providing intratheater aeromedical evacuation and the U.S. Air Force is responsible for global mobility to include aeromedical evacuation. The U.S. Navy is tasked with establishing, maintaining and defending sea bases in support of naval, amphibious, land, air, or other joint operations without specific reference to patient movement. DoD Instruction 6000.11: Patient Movement was updated May 4, 2012 and designated the Commander of United States Transportation Command as the single manager for Patient Movement across the DoD, both intertheater and intratheater, and further tasks the Secretaries of the Military Departments with providing the necessary equipment and personnel to comply with quality care standards as established by Commander of United States Transportation Command. A systems-based approach for patient transport to and from the sea base and across the maritime domain has yet to be standardized. According to the VANGUARD study of 2005, “CASEVAC/MEDEVAC/ERC for each phase of an operation, to include the sea base,” represents a Navy doctrinal capability shortfall. The current operational structure, with patient movement as a component of the larger SAR mission set, does not appear to be adequate. The ERC system utilized in Iraq and Afghanistan was supported by an Army program of record for patient movement in what became a mature operating environment. Despite this, Eastridge et al. document an 87.3% injury mortality in the pre-MTF environment. The Navy rarely operates in a “mature” environment, and lacks the requirement-driven manning, training, and equipment decisions to support ERC that would otherwise accompany a “program of record” designation. Similar to the Mueller et al. recently published article about EMS in the frontier and remote areas in the continental United States, we hope our review of the SAR reports may provide an historical reference of Navy ERC for future operations and logistics planning, guide further research efforts, as well as inform the debate about whether the Navy should officially designate ERC as a program of record. Our objective was to report clinical ERC cases transported by Navy operational assets from January 2012 to January 2015.

METHODS

The study protocol was approved by the Brooke Army Medical Center Institutional Review Board in compliance with all applicable federal regulations governing the protection of human subjects.

Following the authorization of the Commanding Officer of Helicopter Sea Combat Squadron THREE (HSC-3), the de-identified SAR report database capturing SAR reports from January 2012 to January 2015 were obtained through the SARMM office. The SARMM collects SAR reports for all patient transports at sea and from Navy SAR stations in accordance with NTTP 3-50.1. Of note, OPNAV Instruction 3130.6E designates HSC-3 as the SARMM with the responsibility for “promoting policy and standardization in SAR training, equipment, manuals and procedures for all rescue capable units and associated medical personnel.”

Data provided by SARMM per patient included time of transport, ERC provider type, identification of trauma versus medical patients, patient demographics, ALS versus BLS categorization, and interventions performed. ERC provider types included Search and Rescue Medical Technician (SMT), Flight Surgeon (FS), and Other. The Other provider category referred to undesignated corpsmen (non-SMT) not specified in the database. Our study team compiled the summation of total number of patient transports, performed statistical analysis as described below, created an additional provider category of Team defined as a mission having 2 or more providers during transport, and verified ALS versus BLS categorization using an extrapolation of the Centers for Medicare and Medicaid Services definition for ALS care. Specifically, an ALS transport was defined as “the transportation by ground ambulance vehicle and the provision of medically necessary supplies and services including the provision of an ALS assessment or at least one ALS intervention.” Any transport that did not meet ALS criteria was considered BLS. Patient outcomes were not available and, thus, we did not include in the analysis.

We used JMP version 10 (SAS Institute, Inc., Cary, NC) for the statistical analysis. We compared the incidence with \( \chi^2 \) or Fisher Exact test where appropriate. Categorical data were reported as frequencies and percentages. Following a Shapiro–Wilks test for normality, Kruskal–Wallis test was used for nonparametric continuous variables. Continuous variables were reported as medians (interquartile range [IQR]). A \( p < 0.05 \) was considered significant.

RESULTS

For the study interval between January 2012 and January 2015, 428 patient transports by Navy assets were recorded using SAR reports and collected by SARMM. Transport time was 54 (30–78) minutes. Point of origin, destination, and classification of mission type was not captured in this data set. Missions were staffed by an individual provider 78% of the time (61% SMT, 6% FS, and 11% Other), and 22% of the missions were staffed by a Team. Of the 22% (\( n = 93 \)) of missions staffed by a team, 52 were SMT/FS, 13 SMT/Other, 27 FS/Other, and one had SMT/FS/Other. Thus, an SMT was present on a total of 76% of the missions, a FS for 25%, and a non-SMT corpsman 21%. A total of 14 patient transports did not have a documented provider
type. No nurses were documented as the ERC provider for the transports in our data set.

In our data set, patients were 75% male, 30 (IQR 22–50) years of age, 47% active duty military, and 48% trauma injuries (Table I). SMTs transported older patients in comparison to team, FS, and other. In a comparison across provider types, SMTs were more likely to transport civilian patients ($p < 0.0001$) and FS were more likely to transport military patients ($p < 0.0001$). There was no difference in percentage of trauma and medical patients transported across provider types.

Over half (54%) of the missions were considered ALS. Only 13% of patients had an advanced airway device. Oxygen support was provided during 40% of the missions. Half or less of the missions involved providing patients with intravenous access (47%), administration of medications (50%), and cardiac monitoring (49%). SMTs were more likely to staff ALS transports than all other providers (59% vs 40%, $p = 0.0004$). FS were not associated with transport level or patient type. Further, ALS transports were more likely to be medical (ill, not injured) patients than trauma patients (61% vs 46%, $p = 0.002$). Missions with medical patients compared to trauma patients were more likely to involve oxygen support (47% vs 34%, $p = 0.006$), intravenous access (56% vs 39%, $p = 0.001$), medication administration (36% vs 27%, $p = 0.03$) and cardiac monitoring (57% vs 41%, $P = 0.001$). Mission type and procedures performed are summarized in Table II.

**DISCUSSION**

In our review of 3 years of Navy SAR reports, we found 428 documented patient transport missions, over half (54%) of which required ALS care. To the best of our knowledge, this is the first attempt to quantify and describe the scope of the Navy ERC mission. This description is an important tool to inform operational decisions and resource allocation, as a guide for future research initiatives, and to facilitate an evidence-based decision on the future of ERC as an official program of record.

ERC in the Navy is an ad hoc mission, without dedicated and standardized assets. One of the most critical ERC resources is the health care provider, and the Navy has several different types of provider performing in the capacity as caregiver. OPNAVINST 3130.6E specifically designates the SMT for this role, stating "the mission of the naval SMT is to perform aircrew and advanced life support (ALS) emergency medical care functions independent of a medical officer during SAR, combat search and rescue (CSAR), air ambulance, casualty evacuation (CASEVAC), en-route care and/or MEDEVAC missions from rotary and fixed wing aircraft. The SMT shall be a volunteer and be physically conditioned to routinely perform demanding rescues in all operational environments. The SMT’s expertise shall provide fleet commanders with the ability to assist DoD and civilian personnel in distress."

In August of 2015, a letter from the Commander, Naval Air Force, Pacific to the Surgeon General of the Navy requested a requirement for SMTs to be paramedic level trained to support SAR missions (personal communication, 3 August 2015). This request followed a congressional mandate to the U.S. Army to educate, train, and certify their 68W flight medics as critical-care flight paramedics. Although Sanghavi et al. suggest ALS care may result in worse outcomes for prehospital patients for all conditions other than acute

### Table I. Comparison of Patient Characteristics by Provider Type

<table>
<thead>
<tr>
<th>Provider</th>
<th>Overall</th>
<th>Team</th>
<th>FS</th>
<th>SMT</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (Years)</td>
<td>30 (22–50)</td>
<td>25 (21–35)</td>
<td>24 (23–37)</td>
<td>39 (22–58)</td>
<td>30 (21–37)</td>
</tr>
<tr>
<td>Gender (% Male)</td>
<td>75%</td>
<td>75%</td>
<td>100%</td>
<td>73%</td>
<td>83%</td>
</tr>
<tr>
<td>Civilian</td>
<td>52%</td>
<td>47%</td>
<td>4%</td>
<td>66%</td>
<td>19%</td>
</tr>
<tr>
<td>Military</td>
<td>48%</td>
<td>53%</td>
<td>96%</td>
<td>34%</td>
<td>81%</td>
</tr>
<tr>
<td>Trauma</td>
<td>48%</td>
<td>53%</td>
<td>48%</td>
<td>49%</td>
<td>36%</td>
</tr>
<tr>
<td>Medical</td>
<td>52%</td>
<td>47%</td>
<td>52%</td>
<td>51%</td>
<td>64%</td>
</tr>
</tbody>
</table>

**Table II.** Type of Transport and Interventions Performed by Provider Type

<table>
<thead>
<tr>
<th>Provider</th>
<th>Overall</th>
<th>Team</th>
<th>FS</th>
<th>SMT</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>BLS</td>
<td>46%</td>
<td>34%</td>
<td>40%</td>
<td>44%</td>
<td>59%</td>
</tr>
<tr>
<td>ALS</td>
<td>54%</td>
<td>60%</td>
<td>60%</td>
<td>54%</td>
<td>41%</td>
</tr>
<tr>
<td>Ambulatory</td>
<td>40%</td>
<td>39%</td>
<td>48%</td>
<td>28%</td>
<td>46%</td>
</tr>
<tr>
<td>Airway</td>
<td>13%</td>
<td>15%</td>
<td>16%</td>
<td>13%</td>
<td>0%</td>
</tr>
<tr>
<td>Oxygen</td>
<td>40%</td>
<td>41%</td>
<td>40%</td>
<td>44%</td>
<td>25%</td>
</tr>
<tr>
<td>IV</td>
<td>47%</td>
<td>54%</td>
<td>40%</td>
<td>48%</td>
<td>36%</td>
</tr>
<tr>
<td>Medication</td>
<td>50%</td>
<td>27%</td>
<td>48%</td>
<td>32%</td>
<td>27%</td>
</tr>
<tr>
<td>Monitor</td>
<td>49%</td>
<td>48%</td>
<td>40%</td>
<td>54%</td>
<td>30%</td>
</tr>
</tbody>
</table>

Team, two or more providers; NS, not statistically significant. *14 missions did not have a documented provider type. **$p$ value is reflective of four-way comparison across provider types; Age comparison used Kruskal–Wallis test following Shapiro–Wilks test for normality ($p < 0.0001$); Gender, Civilian, and Military status comparison used Fisher Exact test; Trauma and Medical categories comparison used chi$^2$ test.
myocardial infarction, we find the results of Mabry et al. in 2012 more compelling and thus assert our data analysis supports this requirement. Previous studies on ERC providers and interventions in a combat setting have been performed, however this is the first report specific to Navy ERC. ALS care was required in 54% of ERC transports. Over half (52%) of the patients had a primary medical (ill, not injured), not trauma, indication for transport. Civilians comprised 53% of the patients, which arguably reinforces the requirement to meet the civilian practice requirements as set forth in the National EMS Scope of Practice Model. This would expand the requirement for Navy ERC providers to include not only paramedic-level education and training, but also paramedic-level certification, licensure, and credentialing prior to conducting ALS transports.

Based on our review of the SAR reports, FS also perform this function, although according to the Navy Medicine Operational Training Center website, “The naval FS practices preventive medicine first and foremost.” A FS has completed medical school and at least a 1-year internship, before completing a 7-month training curriculum at the Naval Aerospace Medicine Institute. During that 7-month FS course, approximately 4 hours of didactic education is devoted to ERC. (H. Casey, personal communication, 22 July 2015). Other potential providers performing ERC in the Navy include nurses, designated corpsmen, other physicians, or any combination of the above. Navy nurses, SMTs, and physicians may have the opportunity to attend the Joint ERC Course taught at the Army School of Aviation Medicine in Fort Rucker, AL usually as just-in-time training before deployment.

In our study, the SMTs participated in 76% of the ERC missions, FS participated in 25%, non-SMT corpsmen participated in 21% and 22% were conducted by a team of providers. The role of Navy ERC nurses is not clearly defined with this data set. While SMTs are specifically trained to perform the ERC mission, FS, and other corpsmen are usually not. Thus, in the 24% of ERC missions where an SMT did not participate, the health care provider attending the patient had most likely not received Navy training specific for the mission he or she was assigned. Improving the ERC education and training of the FS and non-SMT corpsmen, restricting the use of FS and non-SMT corpsmen as ERC providers, or increasing the availability of SMTs or other trained ERC providers could mitigate this mismatch.

Future research should link patient transport details with patient outcome data to further characterize the provider attributes that can be modified to optimize patient outcomes during the ERC mission. We should also seek to identify the platforms and equipment used for ERC, and attempt to determine if choice of platform and availability of certain equipment and supplies facilitates the provision of ERC and improves patient outcome. A formal database with prospectively collected data routinely verified for data accuracy and completeness would be beneficial in this effort. The DoD Trauma Registry, Joint Trauma System, and the many evidence based recommendations produced through scholarly review of diligently collected clinical data should serve as a template for future work.

Our study has several limitations. It was a retrospective review of a database of clinical records, which may not be complete. We did not have access to the original SAR reports, and the data provided was already abstracted and de-identified. We were unable to clarify certain data elements, including: if procedures documented were performed by the ERC provider or were completed before transport and maintained en route, point of origin or transport destination, and specific training and certifications of individual providers. Patients may have been transported without the completion of a SAR report and therefore would not be included in our data set. Anecdotal reports of other providers (nurses, physicians, other corpsmen) transporting patients and not using the SAR rescue report were not quantified, but indicate the need for a more comprehensive search of medical and operational records to conduct a complete analysis of Navy ERC. Incomplete mission capture could skew the percentages and descriptive statistics reported, but would also indicate a greater total number of ERC missions. However, a larger sample size of patients may have produced different results. Patient outcomes were not available for analysis. Due to the rapidly changing environment of modern naval warfare, our findings may not be generalizable to the Navy ERC mission profile in the future. They may also not be applicable to U.S. Army or U.S. Air Force ERC missions.

CONCLUSION

In our study of Navy ERC, we found 428 documented patient transports in a 3 year period. A majority of these required ALS care. Based on the number of patients transported and the complexity of the care required during transport, we recommend Navy leaders review the historical ERC mission profiles as they consider the value of establishing ERC as a program of record.

ACKNOWLEDGMENTS

We would like to extend our gratitude to the many Navy ERC providers who completed SAR reports and to the SARMN personnel who created and populated the database, specifically HMC Richard Bestwick, HM1 Tom Walsh, and HM1 Richmond Roy. We appreciate Maj (Dr.) Joseph Maddry, Jacob Minnick, and the Air Force En Route Care Research Center staff, for their support in the analysis of the provided data. Finally, we thank LTC (Dr.) Robert Mabry and the Military EMS and Disaster Medicine fellowship staff for their support on this study.

REFERENCES

Use of the Pain Assessment Screening Tool and Outcomes Registry in an Army Interdisciplinary Pain Management Center, Lessons Learned and Future Implications of a 10-Month Beta Test

COL Diane M. Flynn, MC USA (Ret.)*; Karen Cook, PhD†; Michael Kallen, PhD†; COL Chester Buckenmaier, MC USA (Ret.); COL Ricke Weickum, MS USA (Ret.)*; Teresa Collins, CCRC‡; Ashley Johnson, AA‡; Dawn Morgan, RN*; COL Kevin Galloway, AN USA (Ret.); Kristin Joltes, PA-C, CCRP‡

ABSTRACT  Introduction: The U.S. Army Comprehensive Pain Management Campaign Plan was launched in 2010 to improve pain outcomes in military populations. Interdisciplinary Pain Management Centers (IPMCs) were established at every Army medical center, each offering a robust array of treatment options including conventional and complementary and integrative medicine (CIM) pain management therapies. The Pain Assessment Screening Tool and Outcomes Registry (PASTOR) was developed to assess and track biopsychosocial aspects of pain management and to identify best treatment practices. Methods: During a 10-month pilot test of PASTOR at one Army IPMC, active duty patients completed PASTOR at baseline and at significant junctures during their therapeutic course. Results: 322 IPMC patients completed baseline and follow-up PASTOR assessments. The PASTOR outcomes were analyzed for patients who completed a 3- to 6-week CIM program, a 3-week functional restoration program, or both. For most PASTOR domains, a greater proportion of patients who completed both programs reported important improvement compared with patients who completed either program alone. Conclusions: This pilot test demonstrated the utility of using PASTOR in a military IPMC to track biopsychosocial treatment outcomes. These preliminary data will inform future comparative effectiveness analyses of pain therapies among military and veteran populations.

BACKGROUND
Pain is a leading cause of disability among active duty service members and veterans. There is currently no Department of Defense (DoD) or Veterans Health Administration (VHA) screening tool and outcomes repository to promote consistency in pain care. To address these concerns, in 2009, the Army Surgeon General chartered the Pain Management Task Force (PMTF) to evaluate the state of pain management in the military and to determine how best to optimize the quality of life for those with pain. The resulting 2010 PMTF report delineated 109 recommendations, which were incorporated into the 2010 Army Comprehensive Pain Management Campaign Plan (CPMCP). Among the goals of the CPMCP were to 1) establish an Interdisciplinary Pain Management Center (IPMC) at each Army Medical Center; 2) develop an electronic Pain Assessment Screening Tool and Outcomes Registry (PASTOR) in order to propagate evidenced-based best practices; and 3) implement and research nonpharmacological approaches to pain management using complementary and integrative medicine (CIM) approaches. In 2013, the Office of the Army Surgeon General reaffirmed its commitment by including the improvement of pain management as one of the priorities of the Army Medicine 2020 Campaign Plan.

As a result of Army CPMCP, each Army medical center established an IPMC. Each IPMC offers a broad panel of evidence-based CIM pain therapies, including chiropractic, acupuncture, massage, yoga as well as conventional pain therapies, such as medication management, interventional pain care, and physical, occupational, and psychological therapies. The Defense and Veterans Center for Integrative Pain Management was established to coordinate pain management efforts and research across the military and veteran health systems, as called for in the PMTF report. The Defense and Veterans Center for Integrative Pain Management developed the Defense and Veterans Pain Rating Scale (DVPRS) (Fig. 1) and has validated it in military personnel and veterans in inpatient and outpatient settings. The DVPRS was imbedded into PASTOR, which was developed between 2012 and 2014. Beginning in 2014, a PASTOR beta test was conducted at three DoD sites including the Madigan Army Medical Center IPMC, the Walter Reed National Military Medical Center Warrior Transition Clinic, and the Naval Medical Center San Diego Pain Clinic.

The purpose of this study was to analyze descriptive data collected at the Madigan site during the PASTOR beta test in order to provide a foundation for future comparative effectiveness analyses of pain therapies in military populations. Because of logistical challenges of analyzing data collected

*Department of Pain Management, Madigan Army Medical Center, 9040 Jackson Avenue, Tacoma, WA 98431.
†Department of Medical Social Sciences, Feinstein School of Medicine, Northwestern University, 633 North Street Clair, 19th Floor, Chicago, IL 60611.
‡Defense and Veterans Center for Integrative Pain Management, 11308 Rockville Pike, Rockville, MD 20852.

doi: 10.7205/MILMED-D-16-00212
from all three beta test sites, each with distinct clinical operations and patient populations, the current analysis was limited to data collected at the Madigan site.

METHODS
A 10-month PASTOR beta test was initiated in May 2014. During this period, all active duty patients referred to the Madigan Army Medical Center IPMC completed a baseline PASTOR questionnaire at the time of their initial clinic visit. A pain physician then performed a comprehensive evaluation to develop a therapeutic plan. Depending on each patient’s clinical history, the therapeutic plan may have included medication management, interventional procedure(s), and/or a program of interdisciplinary care. Patients were instructed to

FIGURE 1. Defense and veterans pain rating scale.
complete follow-up PASTOR questionnaires at approximately monthly intervals during their therapeutic course.

The Madigan IPMC offers two interdisciplinary treatment programs. The first, the Integrative Modalities Pain Care Team (IMPACT) program, is composed of a 3- to 6-week course of twice weekly physical therapy, occupational therapy, chiropractic treatment, and acupuncture; once weekly psychoeducational classes; once weekly didactic pain management education; and once weekly instruction on myofascial release techniques. Following the completion of the current study, once weekly yoga was added to the IMPACT program but yoga therapy was not included in the current analysis. The IMPACT program totaled approximately 28 hours for the 3-week program, and 56 hours for the 6-week program (Fig. 2).

The other Madigan IPMC interdisciplinary pain treatment program is the intensive outpatient functional restoration program (FRP). The FRP is a 3-week program that includes 4 full days of therapy per week. The first 2 and last 2 hours of each day of therapy occur in a gymnasium working closely with physical therapists, an occupational therapist, and their assistants. Between gym sessions, participants engage in a 1-hour health psychology group and a 1-hour didactic session. The FRP totaled approximately 72 treatment hours over 3 weeks (Fig. 2).

The patient selection for the two interdisciplinary treatment programs was not random. The FRP is recommended for patients who meet baseline fitness entry criteria, which includes the ability to stand up and sit down on the floor independently, to walk on a treadmill for at least 10 minutes at a pace of at least 2.5 miles per hour, to lift 25 pounds from floor to knuckle height and from knuckle to shoulder height, and to carry at least 25 pounds a distance of 40 feet. The IMPACT program is typically recommended for patients who do not meet the FRP entry criteria. Therefore, at baseline, IMPACT patients are generally less fit than the FRP patients. Participants of the IMPACT program who show substantial improvement are often recommended for FRP following IMPACT completion.

**PASTOR**

PASTOR is a web-based survey of patient reported measures adapted for use by the Military Health System and VHA from PROMIS, an outcomes tool developed by the National Institutes of Health (NIH). PASTOR incorporates the DVPRS, pain interference assessment, neuropathic pain scale, headache assessment, and patient-defined activity goals. In addition,

<table>
<thead>
<tr>
<th>TABLE I. Measures Included in Pain Assessment Screening Tool and Outcomes Registry</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pain Measures</strong></td>
</tr>
<tr>
<td>- Pain intensity, Quality, and Impact</td>
</tr>
<tr>
<td>- Defense and Veterans Pain Rating Scale</td>
</tr>
<tr>
<td>- Graphic Pain Map</td>
</tr>
<tr>
<td>- PROMIS Pain Interference (Computer Adaptive)</td>
</tr>
<tr>
<td>- PROMIS Neuropathic Pain Scale (Short Form)</td>
</tr>
<tr>
<td>- PROMIS Headache (Computer Adaptive)</td>
</tr>
<tr>
<td>- What Are Your 3 Most Important Activities That Are Limited By Pain?</td>
</tr>
<tr>
<td>- How Well Are You Currently Able to Perform (Each) Activity?</td>
</tr>
<tr>
<td><strong>Problem Screening</strong></td>
</tr>
<tr>
<td>- Post-traumatic stress disorder</td>
</tr>
<tr>
<td>- PROMIS Anxiety (Computer Adaptive)</td>
</tr>
<tr>
<td>- PROMIS Depression (Computer Adaptive)</td>
</tr>
<tr>
<td>- PROMIS Alcohol (Computer Adaptive)</td>
</tr>
<tr>
<td><strong>PROMIS Measures of Additional Pain Correlates</strong></td>
</tr>
<tr>
<td>- PROMIS SF v1.0: Global Health (Short Form)</td>
</tr>
<tr>
<td>- PROMIS Fatigue (Computer Adaptive)</td>
</tr>
<tr>
<td>- PROMIS Satisfaction With Social Roles (Computer Adaptive)</td>
</tr>
<tr>
<td>- PROMIS Anger (Computer Adaptive)</td>
</tr>
<tr>
<td>- PROMIS Sleep-Related Impairment (Computer Adaptive)</td>
</tr>
<tr>
<td>- PROMIS Bank v1.0: Physical Function</td>
</tr>
<tr>
<td><strong>Health Utilization</strong></td>
</tr>
<tr>
<td>- Providers Seen-By Type</td>
</tr>
<tr>
<td>- Self-Reported Treatment History and Effectiveness</td>
</tr>
</tbody>
</table>

**FIGURE 2.** The Integrative Modalities Pain Care Team (IMPACT) program and Functional Restoration Program (FRP) weekly schedules.
PASTOR includes screens for the following conditions: post-traumatic stress disorder (PTSD), anxiety, depression, and alcohol misuse. Additional pain correlates including global health, fatigue, satisfaction with social roles, anger, sleep-related impairment, and physical function are also assessed (Table I). PASTOR incorporates computer adaptive testing (CAT) in which a computer algorithm selects items asked of each respondent based on their previous responses. The CAT methodology results in higher measurement precision with a lower response burden than
questionnaires without CAT capability. The PROMIS scores within PASTOR are standardized to the T-score metric with a mean score of 50 and standard deviation of 10, which allows comparison with a reference population.

After patients complete each PASTOR assessment, a four-page provider report is generated which shows the location(s) of pain, clinical alerts for positive screens for PTSD, depression, anxiety and alcohol misuse/abuse, and trends for other PASTOR domains. The PASTOR report also shows normative data of the patients’ reports compared with the U.S. population (Fig. 3).

In addition to completing PASTOR, patients who engaged in the IMPACT and FRP programs also underwent objective assessment of baseline and postprogram functional measures. These functional measures were measured by a physical therapist or occupational therapist and included the following: how much weight in pounds could the participant lift from floor to the level of the knuckles when standing, how much weight could be lifted from knuckle-to-shoulder height, how much weight could be carried a distance of 40 feet, and how many metabolic units could the participant generate on a treadmill using the modified Naughton treadmill protocol.17

**Statistical Methods**

In order to determine patient response to therapy, the minimal clinically important difference (MID) was established for each PASTOR domain. In a population of patients with advanced cancer, Yost et al.18 established MIDs for five of the PROMIS domains used in PASTOR including pain interference, physical functioning, anxiety, depression, and fatigue. For PASTOR domains for which MID has not been empirically determined, coauthors with extensive PROMIS psychometrics experience recommended that the MID be estimated as one-half the standard deviation. To determine the MID for objective functional measures, we consulted with a senior biostatistician who advised that we convert the functional outcomes to the T-score metric with a T-score of 50 and standard deviation. To determine the MID for objective functional measures, we consulted with a senior biostatistician who advised that we convert the functional outcomes to the T-score metric with a mean score of 50 and standard deviation of 10, which allows comparison with a reference population.

Despite the consensus recommendation to use responder analyses for clinical trials of pain therapies, a 2014 review of 162 randomized clinical trials of interventions for chronic pain outcomes research was published by the Initiative on Methods, Measurement, and Pain assessment in Clinical Trials.19 This group was comprised of 40 stakeholders with research, clinical, or administrative expertise in pain outcomes from universities, governmental agencies, a patient self-help organization, and the pharmaceutical industry. The consensus of this group was that pain outcomes research should include assessment of four domains: 1) pain intensity, 2) physical functioning, 3) emotional functioning, and 4) overall improvement. In addition, they emphasized the importance of determining the minimally important change for each outcome measure then measuring the proportion of subjects who respond beyond this threshold. This approach to pain outcomes research is termed “responder analyses” and ensures that outcomes reported consider clinical significance rather than statistical significance alone.

Despite the consensus recommendation to use responder analyses for clinical trials of pain therapies, a 2014 review of 162 randomized clinical trials of interventions for chronic pain outcomes research was published by the Initiative on Methods, Measurement, and Pain assessment in Clinical Trials.19 This group was comprised of 40 stakeholders with research, clinical, or administrative expertise in pain outcomes from universities, governmental agencies, a patient self-help organization, and the pharmaceutical industry. The consensus of this group was that pain outcomes research should include assessment of four domains: 1) pain intensity, 2) physical functioning, 3) emotional functioning, and 4) overall improvement. In addition, they emphasized the importance of determining the minimally important change for each outcome measure then measuring the proportion of subjects who respond beyond this threshold. This approach to pain outcomes research is termed “responder analyses” and ensures that outcomes reported consider clinical significance rather than statistical significance alone.

Despite the consensus recommendation to use responder analyses for clinical trials of pain therapies, a 2014 review of 162 randomized clinical trials of interventions for chronic pain outcomes research was published by the Initiative on Methods, Measurement, and Pain assessment in Clinical Trials.19 This group was comprised of 40 stakeholders with research, clinical, or administrative expertise in pain outcomes from universities, governmental agencies, a patient self-help organization, and the pharmaceutical industry. The consensus of this group was that pain outcomes research should include assessment of four domains: 1) pain intensity, 2) physical functioning, 3) emotional functioning, and 4) overall improvement. In addition, they emphasized the importance of determining the minimally important change for each outcome measure then measuring the proportion of subjects who respond beyond this threshold. This approach to pain outcomes research is termed “responder analyses” and ensures that outcomes reported consider clinical significance rather than statistical significance alone.

**Outcomes Analysis**

Published guidelines on pain management research were considered in our study design. In 2008, a consensus statement on chronic pain outcomes research was published by the Initiative on Methods, Measurement, and Pain assessment in Clinical Trials.19 This group was comprised of 40 stakeholders with research, clinical, or administrative expertise in pain outcomes from universities, governmental agencies, a patient self-help organization, and the pharmaceutical industry. The consensus of this group was that pain outcomes research should include assessment of four domains: 1) pain intensity, 2) physical functioning, 3) emotional functioning, and 4) overall improvement. In addition, they emphasized the importance of determining the minimally important change for each outcome measure then measuring the proportion of subjects who respond beyond this threshold. This approach to pain outcomes research is termed “responder analyses” and ensures that outcomes reported consider clinical significance rather than statistical significance alone.

Despite the consensus recommendation to use responder analyses for clinical trials of pain therapies, a 2014 review of 162 randomized clinical trials of interventions for chronic pain outcomes research was published by the Initiative on Methods, Measurement, and Pain assessment in Clinical Trials.19 This group was comprised of 40 stakeholders with research, clinical, or administrative expertise in pain outcomes from universities, governmental agencies, a patient self-help organization, and the pharmaceutical industry. The consensus of this group was that pain outcomes research should include assessment of four domains: 1) pain intensity, 2) physical functioning, 3) emotional functioning, and 4) overall improvement. In addition, they emphasized the importance of determining the minimally important change for each outcome measure then measuring the proportion of subjects who respond beyond this threshold. This approach to pain outcomes research is termed “responder analyses” and ensures that outcomes reported consider clinical significance rather than statistical significance alone.

**TABLE II.** Minimally Important Difference (MID) Key for Pain Assessment Screening Tool and Outcomes Registry (PASTOR) Measures. PASTOR is a Web-Based Survey of Patient-Reported Measures Adapted for Use By the Military Health System and Veterans Health Administration from PROMIS, an Outcomes Tool Developed by the National Institutes of Health

<table>
<thead>
<tr>
<th>Variable</th>
<th>MID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pain Intensity (DVPRS)</td>
<td>2.5 (0–10 Scale)</td>
</tr>
<tr>
<td>PROMIS Anxiety</td>
<td>T Score Change of 3.75</td>
</tr>
<tr>
<td>PROMIS Depression</td>
<td>T Score Change of 3.75</td>
</tr>
<tr>
<td>PROMIS Fatigue</td>
<td>T Score Change of 4</td>
</tr>
<tr>
<td>PROMIS Anger</td>
<td>T Score Change of 5</td>
</tr>
<tr>
<td>PROMIS Pain Interference</td>
<td>T Score Change of 5</td>
</tr>
<tr>
<td>PROMIS Sleep Impairment</td>
<td>T Score Change of 5</td>
</tr>
<tr>
<td>PROMIS Global Health</td>
<td>T Score Change of 1</td>
</tr>
<tr>
<td>PROMIS Physical Function</td>
<td>T Score Change of 5</td>
</tr>
<tr>
<td>PROMIS Satisfaction with Social Role</td>
<td>T Score Change of 5</td>
</tr>
<tr>
<td>PROMIS Alcohol Use</td>
<td>T Score Change of 5</td>
</tr>
</tbody>
</table>

**FIGURE 4.** Magnitude of change in selected Pain Assessment Screening Tool and Outcomes Registry measures following care in the Interdisciplinary Pain Management Clinic (n = 343). MID, minimal clinically important difference.
low back pain revealed that only 17% of studies included a justification for the definition of responder. Among those studies, the definition of responder varied widely.20 The authors called for a standardized approach to defining response to therapy based on empirically derived MIDs.

More recently, the 2015 report of the NIH Taskforce on Research Standards for Chronic Low Back Pain emphasized the importance of measuring impact of chronic low back pain in terms of pain intensity, pain interference and physical function, and concurred with previous recommendations to use established MIDs to conduct responder analyses.21

RESULTS

During the first 10 months of PASTOR use, 646 IPMC patients completed initial PASTOR assessment and 343 (53%) IPMC patients completed at least one follow-up assessment. This group was composed of patients who received any form of treatment in the IPMC, including medication management and/or interventional therapies, in addition to those who completed the IMPACT program or FRP. The interval between first and last PASTOR assessment ranged from 20 to 287 days. Figure 4 shows the PASTOR outcomes for this group.

To interpret Figures 4 through 7, it should be noted that each bar represents 100% of subjects who completed baseline and follow-up assessment of each domain listed on the x-axis. The legend shows the portion of each bar corresponding to different magnitudes of change between baseline and follow-up. The shaded portion(s) at the bottom of each bar corresponds to proportion of subjects who reported at least the minimal clinically important improvement or two times the minimal clinically important improvement. The central unshaded portion of each bar corresponds to the proportion of subjects who did not experience clinically important improvement or worsening. The shaded portion(s) at the top of each bar corresponds to the proportion of subjects who reported at least the minimally clinically important worsening or 2 times the minimally important worsening. Using anxiety as an example, Table II shows that a T score change of 3.75 has been determined to be the MID. The bottom two portions of the bar corresponding to anxiety in Figure 2 indicates that among the 343 subjects who completed baseline and follow-up assessments of anxiety, 36% reported improvement corresponding to a T score change of at least 3.75 and 18% reported at least twice that magnitude of improvement (7.5 or greater change in T score). The central portion of the bar indicates that 42% of the 343 subjects reported no clinically important change in anxiety. The top two portions of the bar indicate that 25% of subjects reported a worsening of anxiety corresponding to at least 3.75 change in T score and 13% reported at least twice that magnitude of worsening (7.5 or greater change in T score).

Of the 343 patients who completed baseline and follow-up PASTOR assessments, 50 completed the IMPACT program and/or FRP (Table III). The remainder of this analysis will focus on these treatment subgroups.

### Table III. Distribution of Treatment Groups

<table>
<thead>
<tr>
<th>Interdisciplinary Pain Management Center</th>
<th>Number of Patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient Population</td>
<td></td>
</tr>
<tr>
<td>Completed Baseline Pain Assessment</td>
<td>646</td>
</tr>
<tr>
<td>Tool and Outcomes Registry (PASTOR)</td>
<td></td>
</tr>
<tr>
<td>Assessment</td>
<td></td>
</tr>
<tr>
<td>Completed Baseline and At Least 1</td>
<td>343</td>
</tr>
<tr>
<td>Follow-Up PASTOR Assessment</td>
<td></td>
</tr>
<tr>
<td>Completed Integrative Modalities Pain</td>
<td>50</td>
</tr>
<tr>
<td>Care Team (IMPACT) ± Functional</td>
<td></td>
</tr>
<tr>
<td>Restoration Program (FRP) and Pre-/Post-</td>
<td></td>
</tr>
<tr>
<td>treatment PASTOR Assessment</td>
<td></td>
</tr>
<tr>
<td>IMPACT Alone</td>
<td>28</td>
</tr>
<tr>
<td>FRP Alone</td>
<td>8</td>
</tr>
<tr>
<td>IMPACT + FRP</td>
<td>14</td>
</tr>
</tbody>
</table>

Integrative Therapies (IMPACT) Program Outcomes

Among the 28 patients who completed the IMPACT program during the PASTOR beta test and for whom we had baseline and post-IMPACT PASTOR data, the largest proportion of patients showed no clinically meaningful change for most PASTOR measures (Fig. 5). The percentage of subjects who reported at least a minimal important improvement ranged from 10% to 40%, depending on the PASTOR domain measured. For some PASTOR domains, 10% to 20% of participants reported worse outcomes following IMPACT participation. It is uncertain if this represents an actual harmful effect of the program or the natural history of chronic pain in this treatment subgroup. With regard to objective functional measures, however, a higher percentage, ranging from 40% to 100% of participants showed improvement depending on the outcome measure, and none showed worsened functional status.

FRP Outcomes

Among the eight patients who completed the FRP alone and completed baseline and post-FRP PASTOR assessments, none...
reported improvement in pain intensity. Among the four participants who endorsed alcohol use, none reported decreased use and one reported increased use. For all other PASTOR domains, a range of 25% to 63% reported improvement (Fig. 6). A range of 13% to 26% of participants reported worsened depressive symptoms, fatigue, anger, sleep impairment, global health, and satisfaction with social roles following FRP participation. With regard to objective functional measures, between 80% and 84% of participants showed improvement depending on the functional measure, and one patient showed decreased lifting strength and treadmill endurance.

Combination Therapy Outcomes: Integrative Therapies Followed by Functional Restoration

Among the 14 patients who completed 6 weeks of the IMPACT program followed by 3 weeks of FRP, 14% reported improvement in pain intensity and a range of 16% to 43% reported improvement in other PASTOR domains (Fig. 7). Between 7% and 28% of participants reported worsened outcomes in at least one PASTOR domain following treatment. With regard to objective functional measures, between 63% and 88% showed improvement and one patient’s performance on the treadmill (8%) declined.

DISCUSSION

The protocol for this study was reviewed by the Madigan Institutional Review Board and was determined to be exempt from full review. The purpose of the study was to collect preliminary descriptive outcomes data for patients who completed interdisciplinary pain care in the Madigan IPMC to establish the foundation for future comparative effectiveness analyses. This goal was realized in 2014 when the Madigan IPMC was awarded a 3-year grant to conduct a randomized clinical trial of the therapeutic benefit of CIM therapies (chiropractic, acupuncture, yoga, and myofascial release) when added to a FRP.

A strength of this study is that it followed the recommendations for pain outcomes research by including assessment of pain intensity, physical functioning, emotional functioning, and overall improvement, all of which are included in PASTOR. In addition, this study included responder analysis methodology. A weakness of this study is its descriptive rather than experimental design. Patients recommended for the FRP were typically of higher baseline physical conditioning status than those recommended for the IMPACT program, so outcomes of the two programs cannot be meaningfully compared. Also, there was no analysis of other factors which may have influenced response to therapy, such as concurrent interventional or pharmacologic pain therapies, status of military readiness and medical board evaluation, age, rank, and psychological comorbidities.

Despite these limitations, some interesting general observations can be made from the evaluation of our programs using the PASTOR tool. For example, pain intensity was the outcome most resistant to change in all treatment groups. This is in contrast to the findings of a Cochrane systematic review and meta-analysis of multidisciplinary psychosocial rehabilitation for chronic low back pain which found low-moderate quality evidence of long-term pain relief in studies comparing multidisciplinary rehabilitation with wait-list controls and usual care, respectively. Another observation was that regardless of treatment group, some patients showed worse outcomes in some domains, particularly depression, anxiety, and anger. The reason for a worsening of emotional status is unclear. Is it possible that participants experienced negative emotions as a direct result of failure to improve despite expending a large time commitment to therapy, or the worsening may represent the natural history of their underlying chronic pain conditions. Another interesting finding was that objective functional measures were more likely to improve than subjective patient-reported measures.
underscores the importance of including objective functional measures among outcomes reported for FRPs, so that participants can be shown evidence of functional improvement before they may perceive it. Lack of observed therapeutic benefit among some participants may be due to the uncontrolled, descriptive nature of this study or due to insufficient sample size to detect therapeutic benefit. In addition, it is possible that the domains currently included in PASTOR may not be sensitive enough to detect therapeutic benefit of interdisciplinary pain therapies. It has been proposed that assessments of kinesiophobia, pain catastrophizing, pain acceptance, and patient activation may be important to the assessment of treatment effectiveness, and measures of these factors may be incorporated into PASTOR as more experience is gained with the tool.

CONCLUSIONS

This descriptive analysis of PASTOR results among service members who participated in interdisciplinary care in an Army IPMC provides preliminary data for future research to determine prognostic factors for favorable response to selected pain therapies in military populations. PASTOR may serve as an international model for outcomes-driven pain research, resource allocation, and decision support. PASTOR may provide the necessary data for identifying best pain practices, determining DoD and VHA pain care standards, and enhancing patient pain care. The opportunities for population-based research on pain treatments and safety-related issues are enormous and hold great potential for improving pain medicine.

ACKNOWLEDGMENTS

We wish to thank Raywin R. Huang, PhD, for his biostatistics expertise. In addition, we acknowledge the valuable input of CDR Steven Hanling, MC USN, and his research team at Naval Medical Center, San Diego. We also thank the Madigan/RHC-P Informatics team including Deputy Chief Richard Barnhill and data analyst Bobbie Solveson whose support made PASTOR implementation possible. The work was funded by U.S. Army Medical Research Acquisition Activity.

REFERENCES

Comparative Efficacy of Multiple Variables of Mesenchymal Stem Cell Transplantation for the Treatment of Neuropathic Pain in Rats

LiPing Liu, MD, PhD†; Zhen Hua, MD, PhD††; Jun Shen, MD, MSc*; Yan Yin, MD††; Jing Yang, MD, PhD*; Kathleen Cheng*; Aijun Liu, MD, PhD*; Lina Wang, MD, PhD*; Jianguo Cheng, MD, PhD*

ABSTRACT Objectives: The current treatment options for neuropathic pain due to nerve injuries are limited and largely unsatisfactory. Mesenchymal stem cell transplantation (MSC) has shown promise as an emerging therapy for neuropathic pain. However, a number of critical parameters, including the sources of cells, the number of cells, and routes of transplantation, need to be elucidated before it can be tested clinically. Methods: MSCs were isolated from rat bone marrow (rBM-MSCs) and adipose tissue (rAD-MSCs) and characterized by flow cytometry and functional differentiation. Rats with chronic constriction injury of the sciatic nerve were transplanted either intravenously or intrathecally with rBM-MSCs or rAD-MSCs in two different doses. The effects were evaluated by using paw withdrawal thresholds in response to noxious stimulation. The MSCs labeled with Dil dye were traced. A total of 75 Sprague–Dawley rats were used for these experiments. Results: Both intravenous and intrathecal transplantation of MSCs significantly attenuated neuropathic pain. Comparable results were achieved by either rBM-MSCs or rAD-MSCs. No differences were noted between the two doses of cell transplantation. MSCs were found on the surface of the spinal cord and dorsal root ganglia. The animals did not show any signs of toxicity throughout the whole course of the experiments. Conclusions: Both intravenous and intrathecal MSC transplantations were safe and efficacious and both rBM-MSCs and rAD-MSCs are suitable for transplantation.

INTRODUCTION
Chronic pain affects 116 million Americans and costs $635 billion annually for medical expenses and lost productivity.1 It is also a leading cause of short- and long-term disability among active duty and retired military personnel.2 Two of every U.S. armed forces veterans have persistent pain attributable to military service. Ineffective treatment leads to pain-related impairments, occupational disability, and medical and mental complications with long-term costs to the military health and disability systems and society at large.2 Neuropathic pain due to nerve injuries is very common. The current treatment options, such as physical, cognitive, pharmacological, and interventional approaches, are limited and unsatisfactory.3 Managing this population of patients represents a major health care challenge.4,5

Mesenchymal stem cells (MSCs) have the ability to differentiate into adipose tissue, bone, and cartilage.6–8 We, as well as others, have also found that human MSC could differentiate into neuron-like cells with strong expression of neuronal biomarkers and direct evidence of electrophysiological functions.9,10 In addition, stem cells have been shown to be neuroprotective in a variety of nerve injury models.11 Furthermore, MSCs may inhibit neuropathic pain by secreting transforming growth factor-beta (TGF-β)12 and activating the peripheral and central endogenous opioid system and the descending inhibitory pathways.13,14 Indeed, the antihyperalgesic effect of MSC transplantation (MSC-TP) was blocked by naloxone, an opioid receptor antagonist that acts peripherally and centrally. The effect was further reversed by focal downregulation of μ-opioid receptors in the brainstem by RNA interference.13

Although MSCs from different sources or different routes have shown promise in reducing neuropathic pain in various animal models,15 several key questions remain to be addressed. First, what is the comparative efficacy of adipose-derived MSCs (AD-MSCs) vs. bone-marrow-derived MSCs (BM-MSCs) in the treatment of neuropathic pain? AD-MSCs as a source of clinical supply has distinct advantages since adipose tissues are abundant, easily accessible, and less invasive to harvest than obtaining bone marrow (BM) tissue. Second, what is the comparative efficacy of intrathecal (IT) vs. intravenous (IV) transplantation? The advantages of using the IV route include easy access, less invasive procedure, and minimal complications. Third, are there lessons to be learned about the mechanisms of action of MSCs by comparing the two different routes of transplantation? Finding the answers to these important questions is required in order to define and optimize the parameters of transplantation for future clinical trials.
In this study, we isolated and characterized MSCs from the BM and adipose tissues. We demonstrated that IV, as well as IT, MSC-TP effectively reversed hyperalgesia induced by chronic constriction injury (CCI) of the sciatic nerve and that both sources of MSCs were remarkably effective in suppressing neuropathic pain. We have thus provided strong evidence that MSC-TP is efficacious and safe in the treatment of neuropathic pain in rats.

**METHODS**

The research protocols were approved by the Cleveland Clinic Institutional Animal Care and Use Committee. Adult male Sprague–Dawley rats (n = 75) weighing 200 to 250 g (Harlan, Indianapolis, Indiana) were group housed (2/cage) on a 12-hour light/dark cycle with food and water available ad libitum. The animals were allowed to habituate to the housing facilities for at least 1 week before starting behavioral experiments, which typically lasted for 42 days.

**Isolation and Culture of BM-MSCs and AD-MSCs From Rats**

Rat bone-marrow-derived MSCs (rBM-MSCs) were isolated from the BM as described with minor modifications. Briefly, 6-week-old male Sprague–Dawley rats were sacrificed by CO₂ asphyxiation according to Institutional Animal Care and Use Committee guidelines, the femurs and tibiae were removed and washed three times with sterilized 1 × phosphate-buffered saline (PBS). The ends of the tibia and femur were cut by sharp Scissors. A 25-gauge needle was inserted into the BM to flush out the tissue with alpha-Minimum Essential Medium and filtered through a 100-μm filter mesh (BD Bioscience, San Jose, California). Rat adipose-derived MSCs (rAD-MSCs) were isolated from fat tissues, in the abdomen and back, which were chopped into small pieces and digested with 0.1% of collagenase I (Invitrogen, Carlsbad, California) for 30 minutes at 37°C of shaker. Single fat tissue cells were obtained by centrifuge down and passing through a 100-μm mesh. Both sources of MSCs were cultured in alpha-Minimum Essential Medium with 16% fetal bovine serum, 1% L-glutamine, and antibiotic solution (100 μ/mL penicillin–streptomycin) in culture flask and incubated at 37°C with 5% CO₂. The medium was replaced after 24 hours and every 3 to 4 days thereafter. rBM-MSCs and rAD-MSCs were characterized by surface markers through flow cytometry and by differentiation into adipogenic and osteogenic cells. Cell differentiation was tested at passage 4 by following manufacturer’s instructions (Rat MSC differentiation kit, sc020) (Fisher Scientific, Pittsburgh, Pennsylvania).

**Characterization of MSCs by Flow Cytometry**

MSCs were expanded to passage 4 and were examined for expression of specific surface markers using flow cytometry as previously described. Briefly, cultured MSCs were harvested, washed, and resuspended in flow cytometry cell sorting buffer (1% bovine serum albumin and 0.1% sodium azide in 1 × Hank’s balanced salt solution). After blocking with CD16/CD32 Abs at 4°C for 30 minutes, cells were stained for surface markers with directly conjugated Abs in flow cytometry cell sorting buffer at 4°C for 30 minutes. Cells were washed twice and resuspended in 200 to 400 μL of PBS for flow cytometry analysis. Abs used were CD44FITC (Clone: OX-49, BD), CD90.1 BV711 (Clone OX-7, BD), CD45APC-CY7 (Clone: OX-1, BD), CD29BV450 (Clone: HA2/5BD). Flow cytometry analysis was performed with LSRFortessa cytometer (BD Biosciences) equipped with CellQuest software (BD Biosciences); 50,000 events were acquired. Data were analyzed with FlowJo software (Tree Star) (FlowJo LLC, Ashland, Oregon).

**The CCI Model**

CCI was performed in rats following the established surgical procedures. Briefly, under pentobarbital anesthesia (50 mg/kg, i.p.), the right sciatic nerve was exposed at midthigh level by freeing the adhering fascia between the gluteus and biceps femoris. Three ligatures (4/0 chromic gut) were tied loosely around the sciatic nerve at 1-mm intervals above the nerve’s trifurcation. The ligatures constricted only about one-third to one-fourth of diameter of nerve and produced a brief twitch in the muscle around the exposure. The surgical side was defined as the ipsilateral side with its control side as the contralateral side in the following analyses. Muscle and skin layers were finally closed with surgical staples. Sham surgeries involved identical procedures, but without ligation of the sciatic nerve.

**Behavioral Tests**

The sensitivity to noxious stimulation was determined by measuring the paw withdrawal thresholds (PWT) in response to mechanical and thermal stimulation to the right and left hind paws. Animals were handled and habituated before behavioral testing to familiarize them with the environment and to minimize stress. All behavioral tests were performed in the Behavior Core Facility by experienced experimenters who were blinded to the treatment. All animals were weighed once a week, and monitored for any abnormal changes in locomotion, food and fluid take, and survival.

**Thermal Hyperalgesia (Hargreaves Test)**

Rats were allowed to habituate in the environment for at least 60 minutes before the behavioral test. Each rat was placed in a box (22 × 12 × 12 cm) containing a smooth glass floor (Stoelting, Wood Dale, Illinois). The temperature of the glass was measured and maintained at 27 °C ± 0.5 °C. A heat source (Stoelting) was focused on a portion of the hind paw and a radiant thermal stimulus was delivered. The stimulus shut off automatically when the hind paw moved or 20 seconds had passed to prevent tissue injury. The intensity of the radiant heat stimuli was adjusted to obtain either short or long baseline latencies. This allowed...
quantization of the treatment effect (lengthening of the latency, relative to the baseline values, and those of the control groups). In this study, latencies for Hargreaves stimuli at baseline ranged from 7 to 11 seconds. The procedure was repeated 3 times at 5 minute intervals to avoid sensitization and withdrawal latencies were averaged and recorded. The experimenters were blinded to group assignments.

Mechanical Allodynia
Animals were placed in individual 10 × 10 × 15 cm plastic boxes on an elevated metal mesh floor and allowed to acclimate for at least 30 minutes before test. Mechanical sensitivity was tested by using von Frey sensory evaluator filaments (Stoelting). Filaments were applied to the plantar surface of the right hind paw in ascending order of force (0.4–60 g for rats) until the filament bent and was held there for approximately 3 seconds or until a paw withdrawal response took place. On a paw withdrawal response, the filaments were applied in descending order, beginning with the next thinner filament until there was no withdrawal response. The threshold was the thinnest filament to evoke a paw withdrawal response. The procedure was repeated 3 times at 5 minute intervals to avoid sensitization and the withdrawal thresholds were averaged and recorded (mean ± SEM).

IT and IV Transplantation of MSCs
IT transplantation of MSCs was performed in the lumbar region, as described by Chen et al. The rat was shaved in lumbar region of the back under anesthesia (40 mg/kg Phenobarbital Sodium), and placed on a rolled pad so the back was arched. After skin disinfection, the L4-L5 lumbar interspace was identified by palpating spinous processes. A needle (27G) was slowly advanced through the skin over the lumbar region, as described by Chen et al. The rat was shaved in lumbar region, as described by Chen et al. The spinal cord and dorsal root ganglia (DRG) were rapidly dissected and postfixed in 4% PFA. The specimens were washed with PBS and placed in cryoprotection buffer overnight at 4°C until the tissues sank. Afterwards, 30-μm-thick coronal sections of the spinal cord, 20-μm-thick sections of the DRG were cut on a cryostat (Leica Microsystems, Wetzlar, Germany) and mounted on slides and stored at −20°C for future use. Low magnification pictures (10×) of the whole spinal cord were captured by anatomy fluorescence microscope with digital camera (Leica, Image Core, CCF, LRI). Sections of the spinal cord and DRG were washed and mounted with 4', 6-Diamidino-2-Phenylindole mounting medium for examination under the microscope (Leica fluorescence microscope) for the Dil labeled cells.

Statistical Analysis
The PWT to mechanical stimulation were expressed as mean ± SEM. Body weight was expressed as mean ± SEM. Statistical analyses were made using two-way analysis of variance followed by paired comparisons with Bonferroni corrections when comparisons were made between more than three groups. Graphpad Prism 5 (GraphPad Software, Inc, La Jolla, California) was used for all the analysis. p < 0.05 was considered statistically significant (* p < 0.05 compared with PBS; # p < 0.05 compared with 5 × 105 cells).

RESULTS
Isolation and characterization of rBM-MSCs and rAD-MSCs: The forward scatter and side scatter of the rBM-MSCs were morphologically larger than their origin BM cells (Fig. 1A). The rBM-MSCs were negative for hematopoietic surface markers CD45, CD11b, and CD31 (Fig. 1B), indicating that there was no contamination with cells from the hematopoietic lineage. In contrast, the BM cells were positive for stromal cell markers CD90.1 (Thy 1.1), CD44H, and CD29 (Fig. 1B). The rAD-MSCs showed the similar characteristics to BM-MSCs with comparable forward scatter and side scatter that were larger than the hematopoietic lineage cells. They also were negative for hematopoietic surface markers CD45, and positive for stromal BM cell markers CD90.1 (Thy 1.1), CD44H, and CD29 (Fig. 1B). These MSCs also showed the capacity to differentiate into adipogenic and osteogenic cells (Fig. 1C).

IT rBM-MSC-TP reversed hypersensitivity induced by CCI: The mean PWTs to mechanical and thermal stimulation were significantly reduced within 4 days after the transplantation. MSCs were transplanted by either IV or IT 14 days after the CCI surgery which the neuropathic pain developed. Labeled rBM-MSCs with Dil dye were traced 1 day, 5 days, or 8 days after transplantation. Rats were sacrificed by intracardiac perfusion with ice-cold PBS, followed by 4% paraformaldehyde solution under deep anesthesia (sevoflurane, Phenobarbital Sodium 50 mg/kg) at the indicated time (Fig. 1). The spinal cord and dorsal root ganglia (DRG) were rapidly dissected and postfixed in 4% PFA. The specimens were washed with PBS and placed in cryoprotection buffer overnight at 4°C until the tissues sank. Afterwards, 30-μm-thick coronal sections of the spinal cord, 20-μm-thick sections of the DRG were cut on a cryostat (Leica Microsystems, Wetzlar, Germany) and mounted on slides and stored at −20°C for future use. Low magnification pictures (10×) of the whole spinal cord were captured by anatomy fluorescence microscope with digital camera (Leica, Image Core, CCF, LRI). Sections of the spinal cord and DRG were washed and mounted with 4', 6-Diamidino-2-Phenylindole mounting medium for examination under the microscope (Leica fluorescence microscope) for the Dil labeled cells.
FIGURE 1. Isolation and characterization of bone-marrow- and adipose-derived MSCs. (A) Cell size and scatter patterns of BM cells, rBM-MSC, and rAD-MSC. rBM-MSC and rAD-MSC showed similar cell size and scatter. (B) Flow cytometry cell sorting analysis revealed that rBM-MSC and rAD-MSC were CD45, CD31, and CD11b negative, and CD90.1, CD44H and CD29 positive (red line) compared to unstained control (blue line). In contrast, BM cells were positive for all of these cell markers (shifted red line), compared to unstained control (black line). These results represent three individual experiments. (C) MSCs were differentiated to adipose cells with lipid droplets accumulated in the cytoplasm stained with oil red (left) and osteoblast cells stained with Alizarin red (right) in respective media. These data represent three individual experiments. Scale bar: 100 uM.
CCI surgery (Figs. 2A and 2B). In contrast, sham surgery without sciatic nerve ligation did not cause significant changes in the mean PWTs. IT transplantation of rBM-MSCs 7 days after the CCI surgery significantly and persistently increased the mean PWTs compared to the IT PBS control group (Figs. 2C and 2D). The lower dose transplantation group (2 × 10^5 cells) was slightly more effective than the higher dose group (5 × 10^5), at least at several time points in the first 2 to 3 weeks after the transplantation. This small but significant difference disappeared in the later part of the experiments that lasted as long as 42 days. Remarkably, the antihyperalgesia effects of both groups after a one-time transplantation were long lasting without any signs of waning to the end of the experiments. No significant changes were observed in the contralateral hind paw either in response to the CCI surgery or to the MSC-TP (data not shown).

Both rAD-MSCs and rBM-MSCs reversed neuropathic pain: IT transplantation of rAD-MSCs (2 × 10^5) or rBM-MSCs (2 × 10^5) produced persistent and significant antihyperalgesia effects (Fig. 3). The mean PWTs were significantly higher in the transplantation groups in most of the time points, compared to the
PBS control group ($p < 0.05$), 1 week after the transplantation. The differences between the rAD-MSC and rBM-MSC groups were not statistically significant at later weeks of the experiments ($p > 0.05$). The effects in both groups lasted to the end of experiment without any signs of waning (42 days after CCI).

Both IT and IV transplantations of MSCs reversed neuropathic pain: The mean PWTs were significantly and persistently increased the mean PWTs, compared to the PBS control ($p < 0.05$; $n = 5$ transplantation groups; $n = 3$ control group). The differences between the 2 transplantation groups were not statistically significant ($p > 0.05$).

Long-term safety of MSC transplantation: All animals survived the whole course of the experiments (Fig. 6A), regardless the route of transplantation (Fig. 6B).

**DISCUSSIONS**

Several labs, including ours, have reported that transplantation of either rAD-MSCs or rBM-MSCs inhibited neuropathic pain in animal models. However, the optimal conditions for effective transplantation and clinical translation remain to be determined. By comparing the relative efficacy of rAD-MSCs and rBM-MSCs, we demonstrated that both sources of stem cells were equally effective and both produced long-lasting antihyperalgesia effects that were consistent in mechanical and thermal tests (Figs. 2 and 3). rBM-MSCs and rAD-MSCs are the best characterized and most commonly used in clinical trials. Although there are some differences between the two types of cells in their gene expression profiles, angiogenic potentials, and secretion of factors, data support the use of both MSC types in various clinical applications. Safety has been largely established in pilot trials and efficacy is being evaluated in a few Phase III trials (e.g., for Crohn’s disease). Although significant antihyperalgesia effects were achieved by both sources of stem cells, AD-MSCs appeared to be a more attractive choice in clinical applications. This type of cells can be obtained in substantially greater quantities (up to 500-fold) and adipose tissue is more abundant, easily accessible, and less invasive to harvest than obtaining BM tissue.

We chose MSCs for transplantation for several reasons. MSCs are considered the safest type of stem cells for transplantation. Both rBM-MSCs and rAD-MSCs have been safely tested in clinical trials to treat spinal cord injury, cardiovascular disease, Parkinson’s disease, and diabetes. Both types of cells are readily expanded in culture to generate a sufficient supply of cells for clinical use. They are essentially nonimmunogenic. In addition to the powerful immunomodulatory effects, these cells may activate the endogenous opioid systems. Indeed, the antihyperalgesia effect of MSC-TP was blocked by naloxone, an opioid receptor antagonist that acts peripherally and centrally. The effect was further reversed by focal downregulation of μ-opioid receptors in the brainstem by RNA interference.

Although a number of advantages have been proposed for IT MSC-TP, our data showed that IT and IV routes of MSC-TP were equally effective in attenuating neuropathic pain (Fig. 4). There was a trend that IT MSCs might have a stronger antihyperalgesia effect than IV MSCs in the later part of the experiments, but the differences were not statistically significant. The IV route would be a preferable if future studies confirm that both routes are equally effective. It is less invasive and easier to use clinically. However, since the cells transplanted by this route are largely trapped in the lungs and complement activation may injure MSCs, the long-term antihyperalgesia effect remains to be determined in even longer term studies. Autologous MSCs from the same individual may be a preferable choice.
FIGURE 5. The fate of MSCs from the IT and IV transplantations. (A) Microscopy of labeled MSCs in vitro. MSCs were labeled with Dil dye (red, left) and stained with 4′, 6-Diamidino-2-Phenylindole for nucleus (blue, middle). Double labeling of MSCs is shown on the right. (B) Microscopy of the lumbar area of the spinal cord, demonstrating MSCs residing in the pia mater surrounding the dorsal side of the spinal cord and the DRGs, 1 to 8 days after transplantation. Cross section of the DRG (C) and the spinal cord (D) showing MSCs residing in the pia mater outside of DRG and spinal cord parenchymal 1 day after MSC-TP. Scale bar: (A) 25 μm, (B) 1 cm, and (C and D) 25 μm.
be considered as these cells suffered less cellular injury than allogeneic MSCs after contacting serum.42

The current strategies to manage neuropathic pain induced by focal nerve compression and inflammation are not satisfactory. For example, compression and/or inflammation of the nerve roots due to intervertebral disc herniation or foraminal stenosis of the spine is one of the most common pain conditions in both military and civilian populations. Epidural injection with local anesthetics and steroids is widely used in the United States and in the world to treat this condition.43 Typically, this procedure only provides short-term pain relief in some patients. Serious complications have been associated with epidural steroid injections, such as meningitis due to drug contamination and paralysis and death due to spinal cord and brainstem infarction, likely caused by the particular steroids used.44,45 MSC-TP could be an attractive alternative. Our data suggest that both systemic and IT delivery of MSCs produced similar attenuation of pain hypersensitivity after constrictive nerve injury. Remarkably, the effects lasted until the conclusion of the study with no signs of waning. Compared to steroid injections, MSC-TP may provide safer, more effective and longer-lasting pain relief by activating the endogenous opioid system, in addition to its potential immunomodulatory, neuroprotective, neuroregenerative, and other mechanisms. There is a downregulation of central opioid receptors after ligation of the sciatic nerve.46 It seems that activation of the peripheral opioid receptors plays a major role in the early stage of transplantation, whereas activation of the central opioid receptors plays a predominant role in the later stages.13

The roles of the MSCs found in the surface of the spinal cord and DRG after IT transplantation remain to be elucidated. It is noteworthy that these cells were located on the pia mater of the spinal cord or DRG and were extremely close to rich vasculatures. It is conceivable that these cells, with their remarkable ability to sense chemical cues, may interact with cells in the circulation and release such molecules as interleukin 10, leukemia inhibitory factor, and TGF to the circulation or cerebrospinal fluid. They may even be able to migrate to other locations such as the injury site. An important recent finding is that exposure of MSCs to serum activated complement of the innate immunity and led to diminished viability and function of MSCs.42 Although all three known intrinsic cell-surface complement regulators were present on MSCs, activated complement overwhelmed the protection of these regulators and resulted in MSC cytotoxicity and dysfunction. All three complement activation pathways were involved in generating the membrane attack complex to directly injure MSCs. Thus, MSC viability and/or function are greatly reduced after IV infusion, despite their potent immunosuppressive activity displayed in vitro. This finding may explain why the results of recent clinical trials with stem cells have been disappointing.42 Theoretically, IT MSC-TP negates direct cell contact of MSCs with the immune cells and minimizes exposure to serum contents. Cells in the IT space may survive for a longer period of time,47 since they are largely shielded from the immune system. The central nervous system is relatively immune privileged, in which the activity of adaptive immunity mediated by lymphocytes is scarce.48 Also, the cerebrospinal fluid is a very suitable medium to support stem cell survival.49,50 Furthermore, since there is evidence that activation of opioid receptors is an important mechanism of the MSC-mediated antihyperalgesia effect, the IT approach represents an easy access to the spinal cord for central antihyperalgesia effects. However, since the effects of IV MSCs were as strong and long-lasting as IT MSCs (Fig. 4), all the theoretical reasoning has to be reconsidered and new mechanisms of action to be examined. There are a number of potential mechanisms for the antihyperalgesia effects. MSCs are known to be immunomodulatory through regulation of immune cells by mechanisms that include both direct cell contact and release of soluble factors through a paracrine mechanism.27,51 Cell contact allows MSCs to directly interact with and regulate immune cells and

FIGURE 6. Different types of MSCs transplanted with different routes did not affect body weight gain. (A) Cell types: neither rAD-MSCs nor rBM-MSCs negatively affect body weight gain, compared to the PBS control (p > 0.05). (B) Administration routes: neither IV nor IT transplantation negatively affect body weight gain, compared to the PBS control (p > 0.05).
inflammatory responses that play critical roles in nociception, regeneration, and tissue repair. The paracrine release of soluble factors such as interleukin 10 (IL-10), leukemia inhibitory factor (LIF), and TGF-β is another mechanism by which MSCs modulate inflammatory and immune processes. The antihyperalgesia effects of MSCs may also be mediated by activation of opioid receptors in the peripheral and central nervous systems.

The relationship between number of cells transplanted and the therapeutic effects may not follow the traditional dose-response curve. We tested two different doses (0.5 and 0.2 million cells) and found that more cells (0.5 million) were not associated with better outcomes (Fig. 6). In contrast, there was a trend that the lower dose group had stronger antihyperalgesia effect. This is consistent with the results in mice from a recent report. In fact, compared to the mice ger antihyperalgesia effect. This is consistent with the results.

CONCLUSIONS

We have demonstrated that IV or IT transplantation of a single low dose of AD-MSCs or BM-MSCs produced a robust and long-lasting inhibition of neuropathic pain in rats. They were traced to the surface of the spinal cord and DRG ipsilateral to the sciatic injury when administered intrathecally. The therapy was safe without any signs of toxicity. Transplantation of AD-MSCs or BM-MSCs has great potential to emerge as an innovative, safe, efficacious, and cost-effective therapy for the treatment neuropathic pain or other chronic pain conditions. Based on our results, we suggest that IV transplantation of AD-MSCs is a preferred strategy to consider in future clinical trials for the management of neuropathic pain.

ACKNOWLEDGMENTS

This work is supported by grants from the Department of Defense (10669042) and Cleveland Clinic Anesthesiology Institute. We thank Dr. Yan Li for assistant with MSC isolation and characterization.

REFERENCES


Chondrotoxicity of Liposomal Bupivacaine in Articular Chondrocytes: Preliminary Findings

CPT K. Aaron Shaw, MC USA; CPT Peter C. Johnson, MC USA; CPT Steve Zumbrun, MS USA; Augustine H. Chuang, PhD; Craig D. Cameron, DO

ABSTRACT Objective: The chondrotoxicity of local anesthetics has been previously recognized. Recent introduction of a liposomal formulation of bupivacaine has been found to significantly improve postoperative pain control but its effect on chondrocyte viability has yet to be investigated with this new formulation. We sought to assess the in vitro chondrotoxicity of liposomal bupivacaine. Methods: Chondrocytes were isolated from articular cartilage and grown in culture medium. Cultured chondrocyte-derived cells (CDCs) were treated with 0.9% normal saline solution, 0.5%, 0.25%, and 0.13% bupivacaine and ropivacaine, 1.3% liposomal bupivacaine for 1 hour. Following treatment, cells were washed and incubated in media for 23 hours. The CDCs were then harvested and viability was assessed by flow cytometry using SYTOX green dead cell stain. Results: Treated CDCs demonstrated a dose-response effect for chondrocyte viability when treated with bupivacaine, ropivacaine, and liposomal bupivacaine. Liposomal bupivacaine demonstrated the highest chondrocyte viability following treatment. Ropivacaine demonstrated higher chondrocyte viability than bupivacaine. Conclusion: Following 1 hour of treatment, liposomal bupivacaine demonstrated the highest chondrocyte viability. Chondrocyte viability was inversely proportional to anesthetic concentration.

INTRODUCTION Intra-articular infusions of local anesthetic agents are common procedures performed by the practicing orthopedic surgeon. Various agents, such as bupivacaine, lidocaine, and ropivacaine, have been used in conjunction with a steroid for degenerative arthritis treatment or for postoperative pain management. Bupivacaine is the most commonly used agent, as well as the best studied in orthopedic practice. Despite its common use, bupivacaine has been associated with toxic effects to articular chondrocytes, resulting in histopathologic changes similar to those of osteoarthritis. Chu et al identified the chondrotoxicity of 0.5% bupivacaine to articular chondrocytes in vitro. Piper and Kim assessed the chondrocyte viability of 0.5% bupivacaine as well and 0.5% ropivacaine, finding that ropivacaine was significantly less chondrotoxic than bupivacaine using an in vitro analysis of human chondrocytes. Despite this data, no clinical reports of chondrolysis following single intra-articular injection use have been reported, rather chondrolysis has been associated with continuous intra-articular pain catheter infusions. Recently, a multivesicular liposomal formulation of bupivacaine has been designed and introduced for clinical use. This liposomal formulation has been shown in clinical studies to demonstrate prolonged analgesia, up to 72 hours following infusion, potentially eradicating the need for continuous intra-articular infusions. Clinical results after local infiltration of surgical wounds have identified liposomal bupivacaine as a safe and effective technique for postoperative pain management. Additionally, it has been shown to result in significantly lower cumulative pain scores and decreased opioid consumption compared with standard bupivacaine. Given its demonstrated efficacy, liposomal bupivacaine use in periarticular infusions warrants further investigation. While its parental formulation is known chondrotoxic agent, the effect of the liposomal derivation on chondrocyte viability should be investigated before in vivo use. The purpose of this study was to perform an in vitro assessment of chondrocyte viability following exposure to liposomal bupivacaine.

MATERIALS AND METHODS Analysis of chondrocyte viability was performed according the protocol described by Piper and Kim. Articular cartilage was harvested from fresh bovine stifle joint. Cartilage was isolated and processed under sterile conditions. The cartilage was removed from bone with a sterile scalpel and digested in sterile 0.2% hyaluronidase (Sigma-Aldrich, St. Louis, Missouri) for 20 minutes at 37°C followed by sterile 0.1% collagenase (Sigma-Aldrich) for four to 6 hours at 37°C. Following digestion, chondrocytes were plated in monolayer culture with fresh media into a 75-cm² flask at a density of 10⁴ cells/cm².

Twenty-four hours before experimental treatment, culture specimens were visualized under phase microscopy.

*Dwight D. Eisenhower Army Medical Center, Department of Orthopaedic Surgery, 300 E Hospital Road, Fort Gordon, GA 30905.
†Dwight D. Eisenhower Army Medical Center, Department of Clinical Investigations, 300 E Hospital Road, Fort Gordon, GA 30905.

This article was presented at the Military Health System Research Symposium, Fort Lauderdale FL, August 17–20, 2015; Georgia Orthopaedic Society Annual Conference, Sea Island, GA, October 1–4, 2015; and DDEAMC Annual Resident Research Conference, Fort Gordon, GA, May 21, 2015; and was accepted for oral presentation at the AAOS Annual Meeting, Orlando, FL, March 1, 2016.

The authors are employees of the U.S. Government. This work was prepared as part of their official duties and, as such there is no copyright to be transferred.

doi: 10.7205/MILMED-D-16-00079
to verify a cell morphology consistent with differentiated chondrocytes (Fig. 1), and the cells were replated into a 6-well Falcon plate at a density of $10^5$ cells per well. Only first-passage chondrocytes will be used. Cultured chondrocytes were maintained in high-glucose Dulbecco’s Modified Eagle Medium, 10% fetal bovine serum, 1% penicillin/streptomycin, and 1% Fungizone and were kept in an incubator at 37°C with 5% CO₂. The medium was changed every 3 or 4 days.

**Experimental Groups**

Cultured chondrocyte-derived cells (CDCs) were subdivided into 1 of 4 treatment groups consisting of 0.9% normal saline solution (Baxter, Deerfield, Illinois), 0.5% bupivacaine (Hospira, Lake Forest, Illinois), 0.5% ropivacaine (Fresenius Kabi, Lake Zurich, Illinois), and 1.3% liposomal bupivacaine (Pacira Pharmaceuticals, Parsippany, New Jersey). All samples were treated according to the same protocol. Specifically, culture medium was aspirated and 200 μL of the treatment solution was added to each well. Samples were incubated in 5% CO₂ at 37°C for 1 hour and the treatment solution was aspirated and fresh culture medium was added. Samples were returned to the incubator, and chondrocyte viability was measured after a 24 hour incubation period.

**Chondrocyte Viability Analysis**

Following the incubation period, culture medium was aspirated from the wells, CDCs were washed once with 1× phosphate-buffered saline solution and detached from the wells with 0.25% Trypsin/0.53 mM EDTA (ATCC, Manassas, Virginia). One milliliter of cell suspension was placed into flow cytometry tubes and 1 μL of SYTOX green dead cell stain (Invitrogen, Carlsbad, California) was added to each tube and mixed in the dark for 20 minutes at room temperature. Samples were assessed with flow cytometry using an Accuri C6 Flow cytometer (BD Biosciences, San Jose, California) at a collection rate of 100 μL/min and fluorescence emission detected using a 530/30 bandpass filter. Emission data were used to determine the chondrocyte viability, recorded relative to amount of chondrocyte death. Experimental groups were compared using student t-tests. The effective concentration of the liposomal bupivacaine was calculated, in milligrams, based on the solution concentration and the pharmacokinetic profile of liposomal bupivacaine reported as demonstrating 3% free bupivacaine. Statistical significance was predetermined as $p < 0.05$.

**RESULTS**

Chondrocyte viability is summarized in Table I based on treatment group. Testing demonstrated a relative lack of chondrotoxicity for the liposomal bupivacaine group (4.8% ± 1.9 nonviable cells), whereas the bupivacaine and ropivacaine groups demonstrate significant levels of chondrocyte death.

When compared against the untreated control CDC’s, Figure 2, there was no significant difference in the percentage of nonviable cells ($p = 0.12$) for 1.3% liposomal bupivacaine. The bupivacaine and ropivacaine treatment groups, demonstrated significantly lower chondrocyte viability ($p < 0.05$) when compared against liposomal bupivacaine.

**TABLE I.** Summary of Viability Testing Based on Treatment Group

<table>
<thead>
<tr>
<th>Treatment Solution</th>
<th>Nonviable Cells (%)</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>1.9</td>
<td>0.06</td>
</tr>
<tr>
<td>1.3% Liposomal Bupivacaine</td>
<td>4.8</td>
<td>1.9</td>
</tr>
<tr>
<td>0.5% Bupivacaine</td>
<td>32.05</td>
<td>9.1</td>
</tr>
<tr>
<td>0.5% Ropivacaine</td>
<td>23.4</td>
<td>4.1</td>
</tr>
</tbody>
</table>

**FIGURE 1.** Chondrocyte derived cells visualized under phase microscopy at 100 x magnification.

**FIGURE 2.** Result of experimental protocol demonstrating comparative chondrotoxicity of liposomal bupivacaine, bupivacaine, and ropivacaine following 1 hour exposure. Bup = bupivacaine; Rop = ropivacaine; Lip. Bup = liposomal bupivacaine.
In calculating the effective concentration of the liposomal bupivacaine treatment, we determined there to be 0.39 mg of bupivacaine in comparison to 5 mg in the standard bupivacaine group.

**DISCUSSION**

The chondrotoxicity of local anesthetic agents has been well recognized in the literature. Reports of chondrolysis following anesthetic administration through intra-articular infusion pumps have been made for both the shoulder and knee. From these early reports, many in vitro analysis have been conducted, indicating the influence of the specific anesthetic, anesthetic pH, presence of synovial fluid, solution preservative, and combination with epinephrine. In vivo animal studies investigating the long-term effects of bupivacaine on chondrocyte viability showed that exposure resulted in decreased proteoglycan synthesis and content 3 months following exposure, as well as decreased cell density at 6 months but no significant difference in viability in comparison to controls.

Similar to previous studies, we found that bupivacaine was more chondrotoxic than ropivacaine. When including the liposomal formulation of bupivacaine, it was found to exhibit the least chondrotoxicity after short-term exposure. This finding could be the result of the delayed release of the bupivacaine from the liposomes resulting in an overall decrease in exposure to the chondrocytes, demonstrated in the difference in effective bupivacaine concentration, 5 mg bupivacaine versus 0.39 mg liposomal bupivacaine. The peak concentration of bupivacaine released from liposomal bupivacaine has been demonstrated within the first hour after administration, however, as we did not perform an assay at the conclusion of the treatment exposure period, we were unable to incorporate this analysis into the effective bupivacaine concentration. Extrapolating from previous studies, the more alkaline pH, ranging from 5.8 to 7.4 per product insert, could additionally be diminishing the chondrotoxicity as it has been previously shown that a pH <5 results in chondrotoxicity.

Liposomal bupivacaine has seen growing use in orthopedic surgery, particularly in arthroplasty. Previous studies have demonstrated that a single administration can provide decreased postoperative pain intensity scores, decreased narcotic consumption, decreased length of stay, and reduced hospital costs. Additionally, it has been found to be safe with no reports of cardiac adverse effects and no wound complications.

This study has several limitations. As an in vitro analysis utilizing bovine chondrocyte, these cells may not represent human chondrocyte response to the treatment solutions. Additionally, the cultured chondrocytes have the tendency to transform into mesenchymal cells following prolonged exposure. Most importantly, this study focused on the short-term treatment effects of liposomal bupivacaine. Previous studies have identified that this liposomal formulation allows for sustained release of bupivacaine, detectable in the blood for up to 96 hours following a single administration with the peak in bupivacaine release documented within the first hour after exposure. The lack of quantification of the bupivacaine concentration at the completion of the study precluded us from performing a more detailed analysis of chondrotoxicity relative to the effective bupivacaine concentration. Ongoing research is needed to characterize the long-term treatment effects of liposomal bupivacaine.

**CONCLUSION**

Single dose exposure of 1.3% liposomal bupivacaine demonstrates minimal chondrocyte toxicity following 1 hour of treatment, with significantly less chondrotoxicity in comparison to 0.5% ropivacaine and bupivacaine. Further research is needed to assess the long-term viability of chondrocytes following exposure to liposomal bupivacaine as well to perform in vivo testing.

**REFERENCES**

9. Chu CR, Izzo NJ, Papas NE, Fu FH: In vitro exposure to 0.5% bupivacaine is cytotoxic to bovine articular chondrocytes. Arthroscopy 2006; 22(7): 6939.

Self-Reported Sleep During U.S. Navy Operations and the Impact of Deployment-Related Factors

Elizabeth Harrison, PhD; Gena L. Glickman, PhD; Shiloh Beckerley, PhD; Marcus K. Taylor, PhD

ABSTRACT  Sleep disruption is a growing concern among military personnel. Very little is known, however, regarding sleep and associated factors in military members serving in combat environments. We sought to quantify the prevalence of sleep disruption among military personnel serving in a combat zone during Operation Enduring Freedom, utilizing a cross-sectional survey of active duty and reserve U.S. Navy personnel in the Afghanistan combat theater (N = 6,118). Survey sleep measures included total hours of sleep per day, total hours of sleep needed to feel well rested, difficulty falling asleep, and difficulty staying asleep. Other reported outcomes included accidents related to the mission. Participants reported an average of 5.9 hours of sleep per day despite needing an average of 6.8 hours to feel well rested. Fifty-seven percent reported insufficient sleep, and this interacted with mission type. Sleep disruption was associated with number of prior deployments, as well as total number of months in a combat zone. Further, those who reported less sleep were more likely to report causing an accident or error that affected the mission. This article documents basic sleep metrics and deployment-related correlates of sleep disruption among military members in a combat zone.

INTRODUCTION  Sleep disruption has been linked to a host of physiological and psychological health problems. Additionally, sleep and circadian disruption lead to increased rates of accidents and errors. In particular, sleep disruption directly impairs cognitive function, negatively affecting performance accuracy, decision-making, and attention and response time. Moreover, these neurocognitive effects appear to accumulate over successive days of disruption. Even chronic partial sleep restriction (6 hours/night) produces serious deficits in working memory and attention, equivalent to those observed after 24 hours of sleep deprivation or with a blood alcohol content of 0.1%. Importantly, these deficits are not reflected in subjective sleepiness scores, indicating that individuals tend to underestimate the effect that sleep deprivation has on their own level of functioning. Certain populations, such as those working extended, night, or rotating shifts, are at an increased risk for sleep and health disorders. In light of the accumulating evidence of adverse effects of sleep disruption on health and performance, the American Academy of Sleep Medicine and the Sleep Research Society have issued a recommendation of between 7 and 9 hours of sleep per night for healthy adults. During deployment, service members are often required to be alert and functional for irregular or long periods of time under demanding circumstances. Therefore, it is important to assess sleep in deployed service members and to determine what, if any, deployment-related factors might mitigate the myriad negative consequences of disturbed sleep.

Previous reports have indicated that ∼56% of deployed naval service members are sleep deprived. Given the reluctance of expeditionary sailors to report and seek help for psychological problems, we hypothesized that service members would underestimate their sleep need to a greater degree than civilians. Further, in view of the known relationships between stress, post-traumatic stress disorder (PTSD), and sleep, we predicted that mission type, deployment history, and combat zone exposure would be related to self-reported sleep in deployed service members. Finally, we hypothesized that self-reported sleep would also be related to rates of reported accidents or mistakes that affect mission outcomes.

METHODS  Survey data were collected from Navy personnel between 2008 and 2012 as part of the Behavioral Health Needs Assessment Survey. Administration, institutional review and approval, and informed consent procedures are described in detail elsewhere. Mobile Care Teams were instructed to administer the survey near the middle of the deployment when possible. Participants were asked the anticipated length of their current deployment in months, as well as how far into the deployment they were at the time the survey was administered. On average, the survey was given 50.9 ± 0.4% (mean ± SEM) through the reported deployment length.

As a part of the survey, participants’ sleep was assessed using several self-report sleep items from Army Mental Health Advisory Team surveys. Participants were asked about the average duration of sleep per day during the current deployment as well as the amount they felt they needed.
to feel “well rested” (in hours, reported as whole numbers). They were also asked to rate the degree of difficulty they experienced falling or staying asleep over the previous 2 weeks, on a scale ranging from 1 (none) to 4 (very severe). Other survey variables included the number of deployments before the current one, the number of cumulative months spent in a combat zone, and the current mission type. Deployment assignments were categorized as one of the following groups: Provincial Reconstruction Team, command/staff, medical, detainee operations (DETOPS), supply/logistics, and other. A subset of participants was also asked whether their sleepiness had caused any accidents or mistakes that affected the mission.

Data analyses were performed using two statistical software programs, SPSS (IBM, Armonk, New York) and GraphPad Prism (GraphPad Software, Inc., La Jolla, California). Welch’s t-tests or analyses of variance with Games–Howell post hoc tests were performed to correct for unequal variances. All statistical tests were evaluated at the p < 0.05 alpha level.

RESULTS

Sleep Deprivation is Pervasive and Interacts with Mission Type

Confirming earlier reports in a subset of the same population, expeditionary sailors reported an average of 5.9 hours of sleep (n = 6,118). This is 1 hour and 6 minutes less than the minimum recommendation put forth by American Academy of Sleep Medicine and Sleep Research Society. Such levels of chronic partial sleep deprivation are known to cause cognitive deficits, particularly in attention and memory. While individuals across all groups experienced chronic partial sleep deprivation, the degree varied by mission type (Fig. 1A). Those who worked DETOPS missions reported a significantly greater sleep deficit compared with other mission types, with an average of 5.4 hours/day (Welch’s F(5,4531) = 33.45, p < 0.001, Games–Howell post hoc tests all <0.001 for contrasts with DETOPS). Across all mission types, less than one-third of individuals surveyed had the recommended 7 or more hours of sleep per day. When focusing specifically on DETOPS, this number was reduced to less than 14%.

Reported Sleep Need Is Lower than Standard Recommendations and Still Greater than Actual Reported Sleep Duration

Participants reported that the average number of hours of daily sleep they needed to feel well rested was 6.8 hours (median response = 7 hours). To provide a measure of sleep deficit, a difference score was obtained for each individual by subtracting the average number of hours of daily sleep from the amount they reported needing to feel well rested. Consistent with an earlier report by our group in a subset of this population, an overwhelming 57% of service members reported insufficient sleep by this measure (Fig. 1B; n = 6,110). Though reported sleep deficit is high for all mission types, more individuals working DETOPS reported insufficient sleep compared with individuals assigned to other mission types (χ²(5) = 56.52, p < 0.001; Fig. 1B).

Sleep Difficulty Interacts With Deployment and Combat History

Expeditionary sailors reported the level of difficulty they had falling and staying asleep over the past 2 weeks on a
scale ranging from 0 (none) to 4 (very severe). These questions were modeled from two items on the Insomnia Severity Index, wherein multiple indices of sleep disturbance are summed to generate a total score. Accordingly, these two measures were combined for each individual to form a continuous scale of sleep difficulty, with higher scores reflecting greater sleep difficulty. Thus, the scale ranged from 0 (none on both questions) to 8 (very severe for both questions). Mean reported sleep difficulty increased with the number of prior deployments (Welch’s $F(2,2064) = 3.48$, $p < 0.05$); those individuals with $\geq 2$ prior deployments reported significantly more sleep disturbance than those with only one previously or never before deployed ($F(2,2064) = 133.7$, $p < 0.001$). Related to that finding, the number of hours of daily sleep decreased with the number of prior deployments ($F(2,5983) = 10.11$, $p < 0.001$); individuals who had been deployed $\geq 2$ times reported significantly less sleep than those who had only one previously or never before deployed (post hoc t-tests with Bonferroni adjustment, $t = 3.32$, $p < 0.01$ for 1; $t = 4.46$, $p < 0.001$ for none; Fig. 2B).

Sleep difficulty was greater for those individuals who had served $7$ or more cumulative months in a combat zone than for those with 6 months or less (Student’s $t$-test with Welch’s correction, $t(770) = 2.85$, $p < 0.01$, Fig. 3). Similarly, there was a pattern demonstrating a decrease in the number of hours of sleep with increased cumulative combat zone exposure, though it was not statistically significant ($t(1075) = 1.09$, $p = 0.28$). Effects of combat zone exposure and number of deployments on sleep were not confounded with age or rank in this sample, since average hours of daily sleep had small but significant positive correlations both with age ($r = 0.084$, $p < 0.001$) and with paygrade ($r = 0.213$, $p < 0.01$).

Sleep-Related Accidents Affecting Mission Are Related to Sleep Duration and Difficulty

A subset of expeditionary sailors surveyed were asked, “During this deployment, have you had an accident or made a mistake that affected the mission because of sleepiness?” The number of individuals reporting such an accident increased as the average number of reported daily hours of sleep decreased ($\chi^2(2) = 14.86$, $p < 0.001$, $n = 1,162$; post hoc for $\geq 7$ vs. 4–6, $p < 0.05$; for $\geq 7$ vs. $\leq 3$, $p < 0.001$; for 4–6 vs. $\leq 3$, $p < 0.01$; Fig. 4A). More than 13% of those who received $\leq 3$ hours of daily sleep reported a sleep-related accident compared with 4% of those who

![FIGURE 2](image2.png)

**FIGURE 2.** Sleep quantity and difficulty and deployment history. Panel A: Average hours of sleep per night decreased with two or more prior deployments. Panel B: Reported sleep difficulty (0–8) increased with number of deployments. For both panels, error bars represent the standard error of the mean for each group. Values in each bar represent the sample size for that group. *$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$.

![FIGURE 3](image3.png)

**FIGURE 3.** Reported sleep difficulty and combat zone deployment. Error bars represent the standard error of the mean for each group. Values in each bar represent the sample size for that group. *$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$.
received 4 to 6 hours and only 1% of those who received the recommended ≥7 hours. Consistent with those findings, an increase in sleep-related accidents was also reported in individuals with a calculated sleep deficit (χ²(1) = 14.08, p < 0.001, n = 1,149; Fig. 4B).

In this subset of our sample, we also obtained a measure of how frequently, on a scale ranging from 1 (never) to 4 (always), individuals felt they slept less than needed specifically because of work hours. Not surprisingly, as the frequency of this measure increased, so did the number of accidents (χ²(4) = 44.67, p < 0.001, n = 1,156; post hoc for always vs. sometimes, p < 0.05; for always vs. seldom and never, p < 0.001; often vs. seldom and never, p < 0.001; all other p values n.s.; Fig. 4C). Interestingly, those individuals reporting sleep-related accidents that affected the mission also had a higher sleep difficulty score (Student’s t-test with Welch’s correction, p < 0.001, n = 1,163; Fig. 4D).

**DISCUSSION**

The present study assessed the relationship between deployment-related factors, including mission type, length of combat zone exposure, and number of deployments on reported sleep. We found that expeditionary sailors chronically experience partial sleep deprivation to a degree known to have negative cognitive outcomes,7,9,10 and further, that the severity varies by mission type. We also determined that expeditionary sailors working DETOPS were at greatest risk for sleep difficulties and deficits.
An overwhelming 57% of service members surveyed here reported insufficient sleep, in sharp contrast to the national average for the civilian population for the same measure (only 33% for workdays and 16% for nonworkdays).21 The large disparity between service members and civilians is even more striking within the context of service members’ lower estimated sleep need. Participants reported that the average number of hours of daily sleep they needed to feel well rested was 6.8 hours. This is approximately 29 fewer minutes of sleep than civilians,21 and lower than the recommended 7 to 9 hours of sleep.12 In the civilian population, there is a consensus with these recommendations; using a similar scale, individuals in the general population estimate their need to be 7.28 hours (median response = 8 or more),21 a value that falls within the lower end of the recommended range.

It may be the case that through training, selection processes, or some other variable, expeditionary sailors actually do need less sleep than civilians to feel well rested. It may also be the case, however, that expeditionary sailors underestimate their sleep need due to a military culture that fosters selflessness, toughness, and a willingness to tolerate discomfort. Importantly, subjective measures of sleepiness have been shown to be dissociable from actual performance on measures of attention and working memory.9 Therefore, while individuals in our sample may have reported feeling well rested with an average of 6.8 hours of sleep, we cannot conclude from this subjective rating that objective performance is at its peak with a lower than recommended amount of sleep. In fact, evidence from other studies in nonmilitary populations suggests that is unlikely.7,9,10,22

In our sample, deployment history (the number of deployments and months spent in combat zones) markedly influenced both reported sleep quantity and sleep difficulty. These results are theoretically consistent with data from the Millennium Cohort Study, which indicate that individuals who have been previously deployed report a significantly shorter sleep time than those who have not; however, the effect of multiple deployments was not explicitly examined in that study.22 It is possible an accommodation response to the negative consequences of sleep deprivation on performance occurs with years of experience, though we have not found evidence to support this. If such a phenomenon does occur, however, it theoretically could help to counter the effects of reduced sleep found with increased deployments observed here. Further studies are necessary to evaluate the influence of sleep debt on performance and readiness across deployments.

Finally, though we did not measure cognition directly, we found evidence that the sleep deficit observed in this sample may have negative consequences on performance and mission outcomes. A previous, publicly available report indicates a similar, negative linear relationship between sleep hours and accidents in Army personnel who served in Operation Enduring Freedom.24 To our knowledge, the present study is the first to demonstrate reported sleep-related accidents that affect the mission for U.S. Navy personnel serving in combat zones specifically. Because health problems are known to be underreported in this population, and sleep deprivation and its resultant cognitive impairments are not always subjectively recognized, we venture that this is likely an underestimate of actual sleep-related incidents. These findings are consistent with the growing body of literature demonstrating a relationship between sleep and performance, including accuracy,5 decision-making,6 and attention and response time.7 It is important to note, however, that due to the wording of our measures, we cannot definitively determine whether reported accidents were a direct cause of impaired performance while awake but sleepy or were a result of individuals actually falling asleep during mission performance. In either case, the short sleep duration observed here is a probable health and safety concern for our warfighters.

The chronic partial sleep deprivation in expeditionary sailors of <6 hours/day may have various other negative consequences. Shorter sleep durations have been linked to myriad physiological and psychological health problems, including increased risk for metabolic, cardiovascular, and mental health disorders.1–3 Furthermore, in a typical nocturnal sleep episode, individuals have more slow-wave sleep at the beginning of the night, and more rapid eye movement (REM) sleep at the end.25 This means expeditionary sailors are not only receiving less sleep overall, but very likely less REM sleep in particular. REM sleep has been linked with emotional memory and “recalibration.”25 Additionally, REM sleep reduction26–28 and disturbances29–31 have been found in individuals with PTSD. Indeed, a relationship between PTSD symptoms and length of sleep was found in a subset of this population.13 These findings, along with others (reviewed by Troxel et al),32 highlight the possibility that the current average daily sleep duration of expeditionary sailors may pose an increased risk for development of PTSD. As these findings are based entirely on self-report, rely on recall for certain items (i.e., number of accidents), and are based on a limited number of measures only, they should be followed up with more controlled studies to objectively quantify the effects observed here via self-report.

While the extent of sleep problems in our military remains a concern, further education regarding sleep need and sleep hygiene may have a strong positive impact and is a relatively low-cost and easily implemented countermeasure. Educational guides could be distributed and briefs could be integrated into curricula in order to provide guidance on optimizing sleep and periodically re-educate individuals about its importance. Topics could include sleep hygiene, the use of stimulants and sleep aids, the biological effects of light at night, and stress management. Such educational programs are beginning to emerge in the form of primarily online resources, including the Department of Defense’s Human Performance Resource Center and the Army’s Performance Triad program, both of which emphasize sleep as a health behavior on par with nutrition and physical activity for optimizing performance. As yet, no such program exists for U.S. Navy service members.
Importantly, our survey data demonstrate lower reported sleep need in active duty service members compared with that found in the civilian population. While this may be partly due to a lack of understanding of what constitutes good sleep health, there may also be a stigma associated with needing sleep or admitting sleep difficulties. Indeed, previous findings among service members have shown a stigma related to reporting physical and mental health issues. Therefore, in addition to reinforcing current standards of sleep need and providing sleep hygiene education, addressing any stigma that may be associated with mental and sleep health may also prove an effective countermeasure. Further, education could be tailored to specifically address the differences in sleep quantity and sleep difficulty by mission type. Finally, surveillance programs employing objective measures of sleep health should be considered.

ACKNOWLEDGMENT
The authors wish to express their profound appreciation for the professional effort and personal sacrifice made by others in collecting the data used in this report. This work (report number 14) was supported by the Bureau of Medicine and Surgery Wounded, Ill and Injured Program, under Work Unit No. 60813.

REFERENCES
Physiological Impact of Platelet Apheresis in Pigs: Oxygen Metabolism and Coagulation

Wenjun Z. Martini, PhD; Cassandra M. Rodriguez, BS; 1LT Jonathan Richardson, MSC USA; James K. Aden, PhD; LTC Andrew P. Cap, MC USA; Michael A. Dubick, PhD

Abstract Introduction: Platelet apheresis is a routine clinical practice, but the physiological impact on the donors has been incompletely characterized. This study measured the effects of platelet apheresis on hemodynamics, oxygen metabolism, and coagulation in pigs to assess its impact before employing the animals in experimental studies. Methods: Forty pigs (39.8 ± 0.6 kg) were anesthetized and catheterized with an apheresis catheter in the femoral vein. During the platelet apheresis process, blood was withdrawn from the pig to separate platelets, and the remaining red blood cells and plasma returned back to the pigs, using the Haemonetics MCS+9000 system. A total of 12 cycles of blood withdrawn and return were performed during the entire apheresis procedure to reduce platelet counts to a target of 50% of baseline. During the process, hemodynamics was recorded in each cycle. Blood samples were collected before and after apheresis to assess changes in oxygen metabolism and coagulation by prothrombin time, activated partial thromboplastin time (STA-R Evolution Stago), and using Rotem thrombelastometry, and platelet aggregation using a Chrono-Log 700 aggregometer. Results: During each cycle of the apheresis, mean arterial pressure (MAP) was decreased and heart rate was increased by blood withdrawal, but both recovered after blood return. On the completion of the apheresis, platelet count decreased from baseline 345 ± 15 10⁹/L to 141 ± 14 10⁹/L and fibrinogen levels were reduced from 124 ± 5 to 99 ± 4 mg/dL (both p < 0.05). Although oxygen delivery remained unchanged, oxygen consumption was decreased from 4.0 ± 0.2 to 3.2 ± 0.0 mL O₂/kg/min (p < 0.05). Rotem alpha (clotting speed) decreased from 79 ± 0 to 69 ± 1° and maximum clot firmness (MCF or clot strength) decreased from 71 ± 1 to 57 ± 1 mm (both p < 0.05). No changes were observed in prothrombin time or activated partial thromboplastin time. Platelet aggregation induced by arachidonic acid or collagen was decreased to 28 ± 6% or 71 ± 3% of baseline values (p < 0.05), respectively. Conclusion: Platelet apheresis caused significant fluctuations in hemodynamics, reduced oxygen consumption, in addition to the compromised platelet aggregation and clotting function expected. The observations warrant consideration in humans undergoing apheresis over extended periods.

INTRODUCTION

Traumatic injury remains one of the leading causes of death.¹,² Uncontrolled hemorrhage is one of the main potentially preventable causes for the mortality in civilian and military settings.³–⁶ Following blood loss, all components involved in the coagulation process are reduced and may be further diluted by resuscitation. Damage control resuscitation has been increasingly recognized and implemented in trauma care over the past decade.⁷–⁹ One important strategy in damage control resuscitation is limiting crystalloids use and increasing blood product use, such as fresh frozen plasma, packed red cells, and platelets. However, limited information is available on the timing and dosage of those products.

Platelets play the vital role of forming the platelet plug in the coagulation process. However, studying the role of platelets in trauma patients is difficult, due to the complexity of tissue injury and simultaneous loss of many coagulation components. Sondeen et al at our institute have developed a modified platelet apheresis procedure to collect platelets, utilizing commercially available platelet apheresis equipment designed for human use.¹⁰ By using the same platelet apheresis methodology, we could deplete platelet levels in swine to investigate resuscitation with blood products in hemorrhaged animals with low platelet counts. In the present study, we characterized the hemodynamic, metabolic, and coagulation alterations to platelet depletion in swine to determine their new baseline physiological status before undergoing hemorrhage and resuscitation with blood products.

MATERIALS AND METHODS

This study was approved by the Institutional Animal Care and Use Committee of the U.S. Army Institute of Surgical Research and was conducted in compliance with the Animal Welfare Act and the Animal Welfare Regulations in accordance with the principles of the Guide for the Care and Use of Laboratory Animals.

Animal Surgical Procedures

We performed apheresis on 40 pigs so it was important to characterize the physiological status of these animals as we assessed their responses of hemorrhage and resuscitation in our subsequent studies. Therefore, data from 40 pigs (crossbreed Yorkshire swine [38.1 ± 1.0 kg], Midwest Research Swine, Gibbon, Minnesota) with apheresis were included in this manuscript. After an overnight fast, the pigs were pre-anesthetized with glycopyrrolate (0.1 mg/kg, IM,
Physiological Impact of Platelet Apheresis in Pigs

Robinul, Baxter Healthcare, Deerfield, Illinois) and tile-tamine HCl–zolazepam HCl (6 mg/kg, IM, Telazol™, Wyeth, Fort Dodge, Iowa). Afterward, pigs were given inhalational isoflurane in 100% oxygen gas (5% for 1 minute then decreased to 3%) for intubation. They were subsequently anesthetized with 1.0 to 1.5% isoflurane in oxygen for catheterization. A catheter was inserted into the carotid artery for measuring MAP, systolic and diastolic blood pressures, and heart rate (HR). A Swan-Ganz catheter was inserted via the jugular vein for mixed venous blood sampling and cardiac output. Another catheter was inserted into the femoral artery for arterial blood sampling for measurement of blood gases and clinical chemistries. The apheresis catheter (Platelet Collection Kit, 994 CF-CPP, Haemonetics Corporation, Braintree, Massachusetts) was inserted in the femoral vein for platelet apheresis and venous blood sampling. Different from humans, pigs have a contractile spleen which can store and release platelets. To better characterize the physiological responses to platelet depletion through multiple apheresis cycles, splenectomy was performed to reduce the confounding effects of platelet sequestration and release from the spleen.

Platelet Apheresis

An automated blood-processing machine (Model MCS+LN9000) for human use with the standard disposable platelethpheresis set (Model 994CF-CPP, Platelet Collection Kit) was used for platelet apheresis in this study. After baseline recording of hemodynamics and blood sampling, the apheresis was performed following the procedures described by the manufacturer, with some of the apheresis settings adjusted as described previously for swine use. During the platelet apheresis, about 500 mL of blood was withdrawn from the pig via the platelet apheresis catheter (referred as blood withdrawal) and mixed with anticoagulant citrate dextrose solution (ACD, Solution A, TERUMOBCT, Lakewood, Colorado), containing 2.45 g Dextrose monohydrate, 2.2 g sodium citrate dehydrate, and 0.73 g citric acid anhydrous per 100 mL. The mixed blood was then centrifuged and separated inside the apheresis machine into red blood cells (RBC), plasma, and platelets. Although the platelets were collected in the disposable platelet apheresis bag, the separated RBC and plasma were returned to the pigs (referred as blood return). The time from the completion of blood withdrawal to the start of blood return was 5 minutes in each cycle of apheresis. To prevent citrate related hypocalcemia, CaCl₂ solution (720 mM) solution was infused to the blood returning to the pigs.

During method development, the platelet counts of the trial pigs ranged about 200 to 300 × 10³/μL. Since a platelet level of 100 to 150 × 10³/μL was considered the critical level for platelet transfusion in trauma patients, we set the target of 50% reduction of platelet counts, etc. The difference between 50% target reduction and the average reduction of 60% in the 40 pigs used is acceptable, since for the subsequent hemorrhage/resuscitation studies we wanted a low platelet count without compromising the animal’s overall physiology significantly. The characteristics of the apheresis process are summarized in Table I.

### Table I. Characteristics of Platelet Apheresis in Pigs

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Numbers of Cycles per Pig</td>
<td>12</td>
</tr>
<tr>
<td>Elapsed Time per Pig (minutes)</td>
<td>235 ± 17</td>
</tr>
<tr>
<td>Total Volumes of Blood Processed per Pig (mL)</td>
<td>7295 ± 71</td>
</tr>
<tr>
<td>Total Volume of Anticoagulant Citrate Dextrose (ACD) Used per Pig (mL)</td>
<td>850 ± 12</td>
</tr>
<tr>
<td>Total Amount of Citrate From ACD Infused per Pig (g)</td>
<td>68 ± 0</td>
</tr>
<tr>
<td>Total Volume of CaCl₂ (720 mM) Infused During Apheresis (mL)</td>
<td>45 ± 1</td>
</tr>
<tr>
<td>Total Volume of Saline Infused per Pig (mL)</td>
<td>1201 ± 4</td>
</tr>
<tr>
<td>Total Volume of Platelet Collected (mL)</td>
<td>400 ± 11</td>
</tr>
</tbody>
</table>

### Platelet Aggregation

Platelet impedance aggregometry was assessed in whole blood samples using a Chrono-Log 700 aggregometer (Chrono-log, Havertown, Pennsylvania). The aggregation was stimulated with either collagen (2 μg/mL) or arachidonic acid (0.5 mM) as agonists, and the area under the curve was used to compare changes of platelet aggregation, as previously described.

### Statistical Analysis

Data are expressed as means ± standard error of the mean and analyzed using SAS statistical software (Cary, NC). A one-way repeated measure analysis of variance using the
PROC MIXED procedure was used to compare changes of oxygen metabolism and coagulation before and after apheresis. A two-way repeated measures analysis of variance was used to compare hemodynamic fluctuations during cycles of blood withdrawal and blood return over time. The statistically significant level was set at \( p < 0.05 \).

**RESULTS**

**Hemodynamics**

All pigs survived the entire platelet apheresis, which lasted an average of 235 ± 17 minutes per pig. In each cycle of the apheresis process, MAP was decreased during blood withdrawal and recovered during blood return (Fig. 1). The magnitude of MAP drop reduced after cycle 3 and stabilized during the remaining cycles (Fig. 1); HR increased during blood withdrawal and returned near baseline level during blood return, with similar change in magnitude through the 12 cycles (Fig. 1). Venous oxygen saturation (SvO\(_2\)) was decreased from 82 ± 1% to 67 ± 2% \((p < 0.05)\) by the 1st blood withdrawal and recovered after the 1st blood return, with similar changing patterns in the remaining cycles (data not shown). Cardiac output decreased from 4.0 ± 0.1L/min to 3.5 ± 0.1L/min at the 1st blood withdrawal and recovered after the 1st blood return, but no significant changes were observed during the remaining cycles (data not shown). On the completion of the 12 cycles of apheresis process, all hemodynamic parameters had returned to baseline values.

**Blood Measurements and Oxygen Metabolism**

As our goal, the platelet counts gradually reduced as the apheresis cycles progressed (Fig. 2). After 3, 6, 9, and 12 cycles, platelet counts reduced to 77 ± 2%, 61 ± 3%, 50 ± 4% and 40 ± 3% of the baseline values, respectively \((all \ p < 0.05)\). As expected, no changes were observed in Hct, hemoglobin (Hgb), or RBC (Table II). Blood glucose was elevated, but lactate and bicarbonate levels were reduced by the apheresis process (Table II). There was no significant change in arterial pH, but base excess (BE) decreased (Table II). Ca\(^{2+}\) and Mg\(^{2+}\) were lower after the apheresis cycles \((p < 0.05)\), while K\(^+\) was unchanged. Total protein and albumin were also lower compared to baseline (Table II).

**TABLE II.** Blood Measurements From Samples Taken Before and After 12 Cycle of Apheresis

<table>
<thead>
<tr>
<th></th>
<th>Before Apheresis</th>
<th>After Apheresis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Platelet counts ((10^3/\mu L))</td>
<td>345 ± 15</td>
<td>141 ± 14*</td>
</tr>
<tr>
<td>Hct ((%))</td>
<td>29.4 ± 0.4</td>
<td>29.8 ± 0.3</td>
</tr>
<tr>
<td>Hgb ((g/dL))</td>
<td>9.8 ± 0.1</td>
<td>9.9 ± 0.1</td>
</tr>
<tr>
<td>RBC ((10^3/\mu L))</td>
<td>5.7 ± 0.1</td>
<td>5.8 ± 0.1</td>
</tr>
<tr>
<td>Glucose ((mM))</td>
<td>4.0 ± 0.3</td>
<td>5.2 ± 0.2*</td>
</tr>
<tr>
<td>Lactate ((mM))</td>
<td>2.2 ± 0.1</td>
<td>1.4 ± 0.1*</td>
</tr>
<tr>
<td>HCO(_3^-)</td>
<td>27.5 ± 0.2</td>
<td>26.2 ± 0.2*</td>
</tr>
<tr>
<td>Arterial pH</td>
<td>7.45 ± 0.00</td>
<td>7.43 ± 0.00</td>
</tr>
<tr>
<td>BE</td>
<td>3.6 ± 0.2</td>
<td>2.3 ± 0.3*</td>
</tr>
<tr>
<td>Ca(^{2+}) ((mM))</td>
<td>1.3 ± 0.0</td>
<td>1.0 ± 0.0*</td>
</tr>
<tr>
<td>Mg(^{2+}) ((mM))</td>
<td>1.8 ± 0.1</td>
<td>1.6 ± 0.0*</td>
</tr>
<tr>
<td>K(^+) ((mM))</td>
<td>3.68 ± 0.04</td>
<td>3.76 ± 0.04</td>
</tr>
</tbody>
</table>

BE, base excess; Ca, calcium; Hct, hematocrit; HCO\(_3^-\), bicarbonate; Hgb, hemoglobin; K, potassium; Mg, magnesium; RBC, red blood cells. Data are expressed as means ± SE from 40 pigs. *\(p < 0.05\) before vs. after.
**Oxygen Metabolism**

Changes of oxygen metabolism are summarized in Figure 3. Oxygen delivery was not affected by the apheresis process, but oxygen extraction was decreased on the completion of apheresis (Fig. 3, A and B). Similarly, oxygen consumption and oxygen demand were also significantly decreased at the end of the apheresis process (Fig. 3, C and D).

**Coagulation**

Blood fibrinogen concentrations dropped from baseline value of 124 ± 5 to 99 ± 4 mg/dL by platelet apheresis ($p < 0.05$). However, there were no significant changes in PT (from 16.0 ± 0.2 to 16.6 ± 0.2 seconds) or aPTT (from 23.6 ± 0.5 to 22.5 ± 0.5 seconds) observed after the full platelet apheresis process.

Coagulation functional profile was compromised by the platelet apheresis process as anticipated by the citrate in the system and the reduction of platelet counts. Clotting initiation time was prolonged from a baseline value of 52 ± 1 to 58 ± 1 seconds after the platelet apheresis ($p < 0.05$). Clotting rapidity (α) and clot strength (MCF) were reduced after the platelet apheresis (Fig. 4). But no significant changes were observed in LI45 during the study (data not shown).

Changes of platelet aggregation (area under the curve) by platelet apheresis are summarized in Figure 5. Arachidonic acid-induced platelet aggregation was decreased from a baseline value of 147 ± 17 to 42 ± 7 (Fig. 5, A1, $p < 0.05$). When adjusted based on platelet counts of 1,000, the observed

* $p<0.05$ before vs. after platelet apheresis

**FIGURE 3.** Changes of oxygen metabolism by platelet apheresis in pigs.

**FIGURE 4.** Changes of clotting speed (alpha) and clot strength (MCF) by platelet apheresis in pigs.
reduction disappeared (Fig. 5, A2). Collagen-induced platelet aggregation was also decreased from baseline value of 301 ± 10 to 235 ± 10 (Fig. 5, B1, $p < 0.05$). However, when adjusted to platelet counts, collagen-induced aggregation was increased (Fig. 5, B2, $p < 0.05$).

**DISCUSSION**

In this study, we assessed the physiological impact of platelet apheresis on the donor in a swine model, using an apheresis instrument designed for humans. The 12 cycles of platelet apheresis effectively reduced platelet counts to 40% of the initial values. All hemodynamic parameters returned to baseline values upon the completion of the platelet apheresis process. However, there were rapid and variable fluctuations of hemodynamics observed during the process. Specifically, MAP dropped and HR rose during each blood withdrawal, but both returned after blood return as expected due to reflex responses. The magnitude of fluctuation of hemodynamic parameters varied, with MAP having the largest fluctuation. Withdrawal of 18% total estimated blood volume caused about a 40% drop in MAP, compared to about a 20% increase in HR and no changes in cardiac output (except during the 1st cycle). Thus, it appears that more sensitive mechanisms are involved in regulating blood pressure in response to changes of blood volume. In addition, the rapid changes and quick return of hemodynamics reflect effective responses in the host to the apheresis process.

The quick recovery of hemodynamics, however, was not observed in oxygen metabolism in this study. On the completion of platelet apheresis, oxygen delivery returned to baseline values, but oxygen consumption and oxygen demand remained decreased. These findings were unexpected as we did not observe any changes in oxygen consumption after 35 or 60% blood loss in our previous pig studies, with only LR resuscitation.13,14 Each cycle of blood withdrawal and return can be considered as a series of hemorrhage and resuscitation, with resuscitation as platelet-depleted whole blood. What we observed in this study reflects the effects of 12 series of hemorrhage and resuscitation. Despite a small volume of hemorrhage (less than 20%) and close to whole blood resuscitation (only reduced in platelets) in each cycle, oxygen consumption was eventually affected after 12 series. The reduced oxygen consumption in this study and sustained oxygen consumption in our previous studies of larger volume hemorrhage and no blood product resuscitation demonstrate that repetitive small blood loss and return has more impact on oxygen metabolism than that of a single much larger volume blood loss with no blood resuscitation. Another possibility is that sustained citrate infusion caused this metabolic shift, either independently of effects on ionized calcium levels or directly due to hypocalcemia. The underlying mechanisms and physiological significance of oxygen metabolic reduction warrant further investigation.

Platelet apheresis also disturbed the coagulation process in this study considering the large drop in platelet counts.
Both clotting speed and clot strength were compromised by platelet apheresis, possibly also due to the reduction in fibrinogen levels and the use of citrate as an anticoagulant in the apheresis process. In addition, platelet aggregation stimulated by either arachidonic acid or collagen was impaired by platelet apheresis. When standardized to platelet counts, collagen-induced platelet aggregation was actually increased, suggesting possible activation from the apheresis process. However, arachidonic acid-induced aggregation, after platelet count standardization, remained unchanged; indicating that collagen-induced aggregation is more sensitive to the apheresis process. This observation is different from what we have observed in responses of anti-inflammatory drugs on platelet aggregation previously. For example, in pig blood samples, arachidonic acid-induced platelet aggregation was completely abolished by four times recommended doses of acetaminophen and/or meloxicam, while collagen-induced platelet aggregation was inhibited to 72% of the control value at the same doses.

Platelet function is dependent on the synthesis of thromboxane A2 (TXA2) from prostaglandin H2, which is generated from arachidonic acid by cyclo-oxygenase (COX-1). The antiplatelet effects of various anti-inflammatory drugs are through inhibition of COX-1 activity and reduction of TXA2 synthesis. Putting these observations together, the different sensitivities of arachidonic acid- and collagen-induced platelet aggregation may suggest that the apheresis process affects the aggregation via different pathways, that is, not through inhibition or COX-1 or TXA2. Future effort is needed to clarify the underlying mechanisms.

Hemodynamics and cardio-pulmonary function were monitored throughout the entire 12 cycles of apheresis in this study. Other than changes of MAP and HR, we did not observe significant signs of vasovagal responses or symptoms related to coronary ischemia. The lack of those signs may likely be due to the small volume of blood withdrawn and rapid return during each cycle of apheresis.

Platelet apheresis is a routine clinical procedure. Most common complications from platelet apheresis are related to citrate-induced hypocalcemia. Ionized calcium (Ca2+) is important for many cellular and physiological processes, including coagulation, muscle contraction and cellular membrane stabilization. When citrate-containing blood returns to platelet donors after apheresis, citrate chelates cations, such as Ca2+ in the circulation and causes the levels of Ca2+ to drop. Consequently, metabolic and cardiovascular complications related to hypocalcemia can develop. In humans, Ca2+ ranging from 1.1 to 1.4 mmol/L is required to maintain normal physiological function. Ca2+ below 1.1 mmol/L is associated with tetany and below 0.8 mmol/L is associated with fatal arrhythmias. In this study, to prevent a citrate related Ca2+ drop, we infused CaCl2 (25 mg/kg/h) during the entire apheresis process, at a rate higher than that recommended for continuous blood loss in pediatric population (10 mg/kg/h). Despite the effort, we were unable to prevent Ca2+ drop in this study. In fact, Ca2+ level was dropped to below normal range after the apheresis process, together with the drop of Mg2+ levels. Thus, a higher dose of Ca2+ should be used for any prolonged apheresis process to maintain Ca2+ homeostasis and prevent citrate related hypocalcemia.

Although the same procedures and equipment were used in this current study as reported previously, we noticed different efficiencies in platelet depletion between this study and the previous one. The difference may be due primarily to differences in size and Hct of the pigs used, since the apheresis operation parameters were calculated by the instrument based on the inputs of pig body weight, Hct, and baseline platelet counts. Differences in pigs would result in different operation parameters, thus, different outcomes. For example, in the previous study by Sondeen et al., pigs of 70 kg and 30 ± 2% Hct were used as donor pigs for platelet collection. Five cycles of apheresis resulted in 40% reduction of platelet counts in those pigs. In the current study, pigs of 38 kg and 29 ± 0% Hct were used. It took 12 cycles to reach 40% of baseline platelet counts, thus a 60% reduction. The differences in pig characteristics, primarily body weight, contributed to the different efficiency of apheresis.

In conclusion, we characterized the physiological effects of platelet apheresis on the donor in a swine model. Platelet apheresis consisting 12 cycles of blood withdrawal and return triggered rapid fluctuations and rapid return in hemodynamics, reduced oxygen metabolism and compromised the coagulation process somewhat. The observations may warrant considerations in humans undergoing apheresis over extended periods.

ACKNOWLEDGMENTS

We appreciate the support received from the Laboratory Support Section at the U.S. Army Institute of Surgical Research in coagulation measurements. This study was supported by the U.S. Army Medical Research and Materiel Command.

REFERENCES

Physiological Impact of Platelet Apheresis in Pigs

The Development of a Probabilistic Dose–Response for a Burn Injury Model

Anthony Iyoho, PhD; Laurel Ng, PhD; Philemon Chan, PhD

ABSTRACT Objective: The objective was to augment a burn injury model, BURNSIM, with probabilistic dose–response risk curves. Methods: To develop the dose–response, we drew on a considerable amount of historical porcine burn injury data collected by U.S. Army Aeromedical Research Laboratory in the 1970s. The experimental parameters of each usable data point served as inputs to BURNSIM to calculate the burn damage integral (i.e., the internal dose) for 4 severities (mild, intermediate, deep second- and third-degree burns). The Binary probability response was constructed and logistic regression was applied to generate the respective dose–response. Historic data collected at the University of Rochester in the 1950s were used for validation. Results: Four dose–response curves were generated, ranging from mild to third degree, with tight 95% confidence bands for mild to deep second degree, and slightly wider bands for third degree. Parametric sensitivity analysis revealed that epidermal and whole skin thicknesses, skin temperature, and blood flow rate have a large effect on predicted outcomes. Conclusions: Addition of dose–response curves provides a critical augmentation to BURNSIM to improve operational risk assessments of burn health. Future recommendations for BURNSIM include the use of body location- and gender-specific parameters with coupling to a thermoregulatory model.

INTRODUCTION
Recent advances in the development of novel nonlethal (NL) multistimuli devices such as flashbangs, have led to a higher burn injury risk; where larger fireballs are produced for more effective visual suppression of hostile targets. To assess the risk of injury for NL systems which use multistimuli devices such as flash bangs, the Human Effects Center of Excellence (HECOE) calculates an aggregate Risk of Significance (RSI), which estimates probability of “significant” injury. Some injury models provide continuous probabilistic dose–response curves while others are based on a threshold value of injury. The latter type of analysis can overestimate the RSI values and prematurely fail the safety requirement of the NL system. Therefore, there is a need to increase the accuracy of the RSI calculation through replacement of threshold injury predictions with continuous probabilistic dose–response estimations. BURNSIM, the primary burn risk model used by a number of Army, Air Force, and Navy agencies including HECOE, uses a threshold prediction and needs to be refined to produce a dose–response prediction.

BURNSIM, originally developed at the U.S. Army Aeromedical Research Laboratory (USAARL) by Dr. Francis Knox and later extended at the Air Force Research Laboratory, is a biomechanically based model that includes bioheat transfer components such as boiling, tissue property dependence on water content, heat flux penetration, and dermal blood flow. BURNSIM calculates thermal transfer through the skin using Fourier’s heat conduction equation, which is solved numerically using the Crank–Nicholson finite difference method where three composite layers (epidermis, dermis, and subcutaneous fat) are considered. Most burn injury models to date have been purely theoretical or validated only against nonsevere human burns, whereas BURNSIM has been demonstrated to compare well with a considerable amount of porcine burns of mild to high severity. The pig has long been used in burn experiments due to the striking resemblance between pig and human skin in regard to physiology, anatomy, and physical appearance.

Burn severity is classified into degrees according to the damage integral calculation. The damage integral is calculated at each depth and governed by a first-order Arrhenius equation that is integrated over the entire duration that skin temperature is above 44°C. For a given heat insult, the burn depth is the skin depth where the damage integral is equal to 1. Therefore, a second-degree burn occurs if the integral of the damage rate (Ω), is greater than or equal to 1 at the epidermis/dermis interface. Similarly, a third-degree burn occurs if Ω ≥ 1 at the dermis/subcutaneous fat layer interface. Additionally, BURNSIM predicts time to pain for six levels of pain intensity. BURNSIM has been validated against hundreds of first-, second-, and third-degree porcine burns; however, there is no human validation for moderate to severe burns.

One key shortcoming of BURNSIM is that it does not provide probabilistic dose–response curves for the different severity of injuries and needs to be modified to include this. The necessary data required to develop such a relationship can draw on the considerable amount of historical data that supported the original BURNSIM model development. Therefore, the objective of this work was to augment the capabilities of BURNSIM by developing probabilistic dose–response curves for various severity levels of burn (mild second degree, intermediate second degree, deep second degree,
and third degree). This work supports the overall modeling and simulation effort that will be incorporated into the Human Effects Modeling Analysis Program sponsored by the Joint Nonlethal Weapon Directorate, HECOE, and General Dynamics Information Technology.

METHODS

The augmentation of BURNSIM to include probabilistic dose–response curves was accomplished via the following tasks. First, burn data to be used for the development and validation of the dose–response curves were identified. Next, the damage integrals (i.e., internal dose) were calculated for each usable experiment of the development data set. Then, logistic regression curves were fitted at varying burn severity levels and validated against independent data sets. Finally, a parametric sensitivity study was conducted to analyze the robustness of the results. There was no change to the formulation and the numerical method of the BURNSIM model.

Historic Burn Data

Burn data were needed for the development and validation of the dose–response curves. We identified data generated by USAARL and the University of Rochester to be used for developmental and validation purposes, respectively. The USAARL data features over 1,500 burns produced on white domestic swine (43 ± 8 kg). The pigs were exposed to two sets of 6 holes on the left and right side, for a total of 12 burn sites on the lateral surface. Before testing, the hairs of the pig were trimmed closely and anesthesia was applied. For each experiment, room temperature was stable but not constant day to day. The thermal source was a NASA-Ames T-3 furnace using JP-4 fuel which delivered 0.7 to 3.92 cal/cm²/s of heat flux. Biopsies, used to determine burn depth, were taken from each site for histological analysis 24 hours postburn. When the depth of burn was not uniform, the maximum depth of damage was used for grading. The burn depth was also corrected after the fact for tissue shrinkage. Refers to USAARL reports for further details about the burns.

Calculation and Validation of Dose Response

To develop the dose–response curves, we relied on the considerable amount of data collected by USAARL in the 1970s. The USAARL dataset features over 1,500 individual burns of varying severity. We separated out bare and blackened skin experiments from the USAARL data (i.e., no clothed skin burns) and removed bad data (i.e., data with missing parameters) to arrive at 810 useful data points. For each data point, many experimental parameters were recorded. Of these parameters, the exposure duration, heat flux, initial skin temperature, core temperature, epidermal thickness, and whole skin thickness used by each experiment were used to simulate BURNSIM to calculate the damage integral at the bottom of the epidermis, one-third into the dermis, two-thirds into the dermis, and at the bottom of the dermis, which corresponds to the threshold of mild second-degree, intermediate second-degree, deep second-degree, and third-degree burns, respectively. The observed microscopic grade (or micrograde) for each experimental point was used to categorize each burn as no burn, first, mild second, intermediate second, second, or third degree (Table I). Microscopic burn grades were evaluated at 24 hours postburn. To construct the binary probability response for at least mild second-degree severity, data points with an observed burn micrograde of 6 or higher were assigned a value of 1 (i.e., 100% probability); data points with micrograde less than 6 were assigned a value of 0 (i.e., 0% probability). A similar analysis was performed for at least intermediate second-degree, at least deep second-degree, and at least third-degree severity using threshold microgrades of 7, 8, and 9, respectively. Of the 810 useful data points, 288 (35.6%), 126 (15.6%), 78 (9.6%), and 57 (7.0%) points were of microgrades 6, 7, 8, and 9, respectively. For mild second-degree,
intermediate second-degree, deep second-degree, and third-degree severity, the binary probability was plotted against the damage integral at the bottom of the epidermis, one-third into the dermis, two-thirds into the dermis, and at the bottom of the dermis, respectively. To produce the third into the dermis, two-thirds into the dermis, and at degree severity, the binary probability was plotted against the corresponding damage integral, which is derived from the logit function:

\[ F(\Omega) = \frac{1}{1 + \exp[-(\beta_0 + \beta_1 \log(\Omega))] \text{ (1)}} \]

where \( \beta_0 \) and \( \beta_1 \) are fitted coefficients of the logistic regression curve, which is derived from the logit function:

\[ \logit(F(\Omega)) = \ln \left( \frac{F(\Omega)}{1 - F(\Omega)} \right) = \beta_0 + \beta_1 \ln(\Omega) \text{ (2)} \]

The fit of the regression curves was verified through a comparison against the percentage of burn within damage integral bins. A logistic regression curve for at least first-degree severity was not constructed because there was not enough data classified as no burn. However, the combined probability of no burn plus first-degree burn can be determined by subtracting the probability of at least mild second-degree burn from 100%.

A validation study was carried out to compare the predictions to other data that were not part of the original data used for BURNSIM dose–response development. For each exposure condition of the validation dataset,16,17 BURNSIM was simulated to calculate the damage integral at the bottom of the epidermis, one-third into the dermis, two-thirds into the dermis, and at the bottom of the dermis. The probability of at least mild second-degree, at least intermediate second-degree, at least deep second-degree, and at least third-degree burn observed for each exposure condition (versus the corresponding calculated damage integrals) was plotted against the corresponding dose–response curves and \( R^2 \) was calculated to assess agreement.

### Parametric Sensitivity

Parametric sensitivity calculations were carried out to evaluate the robustness of the dose–response curves. We identified five parameters: initial skin temperature, core temperature, epidermal thickness, whole skin thickness, and blood flow rate for parametric sensitivity analysis. These particular parameters were selected because they can vary quite significantly. For each of these parameters, mean, minimum, and maximum values were determined as shown in Table II. The minimum and maximum values for initial skin temperature and core temperature were determined by simulating cold (minimum values) and heat-exercise (maximum values) stress with a thermoregulatory model developed by the authors.19 A core temperature of 39.5°C can be achieved during rigorous exercise without reaching heat exhaustion if the ambient temperature is cool.20 Skin thicknesses and blood flow rates are parameters included in the list.

#### Table I. Microscopic Burn Grade (or Micrograde)

<table>
<thead>
<tr>
<th>Severity</th>
<th>Micrograde</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Burn</td>
<td>0</td>
<td>No Thermal Damage</td>
</tr>
<tr>
<td>First Degree</td>
<td>1</td>
<td>Cell Damage Without Acidophilism</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Partial Epidermal Acidophilism</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Complete Epidermal Acidophilism</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>Partial Dermal–Epidermal Separation</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>Complete Dermal–Epidermal Separation</td>
</tr>
<tr>
<td>Second Degree</td>
<td>6</td>
<td>Mild Dermal</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>Intermediate Dermal</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>Deep Dermal</td>
</tr>
<tr>
<td>Third Degree</td>
<td>9</td>
<td>Complete Dermal to Adipose Border</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>Adipose</td>
</tr>
</tbody>
</table>

#### Table II. Mean, minimum, and maximum parameter values utilized during the parametric sensitivity analysis

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Units</th>
<th>Mean</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial Skin Temperature</td>
<td>°C</td>
<td>33</td>
<td>20</td>
<td>38</td>
</tr>
<tr>
<td>Core Temperature</td>
<td>°C</td>
<td>37</td>
<td>35</td>
<td>39.5</td>
</tr>
<tr>
<td>Epidermal Thickness</td>
<td>μm</td>
<td>75</td>
<td>40</td>
<td>200</td>
</tr>
<tr>
<td>Skin Thickness</td>
<td>μm</td>
<td>1,125</td>
<td>500</td>
<td>2,000</td>
</tr>
<tr>
<td>Blood Flow Rate</td>
<td>m(^3)-blood/m(^3)-tissue/s</td>
<td>0.02</td>
<td>0.01</td>
<td>0.1</td>
</tr>
</tbody>
</table>

Cold Stress: \( T_{sk} = 20°C, T_{core} = 36°C, BF = 0.01 \text{ m}^3\text{-blood/m}^3\text{-tissue/s} \)
Heat Stress: \( T_{sk} = 36°C, T_{core} = 38.5°C, BF = 0.1 \text{ m}^3\text{-blood/m}^3\text{-tissue/s} \)

The parameters included the initial skin temperature (\( T_{sk} \)), core temperature (\( T_{core} \)), epidermal thickness, whole skin thickness, and blood flow rate (BF). Cold and heat stressed cases were also analyzed.
the nominal and altered parameters at the exposure condition resulting in 50% probability for the nominal case. The above analysis was also performed for a cold and heat stress case where initial skin temperature, core temperature, and blood flow rate were altered simultaneously (Table II).

RESULTS

Dose–Response Curves

The results of the dose–response development are shown in Figure 1. The binary response of the USAARL data for at least mild second-degree, at least intermediate second-degree, at least deep second-degree, and at least third-degree severity is shown along with the fitted dose–response curves. The probability within damage integral bins for each severity verified that the dose–response curves were a good fit for the binary response data. The 95% confidence bands were narrow for mild second-degree burn and increasingly wider for the higher severities. Particularly, the confidence bands at third-degree burn severity were wide due to the lower quantity of data as compared to lower severities. The parameters of the dose–response curve are summarized in Table III.

Parametric Sensitivity

The sensitivity analyses indicate two categories of parameters can significantly affect the accuracy (up to 50% error) of the burn prediction: (1) location specific boundary conditions (i.e., regional epidermal and whole skin thickness) and (2) initial physiologically based conditions (e.g., skin temperature, blood flow). The results of the parametric sensitivity analysis are shown in Table IV.

The resultant RMSE values show that the regional specific parameters such as epidermal thickness and whole skin thickness have the largest effect on second-(up to 32.5% error magnitude) and third-(up to 49.8% error magnitude) degree burn probabilities, respectively. Increases in epidermal

The results of the validation study are shown in Figure 2. For all severities, the dose–response curves showed good visual agreement with the University of Rochester validation data. Additionally, the $R^2$ values for at least mild second-degree, at least intermediate second-degree, at least deep second-degree, and at least third-degree predictions were 0.82, 0.82, 0.80, and 0.91, respectively; indicating a very good agreement with data.
thickness decrease the probability of second-degree burn (−27.9% error) but has no effect on third-degree burn probability. The lack of effect on third-degree burns was because changing epidermal thickness does not affect the likelihood of third-degree burns if whole skin thickness remains unchanged. Increases in whole skin thickness results in a decrease in the probability of third-degree burns (−40.3% error), but has a minimal effect on second-degree burn probability (up to 4.5% RMSE and 3.0% error).

Physiologic conditions such as initial skin temperature and blood flow can have a large effect on the burn probability and are directly dependent on the operational scenario (e.g. air temperature, humidity, activity). Changes in initial skin temperature have a relatively large effect on the probability of second- and third-degree burns (up to 11.6% and 32% error magnitude, respectively). Increases in initial skin temperature increase the probability of second- and third-degree burns. However, changes in core temperature have a minimal effect on second- and third-degree burns (0 and 2.2% error, respectively). Increases in blood flow have a minimal effect on second-degree burn probability (4.9% RMSE and −0.2% error) but significantly decreases the probability of third-degree burn (−25.1% error). Going from thermoneutral to cold stressed greatly reduces the probability of second- and third-degree burn (−11.6 and −29.4% error, respectively); an effect that is mediated mostly by decreases in initial skin temperature. Going from thermoneutral to heat stressed has a minimal effect on second-degree burn probability (3.6% RMSE and 1.4% error) but significantly reduces the probability of third-degree burn (−17.0% error), an effect that is mediated mostly by increases in blood flow.

**DISCUSSION**

The main objective of this work was to expand BURNSIM to include probabilistic dose–response curves at four burn severities: at least mild second degree, at least intermediate second degree, at least deep second degree, and at least third degree. BURNSIM is one of the most complete burn injury models, implements a number of thermo-physiological mechanisms that alter heat transfer through the skin, and has been validated against severe burn. However, BURNSIM originally did not provide probabilistic dose–response curves for the different severity of injuries, which is needed for risk assessments for military and weapon systems. The resultant logistic regressions curves were shown to be a good fit to the USAARL data when it was reduced to binary data and compared well against the University of Rochester validation data, which imparts confidence in use of the developed dose–response curves or probability risk assessment at different burn severities. In developing the dose–response curves, we used individual burn data from USAARL experiments. We accounted for epidermal thickness, whole skin thickness, initial skin temperature, and core temperature, but there was
In this work, we calculated the dose–response curves at varying degrees of second-degree severity (i.e., mild, intermediate, and deep). Distinguishing second-degree burn severity is important since superficial second-degree burns can be treated on an outpatient basis as long as special regions like the hands and face are not involved; whereas deeper second-degree burns (approximately two-thirds into the dermis) require hospital care. The USAARL data naturally distinguished burns of varying second-degree severity when grading burn injury on porcine skin as shown in Table I. The damage integral was calculated at bottom of epidermis, one-third into the dermis, two-thirds into the dermis, and at the bottom of the dermis for mild second-degree, intermediate second-degree, deep second-degree, and third-degree burns, respectively. The dose–response curves compared well with validation data from University of Rochester, which also distinguished second-degree burns, indicating the location of the calculated damage integrals were appropriate.

In calculating the damage integral, we used BURNSIM’s default damage coefficients. The default coefficients were determined by Weaver and Stoll for the epidermis and by Takata for the dermis. These same default damage coefficients are also used by the American Society for Testing and Materials F1930 standard for evaluating the protection provided by flame resistant clothing, and have also been used in other burn injury models. 5 The fitted dose–response parameters shown in Table III are calculated based on the default damage coefficients. Therefore, if a BURNSIM user significantly altered the damage coefficients, the current dose–response curves would be invalid. Thus, the dose–response curves developed in this work require the use of BURNSIM’s default damage coefficients.

In this work, the threshold of second-degree burns corresponded to micrograde of 6 (superficial dermal damage), which was revised from the previous threshold definition of 4 (partial-epidermal separation). 14 By definition, a second-degree burn entails damage into the dermis; using micrograde 6 as the second-degree threshold is consistent with this definition. This revision resulted in better agreement with the University of Rochester validation data compared to the original definition.

The parametric sensitivity analysis revealed a need for specificity of regional epidermal and whole skin thicknesses for an improved comprehensive burn injury risk assessment.

### TABLE IV. Results of the Parametric Sensitivity Analysis

<table>
<thead>
<tr>
<th>Parameter Change</th>
<th>Second Degree</th>
<th>Third Degree</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RMSE (%)</td>
<td>Error (%)</td>
</tr>
<tr>
<td>Minimum Skin Temperature</td>
<td>10.4</td>
<td>–11.6</td>
</tr>
<tr>
<td>Maximum Skin Temperature</td>
<td>3.9</td>
<td>3.0</td>
</tr>
<tr>
<td>Minimum Core Temperature</td>
<td>0.2</td>
<td>0.0</td>
</tr>
<tr>
<td>Maximum Core Temperature</td>
<td>0.3</td>
<td>0.0</td>
</tr>
<tr>
<td>Minimum Epidermal Thickness</td>
<td>18.0</td>
<td>32.5</td>
</tr>
<tr>
<td>Maximum Epidermal Thickness</td>
<td>17.0</td>
<td>–27.9</td>
</tr>
<tr>
<td>Minimum Skin Thickness</td>
<td>3.3</td>
<td>–2.0</td>
</tr>
<tr>
<td>Maximum Skin Thickness</td>
<td>4.5</td>
<td>3.0</td>
</tr>
<tr>
<td>Min Blood Flow</td>
<td>0.5</td>
<td>0.0</td>
</tr>
<tr>
<td>Max Blood Flow</td>
<td>4.9</td>
<td>–0.2</td>
</tr>
<tr>
<td>Cold Stress</td>
<td>10.2</td>
<td>–11.6</td>
</tr>
<tr>
<td>Heat Stress</td>
<td>3.6</td>
<td>1.4</td>
</tr>
</tbody>
</table>

RMSE between the nominal and altered cases was calculated to determine the sensitivity with respect to each parameter over the whole probability range (0–100%). The error at the exposure condition resulting in 50% probability for the nominal case was also calculated (“error”) and is indicative of the magnitude and direction of change.

The Development of a Probabilistic Dose Response for a Burn Injury Model

MILITARY MEDICINE, Vol. 182, March/April Supplement 2017

207
Epidermal thickness can vary five-fold from 40 μm (chest) to 200 μm (hand),21 which we have shown will produce up to 32% error in second-degree burn outcome, if the correct skin thickness is not chosen. Likewise, whole skin thickness can vary four-fold from 500 μm (face) to 2000 μm (back),21 which can produce up to 50% error in third-degree burn outcome. Regional skin thicknesses can also vary considerably between genders.21 Therefore, it is paramount to account for differences in epidermal and whole skin thickness between body regions and gender when predicting burn outcomes.

It should be noted that we used a maximum skin temperature of 2,000 μm for our parametric study; however, values greater than 4,000 μm have been observed elsewhere. The use of 4,000 μm results in a less conservative prediction (i.e., lower probability) of third-degree burn. Currently, the use of 2,000 μm results in a reduction of third-degree burn probability (~40%) from the nominal case of 1,125 μm. Using 4,000 μm would result in an even greater reduction.

The parametric sensitivity analysis also revealed a need for accurate assessment of initial skin temperatures and blood flow conditions. Our analysis showed that variations in initial skin temperature and blood flow can produce up to 32 and 25% error in third-degree burn outcome, respectively. Skin temperature and blood flow will vary according to body region and physiological stressors (e.g. cold or heat stress). We showed that typical heat and cold stress can produce up to 12% and 30% error in second- and third-degree burn outcomes, respectively. Berkley et al.25 found a similar magnitude difference in burn outcomes for pigs exposed to cold room temperatures (which reduced skin temperature by approximately 10°C) before heating in comparison with normal room temperatures. Therefore, BURNSIM should be integrated with a thermoregulatory model that can provide skin temperatures and blood flow rates for the whole body by accounting for the effects of ambient temperature, humidity, clothing, activity level, and gender, all of which are part of the description of any operational scenario. The use of a thermoregulatory model to predict blood flow rate is especially important since a typical user would not know this value. Any human thermal model that predicted regional skin blood flow rate could provide this input to BURNSIM. The authors have developed such a model under the sponsorship of the U.S. Army Medical Research and Materiel Command.19

BURNSIM’s probabilistic dose response has been applied to thermal injury predictions from furnace fires (i.e., USAARL data) and carbon arc sources (i.e., University of Rochester data); but has application to other types of thermal injury. BURNSIM can already be simulated for purely radiant, convective, and conductive source (or for any combination of these heat transfer modes) and contact burns. BURNSIM also considers heat flux penetration and the power spectrum of the heat source; and is therefore applicable to predicting thermal injury from lasers and heat lamps (e.g. xenon arc and quartz). The use of an internal dose (in our case, the damage integral) allows for injury predictions for any given heat source as long as the insult can be characterized.

CONCLUSION
In this work, we have augmented the output of BURNSIM by developing probabilistic dose–response curves at four burn severities. To our knowledge, this is the first burn injury model to calculate a probability risk. Dose response is needed for probabilistic risk assessments for military and weapon systems; this improves on the current all-or-nothing prediction of the legacy BURNSIM software. This capability allows designers to more adequately understand the trade-offs (i.e., effectiveness vs. burn risk) when developing NL weapons or assessing occupational or training environments.

For added functionality, we have also developed a software application for HECOE that integrates the dose–response curves with BURNSIM. Developing probabilistic dose–response curves addresses one of the top needs in the current burn injury model, but future improvements including parameter regionalization, gender generalization, and thermoregulatory model integration should be made to improve the accuracy of predicted burn outcomes.

ACKNOWLEDGMENTS
This work was funded through General Dynamics Information Technolog (GDIT) under contract FA8650-13-D-6368/0004. The opinions or assertions contained herein are private views of the authors, and are not to be construed as official or as reflecting views of the Department of the Army or the Department of Defense. Cleared by JNLWD for public release on February 5, 2016.

REFERENCES


A Novel, Inexpensive Method to Monitor, Record, and Analyze Breathing Behavior During Normobaric Hypoxia Generated by the Reduced Oxygen Breathing Device

Leonard A. Temme, PhD*; Paul St Onge, PhD†; Mark Adams, MBBS*; David L. Still, OD, PhD*; Jonathan K. Statz, MS‡; Steven T. Williams, MS†

ABSTRACT Objectives: Since hypoxia remains one of the most important physiological hazards the aviation environment poses, continues to be identified as a contributing factor to many aviation incidences and mishaps. While several types of hypoxia are commonly differentiated, the most important for aviation is hypoxic hypoxia, which occurs when the reduced partial pressure of O₂ in the inspired air effectively reduces the O₂ available in lung alveoli to such an extent that the O₂ concentration in arterial and capillary blood is reduced. Because of the risks hypoxia poses, aviation personnel are trained to recognize the symptoms of hypoxia so they can implement appropriate procedures and countermeasures. Such training is necessary for aviators in pressurized aircraft equipped with supplemental oxygen systems in which equipment failures may occur. Aviators in nonpressurized, rotary-wing aircraft may sustain sufficient hypoxia to affect cognitive and sensory functions. Consequently, hypoxia familiarization remains an essential component of current United States military aviation training.

A common way to generate hypoxia is to reduce the percent of O₂ inspired while keeping the air pressure constant. This type of hypoxia involves offsetting the O₂ with a physiologically inert gas, usually N₂. For example, while keeping air pressure constant at sea level, increasing by 10% the N₂ content in the inspired air to 88.09% and reducing by 10% the O₂ content in the air to 11%, produces a resultant partial pressure of O₂ approximating that found at an altitude of about 16,000 ft above mean sea level (MSL). This method, used commonly in aviation training and research, generates a form of hypoxia called normobaric hypoxia (NH) because it does not involve a change in barometric pressure. Equipment frequently used to generate NH is the Reduced Oxygen Breathing Device (ROBD), a commercial, off-the-shelf, computer-controlled, gas-blending system specifically designed for aviation training and demonstration.

The standard tool for monitoring hypoxia in clinical and research environments is pulse oximetry, a noninvasive measure of the O₂ saturation of arterial blood (SpO₂). In normal healthy people, SpO₂ values at sea level rarely fall below 95%; but the range in SpO₂ between individuals at altitude can be immense, as illustrated in Figure 1, which summarizes the results of a survey of 57 published studies of fit, healthy subjects. Differences in individual physiology may account for at least some of the extraordinary range in the normal SpO₂ values found for a given altitude. One source of the range in individual differences may be the effect acute hypoxia has on respiration. By increasing the volume of air moved in and out of the lungs, either by breathing faster or deeper, a hypoxic individual can dramatically increase SpO₂ from less than 70% to more than 97%. However, this relatively common hypoxic hyperventilatory response (HVR) can have important physiological consequences because it can result in hypocapnia, a reduction in the amount of CO₂ in the blood sufficient to produce respiratory alkalosis. These changes in blood chemistry, precipitated by the HVR, can

*Visual Protection and Performance Division, U.S. Army Aeromedical Research Laboratory, 6901 Farrel Road, Fort Rucker, AL 36362.
†Laulima Government Solutions, Orlando, FL 32826.
‡Oak Ridge Institute for Science and Education, Oak Ridge, TN 37830.

doi: 10.7205/MILMED-D-16-00053
produce cerebral vasoconstriction; and, if breathing is not properly monitored and controlled to allow a normal accumulation of CO₂, the HVR can produce a loss of consciousness due to reduced cerebral blood flow. Consequently, changes in respiration rate can radically alter physiological responses to hypoxia and its apparent symptoms, yet respiration rate is rarely if ever monitored during aviation hypoxia training, other than by visually observing the individual being trained.

The purpose of the present article is to describe a new method to monitor respiration when the ROBD is used to generate NH. The new method involves recording information that the ROBD currently acquires but discards. The new method does not require adding anything additional to the individual being trained, nor does it require any new instrumentation other than a relatively generic computer with a standard RS232 port to download the information from the ROBD.

METHODS

The present article reports new instrumentation and compares it with instrumentation that had been used to collect archived data from a previous study reported from our laboratory.¹⁹,²⁰ To set up and test the new instrumentation, including a data acquisition system that synchronizes multiple asynchronous data sources to a common time base, test data were obtained and compared to known standards and archived data. The U.S. Army Aeromedical Research Laboratory regulatory compliance office determined that the instrumentation effort reported here did not constitute research. The data collection procedures used to collect the archived data were reported previously.¹⁹,²⁰

**Instrumentation**

The ROBD is a commercial (Environics, Tolland, Connecticut), off-the-shelf, portable, computerized, gas-blending instrument that mixes tanked N₂ with tanked MSL air to produce breathable air with O₂ partial pressures comparable to those typically encountered at known altitudes.⁶,¹⁴,¹⁵,²¹,²² These air/N₂ mixes are precise and repeatable, and are used to safely induce NH without risk of barotrauma or decompression sickness associated with hypobaric chambers. The ROBD delivers the mixed air/N₂ gas by hose to the user’s breathing mask. The output port and the user’s mask are part of the breathing loop, which includes a built-in breathing loop pressure (BLP) transducer from which it is possible to record changes in BLP over time.

The ROBD has four modes of operation: Hypoxia Recognition Training (HRT), Flight Simulator Hypoxia Training, Oxygen System Failure Training, and the Positive Pressure Training modes. The present report applies only to the HRT mode; the other three modes are not considered further in this report. In the HRT modes, the ROBD is configured to maintain the BLP to approximately between 0 and 2 inches of water.

The rear panel of the ROBD includes a standard RS232 port that enables two-way communication with a computer that simultaneously controls the ROBD while recording the ROBD data output. The output data include (1) time, (2) current simulated altitude, (3) O₂ concentration, (4) BLP, (5) the elapsed time at the current simulated altitude, (6) the time remaining at the current simulated altitude, (7) the volunteer’s SpO₂, and (8) the volunteer’s pulse rate. These last two parameters are measured with the ROBD’s standard built-in pulse oximeter. These eight response parameters comprised the archived ROBD output dataset as well as the ROBD dataset acquired during the set up and test of new instrumentation.

The new instrumentation included a commercial (Biopac Systems, Inc., Goleta, California), off the shelf, separate system to measure respiration via thoracic chest movements as well as a system to measure the O₂ and CO₂ concentration of each expired breath. The system included an MP150 data acquisition hardware unit, the RPS100c respiration rate amplifier with Respiration Belt transducer, and the O2100C Oxygen and the CO2100C Carbon Dioxide measurement modules. The traces in Figure 2 were generated during integration, calibration, and validation of this equipment. These Biopac data were integrated and synchronized with the ROBD data.

The instrumentation goals included ensuring that the ROBD and the Biopac data were integrated accurately onto a data-harvesting computer synthesizing the separate data streams onto a common time base.

**Data**

**Acquired Data**

One of the research staff served as a one-time proxy to provide realistic physiological signals to drive the ROBD BLP and the Biopac sensors and recording instrumentation. That
single test of equipment, which was limited to MSL air (21% O₂), was the full extent of the new data collection reported here. The 8 fields of recorded ROBD data were exported at a rate of 3 Hz, a value selected based on previous experience with the ROBD. Biopac data were collected at 1,000 Hz and processed with the Biopac Acknowledge analysis software system and custom software. The entire procedure was limited to ensuring that the Biopack instrumentation and ROBD BLP instrumentation worked as intended. Since these acquired data are based on a single case, statistical comparisons are inappropriate.

Archived Data

The archived data were collected from military aviators in support of an approved study of the effects of hypoxia on human vision.¹⁹,²⁰ For that study, a data-harvesting computer running the data collection script recorded the exported ROBD data over the RS232 at a recording rate of 0.5 Hz. The eight data fields listed above comprised the exported dataset and were processed with MATLAB (MathWorks, Inc., Natick, Massachusetts) and Microsoft Excel (Microsoft Corp., Redmond, Washington). These archived data are from a study in which the subjects breathed either MSL air or a normobaric equivalent to the partial pressure of O₂ approximating 14,000 ft above MSL (13% O₂).¹⁹,²⁰ Thus, the ROBD was used for both simulated altitude conditions, and all testing was conducted in a ground level (301 ft/92 m) laboratory setting. The archived ROBD data, recorded at 0.5 Hz is adequate for measuring a normal breathing rate based on the Nyquist criterion, which indicates data should be recorded at a sampling rate that is at least twice as fast as the highest frequency of interest. Typical adult breathing rate is between 8 and 15 breaths per minute; values higher than 20 breaths per minute (tachypnea) are abnormal.

RESULTS

Typical BLP data recorded from one subject breathing 13% O₂ air or MSL air is shown in Figure 3. The abscissa displays time in terms of data point number, which, at 0.5 Hz, produces a 2-second interval between successive data points. The ordinate is BLP dimensioned as inches of H₂O. The MSL condition lasted about 38 minutes producing about 1,160 data points, whereas the 13% O₂ condition lasted about an hour producing about 1,800 points; this difference in duration is evident in the figure since the MSL data trace ends before the 13% O₂ data trace. Despite the time-scale compression, it can be seen that BLP recorded at MSL ranged from nearly 0 to about 1.5 inches H₂O, whereas BLP recorded during the 13% O₂ condition ranged from about 0 to about 1.1 inches H₂O. The difference in BLP range between the two conditions seems relatively constant over the duration of the study. The pressure of standard atmosphere is 10.3 m (33.8 ft) H₂O, so the observed changes in BLP were about 0.4% of standard atmospheric pressure.

Since the abscissa scaling makes it difficult to resolve fine patterns in the data, the data were magnified in Figure 4 by plotting the 30 data points that comprise minute 15 for both the MSL and the 13% O₂ trace. The figure shows a waveform consistent with a breathing pattern. Visual inspection shows that during this 1-minute period there were about 7 to 9 breaths for either condition. For the MSL condition, the average minimum BLP was about 0.38 inches H₂O, whereas the average maximum BLP was about 1.2 inches.
Minimum and maximum BLP for the 13% O2 condition were about 0.50 and 0.97 inches H2O, respectively. Thus, the peak-to-peak difference in BLP for the MSL condition is approximately twice that of the 13% O2 condition. This difference in pressure was found to be consistent, with the 13% O2 condition providing lower maximum pressures. This BLP difference between concentrations did not impact testing or the subjects’ experience; however, for further development of the possible applications of the BLP it is important to consider the coupling of BLP offset and modulation with the air/N2 ratio.

Although BLP can provide an intuitive sense of the data, it may be more useful to convert the BLP into breathing rate or average breaths per minute, by programing a peak counter to calculate the number of breaths within a given period of time. The BLP data displayed in Figure 3 were converted to breathing rate by averaging over a 2-minute wide sliding window translated over time in 0.5-minute steps. These average breaths per minute are displayed in Figure 5, which is a visualization of the data showing trends in breathing rate over time. Breathing rate for both conditions ranged between 3.5 and 14 breaths per minute, and both conditions average close to 8 breaths per minute, or 1 breath every 7.5 seconds. In addition, trends can be observed as systematic changes occurring over each condition. For example, the MSL condition demonstrated an increasing breathing rate trend between minutes 10 to 25 followed by a downward trend over the next 8 minutes.

**Acquired Data**

We are extending our laboratory capabilities by increasing the number of monitored physiological variables. One of the goals of this instrumentation is to record, import, and save several different simultaneous streams of data on a common time base. Because the HVR can impact CO2, we have included standard, commercial chest strap instrumentation for monitoring breathing rate via chest expansion as well as instrumentation to record expired O2 and CO2 on a breath-by-breath basis using the Biopac system.

The upper and lower panels of Figure 2 each display 1 minute of physiological recordings as a function of time on the abscissa. In both panels, the top tracing reflects breathing behavior measured with a chest-strap strain gauge transducer for thoracic expansion. The second trace is the ROBD BLP. The third trace displays expired O2, whereas the fourth trace is expired CO2. Thus, Figure 2 displays 2 different minutes of breathing behavior using the same four response variables.

In the upper panel of Figure 2, the tracings of the chest strap and BLP mirror each other. The chest strap measures the size of the thorax, whereas the BLP measures ROBD system pressure. During inhalation, the thorax expands, which is displayed as an increase in the chest strap trace, and pressure within the ROBD decreases as the lungs fill with air, corresponding to the decrease in BLP value. Conversely, when exhaling, the thorax gets smaller and ROBD system pressure increases. The inverse relationship between chest expansion and BLP in the upper panel corresponds with expected changes in O2 and CO2 with each breath, the lower pair of traces.

It may be noted that the chest strap, O2, and CO2 tracings were exported at 1,000 Hz, whereas the BLP was exported at 3 Hz. Thus, the synchronization of the data output was verified, one of the primary goals of this instrumentation effort. In addition, the new data acquisition system improved the recording rate for the BLP. The system had collected the archived data with a sampling rate of 0.5 Hz, whereas the newly installed system successfully records BLP at 3 Hz, which is adequate to cover all anticipated breathing rates.

More germane to the issue of the value of the BLP as an indicator of respiration is the consistent relationship between BLP and the other variables. In the upper panel of Figure 2, this relationship is consistent, although there are some phase shifts among the four tracings, shifts that reflect lags or latencies or differences in acquisition rate among the measures rather than errors in instrumentation. On the other hand, the chest strap data trace in the lower panel of Figure 2...
is essentially flat. This reflects the fact that the chest strap transducer had inadvertently slipped or loosened, compromising the quality of the data; however, the BLP was unaffected. Thus, BLP may be a more resilient measure of breathing behavior than is the chest strap.

The expired O2 and CO2 graph in the bottom of the lower panel of Figure 2 are also noteworthy. They show that the third complete expiration is nearly twice the duration of the first two expirations; the fourth through seventh expirations are about the same length as the first two; and the eighth expiration is nearly as long as the third, but with a slight double notch; the ninth breath is normal again; and the tenth is truncated but begins its rise time normally. The periods of prolonged expiration occurred when the laboratory staff member intentionally perturbed the signal by speaking aloud a string of numbers. These tracings clearly show that the person was speaking, inhaled four times, spoke again, and then inhaled twice. During the second speaking period, slight catch breathes were evident in the BLP as well as in the expired gas tracings. Much of this information was lost in the chest strap tracing, even though the chest strap is designed specifically to monitor breathing behavior.

**DISCUSSION**

Archived data obtained during a study of the effects of NH on vision and visual performance\textsuperscript{19,20} were reanalyzed to examine BLP. The motivation for this evaluation of BLP derived from an appreciation of the paucity of breathing rate information associated with the large range of SpO2 values typically reported in the hypoxia literature.\textsuperscript{16} The possibility that the HVR could help explain the large SpO2 range (Fig. 1) suggested that the BLP might be a cost-effective way of monitoring breathing rate when hypoxia is generated with the ROBD. Furthermore, if hyperventilation is indeed a contributing factor to increases in SpO2, increased breathing rate could be important for identifying potential cognitive problems associated with hypocapnia, CO2 blood saturation, pH balance, and cerebral blood flow.

The fundamental conclusion of the reanalysis was that the BLP data were reliable for breathing rates of 8 to 15 breaths per minute. Although the 0.5 Hz sampling rate is adequate for normal breathing rates, this sampling rate is inadequate for accurately measuring tachypnea, breathing rates above 20 breaths per minute. The primary motivation for evaluating the BLP was to identify changes in breathing rate that might explain individual differences in SpO2 values, but hyperventilation at altitude could become tachypneic. Because of the marginal sampling rate, further analyses with these archived data seemed moot.

On the other hand, the data set shown in Figures 3, 4, and 5 was typical of the data obtained from all the subjects,\textsuperscript{19,20} which strongly suggests that the BLP is potentially an informative, useful measure of respiration. This is a particularly important point because all ROBDs currently generate the BLP and make it available for export, yet historically these data have not been used to objectively monitor breathing. The quality of the archived BLP data seems extraordinary since these data were all incidental and were not collected for analysis.

The Biopac data reported here are also incidental; they were not intended as data, but merely as a guide for assembling laboratory capabilities and ensuring that the equipment was functioning as expected. These data were not to be used in any systematic fashion, which makes the resilience and orderliness found in the O2 and CO2 data all the more impressive. The chest strap data were less robust than the BLP data, but to be fair, there were no efforts made to optimize any of these measurements; therefore, it is possible that the chest strap could be implemented with robustness comparable to the BLP. Regardless, the chest strap is a separate measure that requires placement on the individual’s body, as well as its own preparation, implementation, methodology, cost, and integration. The ROBD can eliminate all that additional overhead. Collecting BLP data requires no additional instrumentation; a generic Windows computer with an RS232 port is sufficient. The software scripts necessary to extract the data from the ROBD were provided by Environics and documented in the ROBD programming manual.\textsuperscript{15}

The ROBD manual describes the BLP as intended to remain between 0 and 2 inches of H2O during the HRT mode, the mode relevant to the present report. The characteristics of BLP over time and under different conditions have yet to be determined. For example, the amplitude modulation and the mean value of the BLP reported here are different between the MSL air and 13% O2, with the latter producing a lower mean and a smaller range BLP. Thus the BLP is coupled with the air/N2 mix. This dependence of BLP mean and modulation on the air/N2 mix was the reason that the analysis reported here was limited to the assessment of breathing rate, which required the detection of rising and falling edges.

A next step in this work could be to explore the repeatability and stability of BLP as a dependent measure. The Biopac respiration monitoring equipment and the new data acquisition system would allow this work to be carried further. The improvement in sampling rate from 0.5 to 3 Hz for the BLP ensures that all potential breathing rates likely to be encountered will be adequately sampled. It is possible that given sufficient justification and importance, the BLP functionality could be enhanced by ensuring stability over time and comparability of different air/N2 mixes. Additional improvement to the usefulness of the BLP for monitoring respiration with the ROBD could include the development of specific software analysis applications and displays that enable the near real-time evaluation of breathing behavior so that the ROBD might monitor and display respiration rate, respiration frequency, spectra, tidal volume, and minute ventilation in a way comparable to the ROBD’s current display of pulse oximetry. It may also be possible to increase data...
sampling rates high enough to evaluate breathing rate variability on a breath-by-breath basis. Such added functionality may be particularly important for evaluating, controlling, and reducing hyperventilation and its effects, thereby improving ROBD effectiveness as a research and training tool for hypoxic hypoxia.

ACKNOWLEDGMENTS
We would like to thank Ms. Jessica Cumbee for her editorial contributions to this manuscript. This work was supported by an award from the U.S. Army Medical Research and Materiel Command, Combat Casualty Care, Neuro Trauma Research Program, for the project titled Uncovering Latent Deficiencies Due to Mild Traumatic Brain Injury (mTBI) by using Normobaric Hypoxic Stress.

REFERENCES
Ultrasound Versus Landmarks for Great Toe Arthrocentesis

MAJ Jason F. Naylor, SP USA; MAJ Kenneth B. Dekay, MC USA; MAJ Benjamin P. Donham, MC USA; LTC Brian T. Hall, MC USA

ABSTRACT
Background: Several studies have demonstrated ultrasound (US) is superior to traditional landmark (LM)-based techniques for large and medium joint aspiration; however, no studies of sufficient size have evaluated these interventions in the smaller toe joints. The purpose of this study was to determine if US provides an advantage over LM for successful first-pass aspiration of first metatarsophalangeal joint (1st MTPJ) effusions. Methods: A crossover, cadaveric trial evaluating the interventions of US and LM. Eighteen emergency medicine residents performed four US and four LM aspirations each of 1st MTPJ effusions simulated in fresh-frozen cadavers. The initial intervention utilized was randomized. The primary outcome measured was aspiration success or failure. A secondary outcome measured was time in seconds taken to complete a successful aspiration. Results: A total of 144 1st MTPJ aspirations were attempted—72 by US and 72 by LM. US was the initial intervention used in 9 of 18 (50%) participants. Fifty-seven of 72 (79.2%) US attempts were successful, while 53 of 72 (73.6%) LM attempts were successful (95% confidence interval 69.5%, 83.3%; p = 0.36). Successful US aspirations took 43.7 seconds (±10.0), whereas successful LM aspirations averaged 34.0 seconds (±24.3). The mean difference in time to successful aspiration was 9.7 seconds (95% confidence interval 20.3, −0.9; p = 0.07). There was no statistically significant difference in success and time between US and LM. Conclusion: In this study, US did not prove superior to LM for first-pass aspiration of 1st MTPJ effusions.

INTRODUCTION
Septic arthritis is an orthopedic emergency with 10 to 25% mortality and 25 to 50% morbidity rates.1 It afflicts 2 to 5 adults, 5.5 to 12 children, 28 to 38 patients with rheumatoid arthritis, and 40 to 68 individuals with joint prostheses per 100,000 persons annually.2 Two to five percent of adult septic arthritis occurs in the joints of the toes.1–5

Arthrocentesis is mandatory if septic arthritis is a diagnostic consideration. Imaging and laboratory testing outside of synovial fluid analysis cannot rule out joint infection.6–9 If septic arthritis is diagnosed, gram stain and culture results of joint aspirate guide antimicrobial therapies. Alternative diagnoses, such as gout and pseudogout, are also based on synovial fluid analysis. Successful arthrocentesis, however, can be challenging in small joints.

Landmark (LM)-based aspiration of the first metatarsophalangeal joint (1st MTPJ) has a success rate ranging from 81 to 91%.10–12 Successful aspiration was defined as sufficient synovial fluid aspirated to allow crystal analysis. Failed arthrocentesis may have resulted from extra-articular needle placements or small intra-articular volumes typical of toe joints.13 Manadan et al. surveyed 10 rheumatologists with a mean of 17.9 years (range, 3–32 years) of clinical experience and the self-reported success rate for 1st MTPJ arthrocentesis was 37% (range, 0–80%).14 In a pilot study, Balint et al. compared ultrasound (US) to LM for aspiration of several joints; however, at most only 1 MTPJ was attempted by LM and 4 by US.15 Their success rates were 0% (0/1) for LM and 100% (4/4) for US.

No studies of sufficient size have compared US to LM for aspiration of 1st MTPJ effusions. However, studies have demonstrated US is superior to LM for injection of toe joints and aspiration of medium and large joint effusions.16–20 Improved aspiration success rates provide benefits in terms of cost, patient comfort, and clinical outcome. The purpose of this study was to determine if US provides an advantage over LM for successful first-pass aspiration of 1st MTPJ effusions.

METHODS
Study Oversight
Local Institutional Review Board approval was obtained for this cross-over, cadaveric trial. All study subjects were donated for and consented to medical research, and all study participants voluntarily partook in this investigation. All authors have no conflicts to report.

Subject Selection
Study subjects were fresh-frozen cadavers. The inclusion criteria were a donated cadaver and intact feet. The exclusion criteria were 1st MTPJ fusion and less than 30 years of age. Thirty years of age was selected based on reported ages in previous studies and prevalence rates of gout in the United States.10–12,15,21

Study Participants
Study participants were residents from two emergency medicine residency programs. The exclusion criterion was previous experience aspirating five or more 1st MTPJ effusions.
A 7-minute block of instruction on US and LM-based aspiration techniques was given to all study participants before aspiration attempts; however, they were not allowed to practice either technique.

**Study Design**

Each participant attempted eight total 1st MTPJ effusion aspirations, four with US guidance and four by LM. Participants were randomized to the first intervention attempted; however, after the first attempt participants were free to choose which intervention to perform, as long as on each cadaver one joint was attempted with US and the other joint by LM. Participants were allowed only one attempt per joint, which started when the needle entered the skin and ended once the needle was withdrawn from the skin or when the time limit of 180 seconds was reached.

Aspirations were performed without assistance and attempted with a 22 gauge × 1.5 inch hypodermic needle attached to a graduated 1-mL syringe. All participants were required to perform US attempts with active guidance using a Sonosite M-Turbo Portable US Machine and 13-6 megahertz linear transducer in the long axis without standoff and with needle entry from a dorsomedial to medial approach. For LM attempts, participants were instructed to hold and manipulate the great toe (flex 15–30° and apply distal traction) with dorsomedial to dorsal needle entry. Before starting an attempt, participants were allowed to prepare equipment by attaching the hypodermic needle to the syringe and applying acoustic coupling gel to the surface of the transducer. Participants, however, were not permitted to identify LMs or needle insertion sites prior to time starting.

Study investigators simulated 1st MTPJ effusions by injecting 2–3 mL of a mixture of methylene blue dye and saline into each joint immediately before study attempts. The dorsolateral approach was utilized to preserve the dorsomedial to medial aspects for study participants. Investigators used US to confirm simulated joint effusions before participant attempts. Investigators flushed needle and syringe assemblies with clear saline in between each attempt to ensure no dye mixture was retained from a previous attempt.

**Outcome Measures and Definitions**

The primary outcome of the study was success. An attempt was considered complete once the needle tip was withdrawn from the skin or the maximum allotted time, 180 seconds, elapsed. Successful aspiration was defined as the presence of dye mixture inside the aspiration syringe regardless of volume and the procedure was completed within the allotted time. Failed aspiration was defined as no dye mixture within the syringe or exceeding the given time limit to perform the procedure.

A secondary outcome measured in the study was time taken to complete a successful aspiration. Time was measured in seconds from 0 to 180 with a stopwatch. Time started when the study investigator stated, “You may begin.” Time ended when the needle tip was withdrawn from the skin or when the maximum time was reached.

**Statistical Analysis**

All statistical tests were two-tailed and based on an alpha of 5%, a beta of 20%, and a level of confidence of 95%. IBM SPSS Sample Power 3.0 was used to estimate the sample size. For the outcome of success, the US success rate was estimated at 90%. A clinically significant difference was defined as a 20% decrease in the LM group to 70%. To detect a 90 to 70% difference required 62 aspirations per intervention for 124 total aspiration attempts.

For the outcome of time, a 30- to 120-second range was estimated as the time needed to complete the task with US and this was assumed to be the 95% confidence interval (CI). Consequently, the estimated mean was 75 seconds with a standard deviation of 22.5 seconds. A clinically significant difference was calculated as 105 seconds, with a corresponding effect size of 4.7 standard deviations. To detect a 4.7 standard deviation difference required five observations per intervention for 10 total successful aspirations.

IBM SPSS Statistical Analysis Package 3.0 was utilized to analyze the collected data. Outcomes for success were assessed with the Pearson $\chi^2$ test with continuity correction, while results for time were evaluated with the Mann–Whitney Rank Sum test to compensate for likely disproportionate influence of outlier data.

**RESULTS**

**Study Subjects and Participants**

A total of eight fresh-frozen cadavers were available for the study. All eight cadavers met inclusion criteria. Four cadavers were randomly selected and zero were excluded. Three cadavers were male and one was female with an average age of 72.25 years (range, 53–87 years) (Table I).

A total of 34 emergency medicine residents volunteered to participate in the study. Eighteen were randomly selected and zero were excluded. Eight participants were in their third year of residency, seven in their second year of residency, and three were in their first year of residency. Twelve of 18 (67%) participants had never attempted an aspiration of the 1st MTPJ before the study (Table I). Of the six who had prior experience, two previous attempts were the most reported.

**Aspiration Success**

A total of 144 joint aspirations were attempted, 72 by US and 72 by LM. The primary outcome measured was success. The success rate with US was 79.2% (57 of 72), while the success rate with LM was 73.6% (53 of 72) (Figure 1). There was no statistically significant difference in success rates between interventions (95% CI 69.5%, 83.3%; $p = 0.56$).
Time for Successful Aspiration

A secondary outcome of time to complete a successful aspiration attempt was measured in seconds. A total of 110 of 144 aspiration attempts were successful, 57 by US and 53 by LM. The mean ± standard deviation of the time to completion was 43.7 ± 31.0 seconds for US and 34.0 ± 24.3 seconds for LM (Figure 2). The mean difference in time to successful aspiration was 9.7 seconds (95% CI 20.3, −0.9; \( p = 0.07 \)). Although Levene’s test was insignificant for variance, a more rigorous nonparametric test was utilized and confirmed the results were not statistically significant.

DISCUSSION

Aspiration Success

The results of this study indicate that US and LM are equally effective for successful first-pass aspiration of 1st MTPJ effusions. The success rate with US of 79.2% was not significantly better than the LM rate of 73.6%. Compared to previous studies, the outcome and success rates for each intervention were unexpected.

US was superior to LM in multiple studies for aspiration of medium and large joint effusions.\(^{15,19,20}\) That was not the case in this study and may be explained by the 1st MTPJ’s capacity to be mechanically distracted by manipulating the great toe.\(^{22}\) Landmark attempts were made while flexing and applying distal traction to the great toe with the free hand. Joint manipulation, on the other hand, was not possible with US since unassisted participants were required to use active US guidance by holding the transducer during the procedure. US guidance, not US assistance (visualization of the target image before the procedure, but no active guidance), was considered to be the typical clinical application and thus utilized.\(^{20}\)

US was superior to LM in a study for injection of small joints.\(^{16}\) No significant difference between US and LM in this study may be due to the previous study utilizing finger joints, while this study utilized the relatively larger 1st MTPJ. Since toe and finger joints are superficial from a dorsal approach and amenable to some degree of manipulation, the larger size of the 1st MTPJ may have enabled improved rates with the LM technique. Other, smaller, MTPJs were not used in this study due to the 1st MTPJ’s clinical significance as the most commonly afflicted joint in gout, which mimics septic arthritis.\(^{23}\)

Previous studies of the 1st MTPJ were performed by experienced rheumatologists.\(^{10-12,14,15}\) As a specialty, rheumatologists periodically, if not routinely, perform small joint injections and aspirations.\(^{9}\) The participants of this study were emergency medicine residents, most of whom (67%) had never performed a 1st MTPJ aspiration. The annual rate of 1st MTPJ aspirations performed by emergency physicians is not known, but is expected to be low. Thus, residents with little experience may be representative of the majority of clinicians that perform 1st MTPJ aspirations in acute care settings.

Previous studies that compared US to LM for arthrocentesis reported multiple attempts or did not explicitly state single attempts.\(^{15,19,21}\) In this study, participants were limited to a single attempt that was terminated when the needle was withdrawn from the skin or the maximum time limit was reached. No attempts were stopped for time. Multiple attempts by both methods, however, were terminated when the participant inadvertently withdrew the needle from the skin while trying to redirect it within the few millimeters of

![Figure 1](image1.png)  
**FIGURE 1.** Results for Aspiration Success.

---

**TABLE I.** Characteristics of Subjects and Participants.

<table>
<thead>
<tr>
<th>Study Subjects</th>
<th>Gender</th>
<th>Age (Years)</th>
<th>Average Age (Years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Male</td>
<td>53</td>
<td>72.25</td>
<td></td>
</tr>
<tr>
<td>2 Female</td>
<td>70</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Male</td>
<td>79</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 Male</td>
<td>87</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Study Participants</th>
<th>Post Grad Year</th>
<th>Number (% Total)</th>
<th>Previous Experience (% Group)</th>
<th>No. of Previous Attempts (Average)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3 (16.7)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>7 (38.9)</td>
<td>1 (14.3)</td>
<td>2 (2)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>8 (44.4)</td>
<td>5 (62.5)</td>
<td>9 (1.8)</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>18</td>
<td>6 (33.3)</td>
<td>11</td>
<td></td>
</tr>
</tbody>
</table>
skin overlying the joint. This occurred more often with US since the participant was looking at the US machine’s screen and not the procedure site while manipulating the needle. Multiple punctures, however, were not permitted because minimal patient discomfort was a desired condition.

Anatomical differences between cadavers and living subjects may have influenced this study’s outcome. All of the previous studies that evaluated aspiration of the 1st MTPJ, regardless of the technique employed, were performed on living subjects.10–12,15 Despite the inherent limitations of cadaver models, they are considered superior to nonorganic replicas for anatomic likeness and procedural training, especially the fresh-frozen cadavers utilized in this study (Figure 3).17,24 Cadavers were selected as study subjects, however, in order to generate the number of exposures required to obtain sufficient power for statistical analysis. Incidentally, none of the previously mentioned studies obtained sufficient power as indicated by omitted p values for all but 1 study that reported a p value greater than 0.05.11

The outcome for success may also be explained by the value used to statistically detect what was considered a clinically meaningful difference. A difference of 20% between US and LM was arbitrarily selected by the investigators. If a 10% difference, on the other hand, was thought to be more appropriate then the outcome may have differed, but detecting it with an average probability derived from this study’s success rates—76.4% from 79.2% for US and 73.6% for LM—would require a total sample of 564. A 20% difference, however, was selected based on investigator opinion that clinicians would consider modifying their practice if one of the techniques improved aspiration success 2 out of every 10 attempts, as opposed to only 1 out of 10—especially for a procedure that is performed infrequently.

The success rates for both US and LM in this study were below the rates previously reported. This study’s US success rate was 79.2% and LM was 73.6%. Previous studies reported a success rate for US of 100% and for LM 81 to 91%.10–12,15 Again, these studies either reported multiple attempts or did not explicitly state single attempts. They were performed on living subjects by experienced rheumatologists and the US success rate was based on a maximum of four 1st MTPJ attempts. Therefore, the success rates in this study may be a better reflection of the actual rates of successfully aspirating a 1st MTPJ effusion on the first attempt when performed by a provider who does not frequently perform the procedure.

**Time for Successful Aspiration**

The results of this study indicate that US and LM perform equally well with respect to time taken to successfully aspirate 1st MTPJ effusions. Fifty-seven successful US attempts took an average of 43.7 seconds, whereas 53 successful LM aspirations took 34.0 seconds. Again, the result of the secondary outcome measured in this study does not reflect the results of previous studies.

Only two studies comparing US to LM for arthrocentesis evaluated performance in terms of time. Time was a secondary outcome in an emergency medicine study that compared US to LM for successful aspiration of knee joint effusions in patients.25 This study found US was superior to LM with respect to time; however, they acknowledged that time was not measured. Instead, the emergency physician who performed the procedure was later asked to estimate how much time it took to perform the procedure. Consequently, the results of this particular study with respect to time are questionable. Time was the primary outcome measured in a cadaveric study comparing US to LM for aspiration of medium-sized joints (elbow, wrist, and ankle), and it found US was superior to LM.26 The 1st MTPJ is superficial about the dorsal aspect of the foot and easily palpated, especially with manipulation of the big toe. The outcome of this study may be explained by the relative ease, thus speed, with which the 1st MTPJ is identified in comparison to the medium-sized joints used in the aforementioned study.

The results of this study indicate that US and LM are equally efficacious for first-pass aspiration of 1st MTPJ effusions. Clinically, providers with limited or remote experience aspirating 1st MTPJ effusions may expect to have similar rates of success and times for completion with either US or LM. The equivalency of both techniques is particularly relevant in clinical settings without access to US machines.
Most emergency departments, however, have US equipment readily available. In these settings, providers may employ US for clinical benefits outside of aspiration success and speed. US has been demonstrated to be superior to palpation for detection of joint effusions, sonographic findings indicative of specific pathologic states of the joint and adjacent structures can inform clinical decision-making, and US visualization of needle placement within a joint effusion may obviate multiple attempts with an apparent dry tap. \(^ {27-29}\)

**CONCLUSION**

In this study, US did not prove superior to LM for first-pass aspiration of 1st MTPJ effusions.

---

![Ultrasound Versus Landmarks for Great Toe Arthrocentesis](image)

**FIGURE 3.** (A) Shows the Sonographic Appearance of the First Metatarsophalangeal Joint (MTPJ) Live Human, and No Effusion. (B) Shows the Sonographic Appearance of the First MTPJ, Cadaver, and No Effusion. (C) Shows the Sonographic Appearance of the First MTPJ, Cadaver, and with Effusion.

---

**REFERENCES**

The Efficacy of Two Trabecular Bypass Stents Compared to One in the Management of Open-Angle Glaucoma

CPT Anton Vlasov, MC USA; LTC Won I. Kim, MC USA

ABSTRACT  Objectives: To compare the outcomes of combined cataract surgery with two trabecular microbypass stents compared to one with patients in open-angle glaucoma. Methods: Patients with primary open-angle glaucoma were included. Primary outcome measures were intraocular pressure (IOP), postoperative medications, and postoperative adverse events. Results: The average patient age was 73.2 years. Thirty-nine eyes had phacoemulsification and one stent implanted (Group 1); 30 eyes had phacoemulsification and two stents implanted (Group 2). A significant reduction in mean IOP from baseline to 12 months was noted for both Group 1 (16.67 ± 4.1 mm Hg to 14.45 ± 3.8 mm Hg; p < 0.0251) and Group 2 (18.33 ± 3.99 mm Hg to 14.31 ± 1.8 mm Hg; p < 0.0014). Group 2 had a greater percent decrease in IOP at 12 months than Group 1, but this difference was not statistically significant. Only Group 2 demonstrated a statistically significant reduction in medication burden at 12 months. There were no serious complications in either group thought to be caused by the microbypass stents. Conclusions: Both groups demonstrated a significant reduction in IOP at 12 months. Group 2 showed a greater percent decrease in IOP; however, it was not statistically significant. Only Group 2 demonstrated a statistically significant reduction in medication burden.

INTRODUCTION
Glaucoma is the second leading cause of blindness worldwide.1 Currently, the only treatment is intraocular pressure (IOP) reduction through pharmacotherapy, laser trabeculoplasty, or incisional glaucoma surgery.1–3 Increased resistance to aqueous humor outflow at the level of the trabecular meshwork is the principal mechanism of elevated IOP. Trabeculectomy is the most commonly performed incisional glaucoma surgery, introduced by Cairns in the 1960s. This is a fistulization procedure that creates a new pathway for aqueous humor to exit the anterior chamber of the eye into the subconjunctival space, creating a filtering bleb. Although very effective at reducing IOP, it carries a very high risk for complications, some of which are sight threatening. Risks include hypotony, choroidal effusions, choroidal hemorrhage, maculopathy, and infections such as blebitis and endophthalmitis.3–5 The complication rate can be as high as 60% at 3 years, and the risk of developing an infection has been reported to be about 1% per year for a functioning bleb.6,7 It is this persistent risk of infection, which can occur years and even decades after the initial surgery, that is most worrisome from the perspective of active duty (AD) personnel. Any bleb-related infection is vision threatening and requires immediate and aggressive treatment by an ophthalmologist. A service member deployed to an austere environment with such an infection may be potentially delayed from reaching a deployed ophthalmologist who sometimes may be located hundreds of miles away. Such a situation could lead to a devastating outcome and as such, service members who receive trabeculectomy surgery are often put on a nondeployable status permanently, sometimes altering the course of their military career. In addition, about 10% of trabeculectomies will fail every year with the overall failure rate approaching 50% at 5 years postsurgery.8

Recently, microinvasive glaucoma surgery has been introduced as an alternative surgical treatment option for patients with mild to moderate glaucoma.9 Microinvasive glaucoma surgery procedures have been shown to have an excellent safety profile with very few complications.9 These procedures include Trabectome (Neomedix, Tustin, California), endoscopic cyclophotocoagulation (Beaver-Visitec Endo-Optiks, Waltham, Massachusetts), and iStent (Glaukos, Laguna Hills, California). The iStent is a titanium shunt that allows aqueous humor from the anterior chamber to bypass the resistance at the level of trabecular meshwork and freely enter Schlemm’s canal.9 Microinvasive glaucoma surgery has the potential to minimize many of the risks of trabeculectomy surgery as described above. Currently, in the United States, only one iStent is approved for placement per eye in conjunction with cataract surgery.9 However, theoretically, implanting more than one iStent has the advantage of allowing more aqueous humor to bypass the trabecular meshwork and access more of Schlemm’s canal and its associated collector channels, potentially further lowering IOP and improving the chances of success. In the Food and Drug Administration trials, there were no significant complications attributed directly to the iStent. Because of this remarkable safety profile, it is likely that implanting more than one device per eye poses little to no added risk to the patient while potentially improving the chances of success. The placement of multiple iStents has been looked at in Canadian studies where multiple implant placement is approved with promising results.10 Inability to be reimbursed for a second iStent has limited this practice in the United States. In the military...
In an "off label" manner if it is safe and beneficial to the patient. The results of the placement of multiple iStents have not been studied in U.S. military treatment facilities by AD physicians on AD service members and beneficiaries. Because of the aforementioned benefits of this form of glaucoma surgery to the AD population, it is crucial to ascertain its effectiveness in our patient population as it is performed by AD ophthalmologists.

**METHODS**

We conducted a retrospective interventional comparative case series review of surgical results of one versus two iStent implantations in conjunction with cataract surgery performed between June 2013 and June 2015. Institutional review board approval was obtained. Included were patients with primary open-angle glaucoma, pseudoexfoliation, and pigmentary dispersion glaucoma at any stage of disease severity and visually significant cataract.

Primary outcome measures were IOP and number of glaucoma medications used, and adverse events at postoperative visits at day 1, week 1, months 1, 3, 6, and 12 after surgery. Both the IOP and the number of medications used were recorded at baseline and at day 1, week 1, week 4, month 3, and month 6 postoperatively. Glaucoma medication use was calculated by adding up each individual medication the patients were on; e.g., if a patient was taking latanoprost eye drops, as well as a combination dorzolamide/timolol drops, the total number of glaucoma medications would be 3.

The preoperative and postoperative IOP values were compared between the two groups, using a paired-sample $t$ test. Statistical analysis was done using the GraphPad software, a web-based statistical tool (Graphpad software; La Jolla, California). The IOP change between visits was compared using Wilcoxon signed rank test. The limit of statistical significance was set at $p \leq 0.05$. All surgeries were performed by five ophthalmologists at the Walter Reed National Military Medical Center, Fort Belvoir Community Hospital, Malcolm Grow Medical Center, and Kimbrough Ambulatory Care Center. Of the ophthalmologists performing these surgeries, three were glaucoma fellowship trained ophthalmologists, whereas the other two were comprehensive ophthalmologists.

**RESULTS**

There were 69 patients that met the inclusion criteria. Thirty-nine patients underwent a single iStent implant (Group 1), whereas 30 patients received two iStents (Group 2). The average age was 73.2 years (range, 52–92), and 60.9% were male; patient demographics are outlined in Table I.

Effects of intervention on IOP and medication use are described in Table II. There was a statistically significant decrease in the IOP between the preoperative measurements and the postoperative values measured at postoperative month 1, month 3, month 6, and month 12 for both groups. Adverse events are outlined in Table III. For all adverse events, there was no statistical difference between the two groups at any postoperative timepoint.

**TABLE I.** Study Demographics

<table>
<thead>
<tr>
<th></th>
<th>One iStent</th>
<th>Two iStents</th>
<th>$p$ Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>74.23 ± 10.2</td>
<td>70.26 ± 9.64</td>
<td>0.0974</td>
</tr>
<tr>
<td>Male</td>
<td>25</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>14</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>Pre-Op IOP</td>
<td>16.67 ± 4.1</td>
<td>18.33 ± 3.99</td>
<td>0.0870</td>
</tr>
<tr>
<td>Pre-Op Medications</td>
<td>2.33 ± 1.4</td>
<td>2.37 ± 1.50</td>
<td>0.9205</td>
</tr>
<tr>
<td>Pre-Op log MAR BCVA</td>
<td>0.32 ± 0.23</td>
<td>0.38 ± 0.25</td>
<td>0.7484</td>
</tr>
<tr>
<td>Post-Op Month 12 IOP</td>
<td>14.45 ± 3.8</td>
<td>14.31 ± 1.72</td>
<td>0.9051</td>
</tr>
<tr>
<td>Post-Op Month 12 Medications</td>
<td>1.74 ± 2.47</td>
<td>1.15 ± 1.10</td>
<td>0.4305</td>
</tr>
</tbody>
</table>

**TABLE II.** Intraocular Pressure and Medications in One iStent vs. Two iStents Cohorts

<table>
<thead>
<tr>
<th></th>
<th>1 vs. 2 Valid $N$</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>$p$ Value Pre-Op vs.</th>
<th>$p$ Value 1 vs. 2 iStent</th>
</tr>
</thead>
<tbody>
<tr>
<td>IOP Pre-Op</td>
<td>39</td>
<td>16.67</td>
<td>3.82</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>IOP POD 1</td>
<td>35</td>
<td>20.17</td>
<td>7.44</td>
<td>0.0128</td>
<td></td>
</tr>
<tr>
<td>IOP POW 1</td>
<td>34</td>
<td>16.78</td>
<td>5.23</td>
<td>0.917</td>
<td></td>
</tr>
<tr>
<td>IOP POM 1</td>
<td>34</td>
<td>14.76</td>
<td>3.77</td>
<td>0.0389</td>
<td></td>
</tr>
<tr>
<td>IOP POM 3</td>
<td>27</td>
<td>14.74</td>
<td>4.77</td>
<td>0.0755</td>
<td></td>
</tr>
<tr>
<td>IOP POM 6</td>
<td>33</td>
<td>14.44</td>
<td>4.27</td>
<td>0.0233</td>
<td></td>
</tr>
<tr>
<td>IOP POM 12</td>
<td>28</td>
<td>14.45</td>
<td>3.96</td>
<td>0.0251</td>
<td></td>
</tr>
<tr>
<td>Med Pre-Op</td>
<td>39</td>
<td>2.33</td>
<td>1.40</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>Med POD 1</td>
<td>35</td>
<td>0.91</td>
<td>1.44</td>
<td>0.0001</td>
<td></td>
</tr>
<tr>
<td>Med POD 1</td>
<td>34</td>
<td>1.29</td>
<td>1.34</td>
<td>0.002</td>
<td></td>
</tr>
<tr>
<td>Med POD 3</td>
<td>34</td>
<td>1.74</td>
<td>2.58</td>
<td>0.2237</td>
<td></td>
</tr>
<tr>
<td>Med POD 3</td>
<td>27</td>
<td>1.59</td>
<td>2.48</td>
<td>0.1324</td>
<td></td>
</tr>
<tr>
<td>Med POD 6</td>
<td>33</td>
<td>1.91</td>
<td>2.35</td>
<td>0.3528</td>
<td></td>
</tr>
<tr>
<td>Med POD 12</td>
<td>28</td>
<td>1.74</td>
<td>2.47</td>
<td>0.2259</td>
<td></td>
</tr>
</tbody>
</table>

IOP, intraocular pressure; POD, Post-Operative Day; POM, Post-Operative Month; POW, Post-Operative Week.
groups. Glaucoma medication number was also reduced in both groups; however, only the two iStent group had a statistically significant decrease from preoperative visit to postoperative month 12 visit. \( p = 0.0066 \). The overall IOP reduction was 13.3% with one iStent implanted and 21.9% in patients with two iStents implanted at 12 months; however, this difference in percent IOP reduction was not statistically significant between the two groups \( p = 0.9051 \). The medication burden reduction was 25% for Group 1 and 48% for Group 2. The reduction was only statistically significant for Group 2.

There were no complications noted during iStent implantation. Four patients in Group 1 developed cystoid macular edema. Additionally, two patients in Group 1 developed increased IOP because of a steroid response, which was treated with topical antihypertensive medications. Finally, one patient in Group 1 suffered a central retinal vein occlusion, which later led to development of anterior chamber angle neovascularization and neovascular glaucoma. There were no complications in Group 2.

DISCUSSION
The recently Food and Drug Administration–approved trabecular bypass device, iStent, in conjunction with cataract surgery has shown promising results in lowering IOP and reducing medication burden in patients with concomitant open-angle glaucoma and cataract. We compared implanting two iStents versus one, and the effect on the IOP and the number of medications required. The data analysis showed that iStent implantation led to a statistically significant IOP lowering between preoperative values and postoperative month 12 in both groups. There was a trend toward greater IOP reduction in the two iStent group (21.9% vs. 13.3%), but this did not reach statistical significance. We did find the two iStent group to show a statistically significant reduction in medication burden.

Medication burden was assessed by the number of individual medications the patient was taking. Although, the authors did not use a questionnaire to specifically evaluate this, other studies have shown that medication compliance decreases with the increase in number of medications a patient has to use.\(^\text{10}\) Noncompliance with antiglaucoma medications can inadvertently lead to glaucoma progression. Figures 1 and 2 show the IOP trend and the medication requirements during the study period.

Some of the shortcomings of our study include a relatively small population size. With a small population size, patients in the study may not truly reflect a general population and the results may not be widely generalizable. Another disadvantage is the study being a retrospective chart review, selection bias is not eliminated, as well as not being able to control for some confounding variables.

Baseline IOP was higher in patients who received two iStents. Although it was not statistically significant, this implies that patients that received two iStents had glaucoma that may not have been as well controlled before surgery as in the patients that received one iStent. The surgeries were performed by multiple surgeons, all with varying degrees of experience in performing these surgeries. There is a learning curve to the adoption of this procedure and the variability in experience among the surgeons may also be a confounding variable. Surgeons who were glaucoma fellowship trained, generally had a lower final IOP when compared with comprehensive ophthalmologists.

CONCLUSIONS
The iStent has shown promising results in the treatment of open-angle glaucoma in patients with visually significant cataract. Our study was the first U.S.-based study to compare the results of one versus two iStent implantations. The studies that preceded ours looking at the implantation of multiple
iStents were performed in either Canada or Europe. Our study confirmed that the implantation of both iStents significantly lowers IOP. Use of multiple iStents in our study led to a decrease in medication burden which corroborates the findings of a previously published study. Decrease in medication burden, leads to better medication adherence, which may prevent glaucoma progression. Although our retrospective chart review did not show a statistically significant advantage in IOP lowering when more than one iStent was used, there was a trend suggesting superiority for multiple iStents over single iStent placement. Our study was consistent with others in confirming that iStent implantation is extremely safe in comparison with traditional glaucoma surgeries such as trabeculectomy. No serious, vision-threatening complications were seen in our study that was directly attributable to the iStent. These characteristics are especially useful in the AD population, as traditional glaucoma filtration surgery has a lifelong risk of vision-threatening infection, which can affect a service member’s deployment status. Using iStents may allow for control of a service member’s glaucomatous disease without limiting his or her military career and eliminate the risk of some long-term vision-threatening complications. Even though our population consisted of mostly older, retired beneficiaries, similar results would be expected if an iStent was implanted in a younger, AD service member with glaucoma. Future studies looking at the efficacy of this type of surgery in the age group of our AD population will be needed. Larger population size and longer term studies are also needed to obtain more definitive results.

REFERENCES

Blast Wave Dynamics at the Cornea as a Function of Eye Protection Form and Fit

Steven T. Williams, MS*†; Thomas H. Harding, PhD*; J. Keegan Statz, MS*‡; John S. Martin*

ABSTRACT A shock tube and anthropomorphic headforms were used to investigate eye protection form and fit using eyewear on the Authorized Protective Eyewear List in primary ocular blast trauma experiments. Time pressure recordings were obtained from highly linear pressure sensors mounted at the cornea of instrumented headforms of different sizes. A custom shock tube produced highly reliable shock waves and pressure recordings were collected as a function of shock wave orientation and protective eyewear. Eyewear protection coefficients were calculated as a function of a new metric of eyewear fit. In general, better protection was correlated with smaller gaps between the eyewear and face. For oblique angles, most spectacles actually potentiated the blast wave by creating higher peak pressures at the cornea. Installing foam around the perimeter of the spectacle lens to close the gap between the lens and face resulted in significantly lower pressure at the cornea. In conclusion, current eye protection, which was designed to reduce secondary and tertiary blast injuries, provides insufficient protection against primary blast injury.

INTRODUCTION Deployed military personnel are at particular risk of a spectrum of ocular injuries caused by blast, including penetrating eye injury, retinal detachment, eye rupture, intraocular hemorrhage, corneal laceration, etc.1 Blast eye injuries are thought to be caused typically by secondary blast effects where foreign objects come in contact with or penetrate the globe, and most investigations and models of eye injuries have focused on such secondary mechanisms of ocular blast injury. Although some clinical cases of primary blast injury (PBI) have been documented,2–4 only recently has PBI gained significant visibility within the research community,5,6 and strong evidence revealing clinically significant primary ocular trauma was published in 2014.7

To help protect against ocular blast injuries, the Authorized Protective Eyewear List (APEL), which incorporates acquisition guidelines for ballistic protection, was adopted in 2006. With growing laboratory evidence that the blast wave may cause ocular damage as well as higher visual system injuries, we began investigating the effectiveness of APEL eyewear in PBI protection. At the 2014 Military Health System Research Symposium (unpublished data), we reported some limitations for spectacles on the APEL in reducing the blast intensity reaching the cornea. While the spectacles reduced peak corneal pressure for head-on blasts, they did not provide adequate protection from PBI at other angles and even potentiated pressure at the cornea in the vast majority of cases. The APEL goggles, on the other hand, provided an average of about 31% better protection and showed significantly better protection against head-on blasts than the spectacles (as much as 115% in some cases). We hypothesized that this result was due to the foam around the goggles creating a better seal against the face. With the exception of the Wiley X SG1 (which had foam around the inside of the lenses), APEL spectacles are characterized by large gaps between the frame and/or lens and the face, and we posit that the amount of blast pressure allowed to reach the eye is a function of this gap size. Herein, we describe experiments designed to evaluate blast wave protection as a matter of form and fit of the protective eyewear to test our hypotheses.

MATERIALS AND METHODS
A shock tube was used to produce blast waves for this study. The shock tube barrel and driver have an inner diameter of 6 inches and a thickness of about 1 inch. The length of the barrel is 122 inches, and it feeds into a catenoidal horn that expands from the 6-inch diameter to a 4-foot-square outlet. A Mylar diaphragm at the tube chamber was used to generate the blast wave. A pressure–time curve of the Friedlander wave produced by the shock tube is presented in Figure 1. The free-field overpressure measured at the test section was 1.6 psi with a standard deviation of 1.64% between tests. The positive phase of the shock tube Friedlander wave was approximately 1.5 msec.

Two headforms of different sizes (15th and 95th percentile based on male head circumference) made from a polyurethane cast were fitted with blast pressure sensors from PCB Piezotronics (Depew, NY) (model no. PCB102B18). The sensors were installed at the center of each eye location, and were flush with a nylon washer representing the corneal...
The headform was fixed on an adjustable mount placed 13.5 inches from the exit of the shock tube horn. A PCB Piezotronics free-field blast pencil probe (model no. PCB137A23) was affixed 17 inches to the side of the headform at eye level to measure the free-field blast pressure and these pressure readings were used to correct for minor shock-to-shock variability. Data were collected with the headform rotated in 30° intervals. Five tests were performed for each protection case at each headform orientation, resulting in five 1-second measurements for the pencil probe, left, and right eye sensors recorded at a 204.8-kHz sample rate. A National Instruments (Austin, Texas) PXI-4498 data acquisition board with anti-aliasing filters was used to record the measurements and was triggered upon the bursting of the shock tube membrane. Testing was done with four types of eyewear (Fig. 2) affixed to the headform with a combat helmet. No additional mounting schemes were used beyond what was provided by the eyewear and the helmet. Separate experiments were conducted without the eyewear in place and also with foam insulation inserted on each spectacle lens to eliminate the gaps between the face and spectacles (Fig. 3). The added protection provided by the foam inserts was then compared to the protection provided by the Flakjak goggles.

RESULTS AND DISCUSSION
To establish a relationship between eyewear fit and PBI protection, three people concurrently took measurements of gap distances between spectacle lens and face at the bottom and side (Fig. 4) as well as the distance from the cornea to the center of the eyewear lens. Table I shows the measured gaps for the small and large headforms with the spectacle lenses. The distance from the cornea to the spectacle lens increased by 3 mm for each set of eyewear on the large headform. Measurements for the frame-to-side metric increased an average of about 2.3 mm for the large headform. The under eye metric showed the least difference between the small and large headform, with a 0.5-mm gap increase for the large headform with the Talon and Sawfly spectacles and a 0.5-mm gap increase for the small headform with the Genesis spectacles. As a means of establishing gap distances as a critical factor in PBI protection, we compared the average of the five peak pressure measurements at each orientation for the small and large headforms taken with the Sawfly and Talon eyewear (Fig. 5). These are polar plots of the average of the five calculated protection coefficients, which are simply a ratio of the peak pressures measured with and without protection at their respective orientations, i.e., a ratio...
less than one denotes protection, whereas a number greater than one denotes intensification. To account for peak pressure variability at the eye between each test ($\sigma = 0.95\%$), the peak pressure measured at the eye for the protected and unprotected case was normalized by the peak pressure measured by the pencil probe for each respective run. Note the degree markings for each eye are from the perspective of the blast source. In other words, the 30° azimuth measurement is with the head turned counterclockwise with respect to the shock tube exit. The reduced gap distances noted for the smaller headform led to better protection coefficients for the Talon spectacles (right plot in Fig. 5). Results were mixed for the Sawfly spectacles (left plot in Fig. 5), but the protection coefficients for the small headform were better in most cases with an average difference of about 7%. This suggests that the protection coefficients were worse for the larger headform because of the larger gaps exhibited between the spectacle lens and face, though other anatomical variations between the headforms such as the difference in brow ridge size would have a nonzero but less significant effect.

Figure 6 shows the Sawfly and Talon protection coefficients for the large headform with and without the foam inserts. The foam provided noticeable improvement in protection, especially for shocks originating from the front and right side. Figure 7 illustrates the similarity in performance between the spectacles with foam inserts and the Flakjak goggles on the large headform, which have the aforementioned foam and rubber seal around the frame.

---

**TABLE I.** Measured Gaps Between Headform and Spectacle Lens

<table>
<thead>
<tr>
<th></th>
<th>Genesis</th>
<th>Talon</th>
<th>Sawfly</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small Headform</td>
<td>High Frame to Side (mm)</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>Under Eye (mm)</td>
<td>11</td>
<td>7.5</td>
</tr>
<tr>
<td></td>
<td>Cornea to Lens (mm)</td>
<td>24</td>
<td>23</td>
</tr>
<tr>
<td>Large Headform</td>
<td>High Frame to Side (mm)</td>
<td>12</td>
<td>10.5</td>
</tr>
<tr>
<td></td>
<td>Under Eye (mm)</td>
<td>10.5</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Cornea to Lens (mm)</td>
<td>27</td>
<td>26</td>
</tr>
</tbody>
</table>

---

**FIGURE 3.** Large headform with foam inserts on the inside of the Uvex Genesis spectacles.

**FIGURE 4.** Under eye (L) and frame-to-side (R) gap sizes for Uvex Genesis spectacles being measured on the large headform.

**FIGURE 5.** Right eye protection coefficients for Sawfly (L) and Talon (R) spectacles for small and large headforms. In this figure and in the other polar plots, the bold circle is at unity. Data points residing inside the unity circle represent net protection and data points residing outside the circle represent an increase in blast energy reaching the cornea when the eyewear is worn.
Figure 8 shows the average protection coefficients for each of the spectacles for four conditions (no foam, foam, side foam out, bottom foam out). Dramatic improvement was provided by the foam inserts (as high as 54%). There was a small difference in performance between the side and bottom foam inserts, with the bottom foam providing slightly better protection in all cases. This could be due to the bottom foam occupying a greater volume of space than the side foam.

CONCLUSIONS

Gap distances between the spectacles and face were greater when using the larger headform, and this provided a test case for assessing gap distance versus PBI protection. As hypothesized, the larger gap size led to increased blast loadings reaching the cornea. When foam inserts were placed on the edges of the spectacles to reduce the gap distances, dramatic improvement in protection was observed. The improvement was comparable to protection seen with the Flakjak goggles.

By removing part of the foam inserts, we attempted to evaluate the protection attributable to a particular gap. The performance of the spectacles was slightly worse when the bottom foam was removed compared to the side foam. Although a clear trend between gap size and protection could not be established with our limited metrics, it is apparent that filling the gap with foam dramatically improved the protection of the spectacles. It is clear from these studies that APEL spectacles are poor at reducing reflected blast pressure at the eye, and reducing or eliminating gaps between the spectacle and face would likely improve performance in this area. That said, there are other protection metrics that could be investigated, such as the positive phase duration of the blast wave and the integrated pressure of this duration. The role of these metrics in PBI is not currently well established; nevertheless, we are investigating the differences of these metrics between each protection case and will be presenting our findings as a supplement to the data exhibited here.

ACKNOWLEDGMENTS

This research was supported in part by an appointment to the Postgraduate Research Participation Program at the U.S. Army Aeromedical Research Laboratory administered by the Oak Ridge Institute for Science and Education through an interagency agreement between the U.S. Department of Energy and the U.S. Army Medical Research and Materiel Command. This effort was supported by core funds from the Medical Research and Materiel Command’s Military Operational Medicine Research Program.

REFERENCES

Does Combat Hearing Preservation Equipment Affect Balance?

**ABSTRACT** Objectives: (1) To investigate whether the occlusion effect and hearing attenuation produced by 3M Combat Arms Ear Plugs (CAEP) affects balance when compared to no hearing protection and (2) to investigate whether the occlusion effect and noise-canceling capabilities of the Nacre QuietPro system affects balance when compared to no hearing protection. Methods: This prospective study collected pilot data for investigation of mechanisms of balance. 20 subjects with normal hearing and no vestibular dysfunction were tested with blackened goggles in three conditions—no hearing protection, CAEP, and with the Nacre QuietPro. Results: A static posturography forceplate was used to measure center of gravity angular acceleration for a period of 20 seconds. The order of the conditions tested was randomized for each individual. Mean angular acceleration and standard deviation (degrees/second) of the three conditions were: (1) no hearing protection (control), 0.65 + 0.19, (2) CAEP, 0.69 + 0.23, and (3) QuietPro, 0.70 + 0.20 (one-way analysis of variance [ANOVA], df = 2, F = 0.38, p = 0.706). Conclusions: The components of an intact balance system include a variety of neural inputs, to include vestibuloocular, vestibulospinal, and labyrinthine afferents. Both the CAEP and Nacre QuietPro are hearing preservation devices utilized during Operation Iraqi Freedom and Operation Enduring Freedom in Afghanistan. Our pilot data show no decrement in balance with utilization of these combat hearing preservation devices.

**INTRODUCTION**

Noise-induced hearing loss (NIHL) is the number one cause of disability affecting active duty, former, and retired members of the U.S. Armed Forces.1 NIHL, with related disease to include tinnitus, has a cost burden to the Veteran’s Administration exceeding $1.2 billion annually.2 NIHL is typically acquired in the form of sensorineural hearing loss as a result of damage to the cochlea, as well as the subsequent nerve conduction pathway from the cochlea to the brain.3 Hearing preservation equipment, largely in the form of ear plugs, is an integral part in preventing NIHL when worn correctly and routinely in the appropriate operational setting. Military service members cite discomfort wearing hearing preservation equipment and decreased ability to communicate as being common reasons for noncompliance.4 From an audiometric standpoint, use of combat hearing preservation equipment decreases speech localization, reception, and intelligibility.5,6 According to a 2006 study of 301 U.S. Navy Atlantic and Pacific Fleet flight deck personnel, 47% reported never wearing hearing protection with only 7% correctly wearing their ear plugs with enough depth to achieve the maximum noise attenuation.7 In addition to these previously published reasons, a common complaint among service members anecdotally is subjective imbalance with the use of these devices, of which the occlusion effect of the ear plugs and subsequent loss of auditory cues may be contributory. Evaluation of this subjective imbalance is difficult in these patients—typically reported after field training and deployment—and often confounded by the presence of headache syndromes, blast exposure (to varying degrees), as well as traumatic brain injury.8,9 There are currently no studies in the literature that test the effect of isolated hearing loss on balance.

The components of an intact balance system are complex, assimilating visual, proprioceptive, and auditory neural inputs. These inputs are mediated by the vestibuloocular, vestibulospinal, and labyrinthine pathways, respectively.10,11 Static posturography is a vestibular test of these integrated inputs recording displacement of center of pressure on a force-measuring platform.11,12 For the purposes of this study, simulated loss of auditory cues by active and passive noise-attenuating, hearing preservation equipment were utilized to simulate hearing loss in the setting of otherwise normal hearing and balance systems. The Department of Defense (DoD) utilizes both passive and active noise reduction for combat hearing preservation. Two widely-used hearing preservation devices are the 3M Combat Arms Ear Plugs (St. Paul, Minnesota) and the Nacre QuietPro system (Trondheim, Norway).

The Combat Arms Ear Plugs (CAEPs) are a set of in-the-ear plugs issued for operational use, which passively attenuate sounds at all frequencies. The Nacre QuietPro system has been targeted for use for Special Forces units of the U.S. Armed Forces because of its ability to combine passive noise reduction (foam in-the-ear plugs); integrated into the radio headset are electronic active noise reduction that attenuates low-frequency ambient noise, as well as high-level impulse noise.
Objectives of this study were to (1) investigate whether the occlusion effect produced by 3M CAEP affects balance when compared to no hearing protection and (2) investigate whether the occlusion effect and noise-canceling capabilities of the Nacre QuietPro system affects balance when compared to no hearing protection.

METHODS
This prospective, randomized study conducted at the Otolaryngology–Head & Neck Surgery Department of Naval Medical Center, San Diego, collected pilot data for the aforementioned objectives. 20 subjects with normal hearing and no previously documented vestibular pathology were tested in three conditions—no hearing protection, CAEP, and with the Nacre QuietPro. The testing condition sequences were randomized to different orders to prevent bias; all subjects performed all three scenarios. Patients served as their own controls in this regard in the no hearing protection setting.

Inclusion criteria for study included 18 years of age or older, screening audiogram with normal hearing (less than or equal to 20 dB HL) in all frequencies, tympanometry demonstrating bilateral type A results. Further, all patients had a normal comprehensive head and neck physical examination. Exclusion criteria included younger than 18 years of age, previous diagnosis of hearing loss, vestibular disorders, chronic ear disease, previous neurosurgical or neurotologic surgery, history of concussions, closed head injury/traumatic brain injury, blast exposure, or mechanical instability.

A static posturography forceplate was used to measure the subjects’ average angular acceleration of their center of gravity for a period of 20 seconds. Forceplates were equipped with six force sensors designed to measure the three force components along the x-, y-, and z-axis, as well as their corresponding moments. This allowed for calculation of angular acceleration about the center of gravity. Testing occurred in a room free from external auditory cues. The patients performed the test in the following manner: (1) standing in a staggered stance at shoulder width to optimize mechanical stability, with weight centered on the forceplate (stance kept for all three scenarios with center of gravity zeroed for each patient), (2) eyes instructed to be open, with blackened goggles worn for the entirety of the experiment to mitigate the effect of any visual cues, and (3) head-positioned level to the ground with a change in head position from 45 degrees left-of-center to 45 degrees right-of-center at a rate of one half-cycle per second (either left-to-right or right-to-left movement); the half-cycle cadence was initially acquired using a metronome, then shut off to prevent any auditory cues during testing. The rationale for this movement was to challenge the balance system in a consistent way to evaluate for any change in angular acceleration. Testing then commenced in the aforementioned three conditions with proper earpiece fit and occlusion visualized by the tester.

RESULTS
20 patients were successfully recruited, with a single patient excluded due to hearing loss. Mean angular acceleration and standard deviation of the center of gravity (in degrees/second) of the three testing conditions demonstrated were (1) no hearing protection, 0.65 ± 0.19, (2) CAEP, 0.69 ± 0.23, and (3) QuietPro, 0.70 ± 0.20 (one-way ANOVA, ρ = 0.706, df = 2). These findings indicate no significant diminution in balance utilizing the hearing preservation devices with respect to no hearing preservation (Fig. 1).

![FIGURE 1](image_url) Mean angular acceleration and standard deviation of the three testing conditions: (1) no hearing protection, 0.65 ± 0.19; (2) CAEP, 0.69 ± 0.23; and (3) QuietPro, 0.70 ± 0.20 (one-way ANOVA, ρ = 0.706, df = 2).
Further, single-subject design analysis also demonstrated that there was not a significant difference when comparing the conditions within each individual.

DISCUSSION
NIHL is a significant clinical condition that impacts the lives of many active duty and retired service members. The condition also continues to be a cost burden to the health systems in which these individuals receive care, even after the completion of their military service. Despite these impacts, the literature reports that there is relatively poor compliance with hearing preservation equipment—even in scenarios of likely significant impulse acoustic trauma, such as a potential blast injury, firing of noise-intense weapons, or aircraft takeoff.6,7

The components of an intact balance system are complex, and include a variety of neural inputs—including visual, auditory, and proprioceptive cues.8 Static posturography is a vestibular test that examines these integrated inputs by measuring displacement of center of pressure on a force measuring platform.9–11 The only previous studies in the literature that study the effect of middle ear pathology or hearing loss on static posturography involve pediatric patients affected with chronic serous otitis media. These data were inconclusive over the impact of the conductive hearing loss in the setting of middle-ear fluid on balance function.15,16 Our study is the first to explore the effect of the loss of hearing and potential occlusion effect caused by the hearing conservation equipment on balance function as measured by postural stability.

The DoD utilizes different types of hearing preservation equipment depending on military duties and subsequent noise exposure. These include the widely distributed CAEP, as well as noise-cancelling hearing preservation devices, used in smaller distribution for communities such as special operations. Of note, not tested within the confines of the study include foam inserts that are used for intermittent exposure, as well as ear-muffs and helmets used by the aviation community. As mentioned earlier, compliance with respect to usage and correct placement of hearing protection is fairly poor secondary to an overall reduction in situational awareness by way of decrease in auditory detection, sound localization, and speech intelligibility.2,5 The overall low compliance is congruent with the prevalence and cost burden of NIHL.1,6,12

Anecdotal reports that are wearing ear protection can impact their subjective balance function may further contribute to low compliances. It is unclear whether this is individual perception or simply “common operational knowledge” imparted over the years of combat. If the latter is the case, then education of a lack of balance impairment occurring with hearing protection could help ease this concern.

Our pilot study suggests no effect of CAEP and noise-cancelling hearing preservation equipment on balance. This was tested specifically in individuals free from hearing loss or preexisting vestibular disorders to prevent confounding clinical history. Further, testing was performed in a staggered stance, free from visual cues to mitigate effects of mechanical instability or visual stabilization.

It should be noted that this work was a small study conducted in a laboratory setting, which may not represent a perfect proxy for a combat environment. In addition, we tested military subjects without any preexisting hearing loss; individuals with hearing loss may be more so affected by the additional reduction in hearing ability produced by ear protection. Expanding this study in a larger cohort of combat-experienced individuals with more realistic settings (background noise and more real-world testing environment) can provide further clarity in this field. Information from this study in addition to education regarding proper usage and fit may also help to increase training and combat usage of current hearing preservation equipment.

CONCLUSIONS
Both the 3M Combat Arms Ear Plugs and Nacre QuietPro are hearing preservation devices currently employed by the DoD at home and abroad, to include Operation Iraqi Freedom and Operation Enduring Freedom in Afghanistan. Our pilot data show no decrement in postural balance function with utilization of these combat hearing preservation devices.

ACKNOWLEDGMENTS
Travis Pfannenstiel, MD, and Chadwick Donaldson, MD, contributed in the conception and development of this research protocol.

REFERENCES


Human Contrast Acuity Variability

COL Morris R. Lattimore, MS USA (Ret.)

ABSTRACT  Background: Army vision standards have varied little from Aviation’s nominal birth. On the basis of classic Snellen acuity, we simply cannot predict threshold skill levels of any one individual(s). A growing number of Army Flight Surgeons, clinicians, and vision scientists have argued for the inclusion of contrast acuity metrics within flight physical standards. Methods: Previous monitoring of operational contact lens utility in 223 Apache pilots, visual acuity data were gathered under two conditions: high illuminance; low illuminance combined with low contrast. Spectacle, contact lens, and aging influences were evaluated. Results: The high-contrast Snellen acuities clustered at 20/15 and 20/20. Low-contrast acuities stretched from 20/25 to 20/125. LogMAR analysis highlighted statistical significance between the two acuity sets (p < 0.001) to an unanticipated data spread. The known underlying mechanisms possibly related to this effect are poorly documented; all such variables collectively explain <30% of the known variation in low-illumination vision. Discussion: Some pilots possessed the capacity to resolve 20/25 lettering under obfuscating conditions; others were adversely influenced by those same conditions. Snellen acuity involves target recognition; contrast acuity detects threshold differences; both aspects can be important. Conclusion: Prescreening under both vision assessment conditions will help identify and select superior vision performers. The validity and predictability of documenting this effect is targeted within planned future research efforts.

INTRODUCTION

Established Army aviation vision standards have varied little from the initial days of the U.S. Army Aviation’s nominal birth, which was after the U.S. Army Air Corps’ transition to the U.S. Air Force in the late 1940s. In 1991, approximately 23% of all Army aviators wore a spectacle correction in order to comply with aviation’s high-contrast Snellen acuity standard of 20/20 in accordance with Army Regulation 40-501; at that time, the U.S. Air Force figure was slightly higher at 27%, while a more recent USAF document reports 41% of its active duty pilots require corrective lenses to perform flight duties. Historically, military aviation research has varied little beyond the utilization of those long-established, standardized clinical tests, with no directed goal toward an improved understanding of natural threshold-level physiological limitations; particularly as these tests relate to the specialized military equipment, fielded in response to traditional or operational combat challenges. Current Army flight standards for vision require only basic high-contrast assessments for near and distant Snellen visual acuity, stereo acuity, confrontation fields, color vision, and phoria or vergence testing. The advantage of high contrast testing is that it easily permits the determination of population norms, as well as the determination of performance standards (with very small standard errors). Beyond the classic Snellen acuity record, we simply do not know what the near-threshold skill levels (in low contrast under dim retinal illuminance) of any 1 individual, or group of individuals might be. For this reason a growing number of Army Flight Surgeons, as well as clinicians and vision scientists have argued for the inclusion of contrast sensitivity/contrast acuity metrics within flight physical standards. Second, passing a flight physical when well rested is not the same as when fatigued. For example, any individual with a well-compensated phoria may easily fuse binocular images when well rested, but suffer from double vision when stressed and/or fatigued.

New operational challenges have been identified primarily as a function of the environmental conditions encountered, and their secondary effect on visual performance requirements. A degraded visual environment (DVE) has been variously described as being under fog, brownout, or whiteout conditions, such that the overall level of illumination is decremented, matched by decreased target contrast. Brownout conditions, also referred to as a DVE have cost the Army numerous rated aviator lives, as well as over $1 billion in rotary-wing aircraft damage, resulting from approximately 800 class A accidents over the 8-year period of 2002 to 2009 (a class A accident involves the possible loss of one or more lives, with aircraft destruction or damage exceeding $2 Million). The potential fielding of a variety of technological solutions in response to the DVE threat is the Aviation Program Executive Officer’s (PEO’s) number one priority. Although these U.S. Army Research, Development and Engineering Command-developed technological countermeasures to DVE have the goal of making landing, navigating, and fighting easier, the countermeasures themselves may task certain individual’s sensory limitations in ways that are not addressed under current physical examination standards. Human visual resolution performance, initially characterized within the academic community in 1972 had more recently been highlighted in 2009 by way of the publication of the Training and Doctrine Command’s Human Dimension White Paper.  

Vision Performance and Protection Division, U.S. Army Aeromedical Research Laboratory, 6901 Farrel Road (Building 6901), Fort Rucker, AL 36362.

This article was presented at the Military Health System Research Symposium, in Fort Lauderdale, FL, August 17–20, 2015.

The opinions interpretations, conclusions, and recommendations are solely those of the author and are not an endorsement of the U.S. Army, the U.S. Department of Defense, or the U.S. Government.

doi: 10.7205/MILMED-D-16-00208
New imaging and cockpit technologies, developed to compensate for engineering-developed DVE limitations, have the expressed goal of enhancing situational awareness under adverse conditions. However, they can potentially serve to increase workload, and may require changes to medical surveillance, as well as require changes in entrance or accession vision standards (e.g., visual performance: under low-contrast conditions; or after refractive surgery; exclusionary concus- sive history; inability to utilize spatial hearing). Existing occupational vision standards fail to reflect the unique visual demands of these newly developed technologies, as well as their modern combat-employment environment. Given those realities, the U.S. Army Aeromedical Activity Director and the U.S. Army Consultant for Aviation Medicine have both requested research into the possible establishment of new aeromedical standards (aligned toward documenting aircrew proficiency and fitness to operate these novel technological systems under diminished environmental conditions). Therefore, the intent of this article is to review background threshold-associated visual function issues as they relate to demonstrated contrast acuity variability under varying conditions of illumination; the goal of which is the development of an evidence-based medical reference for justifying the establishment of new rotary-wing aeromedical vision performance standards. Consequently, the impetus for the preparation of this manuscript stems from the already existing summed scientific evidentiary support for the inclusion of contrast acuity metrics in flight physical requirements.

METHODS AND MATERIALS

A much earlier study on the practical usefulness of contact lens wear by Army aviators had recorded standard Snellen acuity and contrast-based acuity as a part of the subject follow-up process. That specific USAARL Technical Report provided no comparative analysis, since the primary focus of the study, completed in 1992, was the operational flight utilization of contact lenses under deployed combat conditions. However, AR 40-501 has established medically based vision performance standards for accession, as well as for retention in all aspects of Army service. Usage of de-identified data from that original contact lens protocol was approved by the Institutional Review Board of the U.S. Army Medical Research and Materiel Command for use in this expanded data analysis and report.

Each subject had been provided written and verbal informed consent before their original study participation. Examination initially occurred daily then quarterly exams occurred thereafter over the next 3 years. The initial data evaluation involved the utilization of two separate means of assessing subjective visual acuity documented. Acuity testing used a standard Snellen projector, shown onto a highly reflective screen (at 106 lux), and a Bailey–Lovie low-contrast (8%) chart. Acu- ties were determined before, and following every contact lens research evaluation. Additionally, Bailey–Lovie low-contrast acuity was measured at the same 20-foot distance and under conditions of moderate to low illuminance, at 31 lux. Visual acuities were measured in 223 subjects numerous times over 3 years. A logarithmic conversion of the Snellen data, using a minimum angle of resolution (LogMAR) equivalent, gained the benefit of parametric continuous-variable analyses, including the application of inferential statistical analysis (as opposed to the use of nonparametric discrete variable determinations, which are awkward and of limited usefulness).

RESULTS

Figure 1 provides a generalized visual comparison of the acuities of all 223 CONUS subjects, constituting over 2,000 clinical exams. Frequency distribution analyses of the LogMAR converted acuity responses highlighted the vast difference between the two disparate visual acuity tests (high-illuminance–high-contrast acuity versus low-illuminance–low-contrast acuity; \(p < 0.0001\)). Examining the Figure 1 data, it can be seen that any aviator could possess normal Snellen acuity (20/20 or slightly better), yet potentially exhibit anywhere on a continuum from exceptional to poor-quality low-contrast/low-illuminance acuity. Reverting to the original study, high-contrast Snellen visual acuities through contact lenses were not significantly different from high-contrast Snellen visual acuities through the subject’s normally worn corrective spectacles. However, the acuity data in Figure 1, which is entirely through contact lenses under both conditions, reveals a significantly different distribution spread, with the low-contrast, low-illuminance data possessing an expanded data spread.

DISCUSSION

In consideration of the two singular assessments reviewed in this manuscript, a variety of potential sources for error were in operation to varying degrees. Comprehensive higher-order aberrations were evident in all conditions at equal degrees of involvement. Since each subject served as their own control, this issue presented no differential effects. The presence of high-illuminance/high-contrast acuity conditions equated to miotic or small pupil sizes, potentially affecting subject responses under one condition. Under the latter testing condition noncycloplegic influences, such as reducing retinal illuminance via the use of partial filters caused pupillary dilation, thereby indirectly inducing an increase in spherical aberration, a differential confounding factor for the low-contrast condition. Furthermore, defocus or an optical blur in the absence of a dilated pupil, was operant on an equal basis, because both acuity testing methods were performed under the refractive condition of “best visual acuity” or BVA. Thus, other issues regarding the disparate data distribution evidenced for low-contrast/low-illuminance acuities in Figure 1 must be considered.

Extreme data spread in the low-illuminance/low-contrast condition were concluded to not be due to direct spectacle, contact lens, or aging influences. The data spread could be the result of poor overall uncorrected astigmatia control, since the contact lenses were spherical in nature. The contact lens-wearing condition did correlate inversely in a marginally
significant fashion \((R = 0.12)\). Therefore, some of the low-illuminance/low-contrast acuity distribution spread is exaggerated due to uncorrected astigmatism in the contact lens condition. A correlation of 0.12 essentially meant that no more than 2% of the observed low-illuminance/low-contrast acuity data spread had been due to uncorrected astigmatism in soft contact lens wear (i.e., 1.44%). Similarly, Allard et al\(^9\) had documented age-based effects on acuity as a potential source of data contamination; we saw no systematic age related spread in acuities. However, Allard’s subjects were in their 60s and 70s (the decremented effect had been hypothesized as perceptual “noise,” due to advanced age); our subjects in this limited dataset were 19 years to 46 years of age, explaining why we saw no age-based acuity changes.

Since the diameter of the pupil of the human eye changes as a function of retinal illuminance, many studies seek to control pupil size by using an artificial pupil of a constant diameter in order to keep the degree of retinal illuminance constant (thus controlling one of the variables potentially influencing rod vs. cone functioning). However, while an artificial pupil is an excellent means of variable control, it induces an artificial condition that is not normally encountered by the visual system, potentially artificially influencing the experimental outcome. Therefore, control of the pupil size should not be relied on for use under DVE conditions.

In recent years, a number of investigators have sought to develop sensitive means of assessing visual resolution via practical clinical-based testing.\(^{10,11}\) Although contrast sensitivity testing has proven itself as a penetrating visual performance diagnostic, such testing has been primarily isolated to the research realm due to its cumbersomeness to administer and apply. Consequently, a number of practical clinical offshoots have evolved (e.g., the Bailey-Lovie Computerized Low-Contrast Test, and the computer-based and tablet-based Rabin small letter contrast test [SLCT]).\(^{12,13}\) The latter example represents the current USAF standardized contrast acuity test, or have become proven, established tools capable of easy application within the realm of an aviation-based eye care clinical screening program. Their demonstrated benefits are 3-fold: as a measurement of the integrity of both the central and peripheral visual processing centers; as an indicator of detail resolution (pertinent to facial recognition or highly specific tasks); and, as an indicator of general figure/ground function (pertinent to movement within a complex environment). Contrast sensitivity testing has been shown to be superior at predicting a pilot’s performance in detecting small, low-contrast targets in simulators as well as in the field,\(^{14,15}\) which is of potential importance to current military aviation DVE research efforts. Full scope contrast sensitivity testing under cycloplegic conditions had been proposed as a critical visual assessment task integral to the Army’s class 1 flight physical. During a class 1 flight physical, a topically applied 1% cyclopentolate solution will artificially induce paralysis of the ciliary muscle of the eye. The topical

![Figure 1. Acuity response frequency.](image-url)
cycloplegic pharmaceutical primarily inhibits accommodation; a subsequent secondary effect is then seen as pupillary dilation, which can induce some spherical aberration or secondary blur.\textsuperscript{16–19} Initial SLCT research has shown the SLCT’s sensitivity to be more discriminating than traditional visual acuity testing. It is also more responsive to small amounts of blur,\textsuperscript{20} to subtle changes in the luminance of the stimulus,\textsuperscript{21} to vision with 2 eyes compared to one eye,\textsuperscript{22} and for identifying visual differences among pilot trainees.\textsuperscript{23–25} The goal of reviewing the preceding research data had been to examine the practical importance of this type of visual performance tool as a means of quantifying the degree of subtle visual performance differences that high-contrast/high-illuminance Snellen acuity fails to detect. However, all the justifications summarized in the preceding text highlight the importance of testing DVE sensitivities as a normal function of military physical examination.

Computer display-based contrast threshold systems are available for the organization of faster subject screening assessments, which are now included in visual performance planning by this investigator. As part of the “Force 2025 and Beyond” Initiative,\textsuperscript{26} the Army has begun to reprioritize its Science and Technology (S&T) needs. Key to that S&T reprioritization is an increased emphasis on human performance optimization, defined as the process of applying knowledge, skills, and emerging technologies to improve and preserve the capabilities of Department of Defense personnel to execute essential tasks. External comparisons of the two acuity determination methods emphasized statistically significant differences that have potential for future use in identifying superior visual performers. If this research potential is realized, then the establishment of visual performance standards, in the military in general, or in military aviation specifically, will need to be modified to include the full-scope application of these visual performance assessments. The current vision sensitivity recording of only the upper-level or ceiling effect documented by high-contrast Snellen acuity will necessarily be expanded on, in order to document the floor- or threshold-level of absolute visual sensitivity. Given the Training and Doctrine Command’s emphasis on the identification, development, and optimal integration of human capabilities, this manuscript had been prepared in an effort to stimulate such an expanded assessment of Army physical examination, visual performance standards (color standard modernization should not be overlooked either). Long-term approaches toward expanded examination regarding the role of color sensitivity, binocularity, and stereopsis, as well as cognitive visual processing (e.g., reading comprehension, short- and long-term memory, and eidetic memory) would complete the overall physical, and neurological functional analyses required for a complete understanding of one’s ophthalmic health.

A wide variety of additional factors with the potential to influence visual resolution (corneal distortions, lenticular alignment, aspect relationship errors, fusional, and stereoscopic errors, as well as numerous anatomic optical system variations)\textsuperscript{27,28} can cumulatively contribute to reduced image clarity, as well. However, neural processing applications could partially balance those confounding effects from anatomic variation. This neurological adaptational ability has previously been identified as a critical factor related to visual recovery from refractive surgery.\textsuperscript{29} The conceptual framework for providing a global assessment of threshold-linked visual performance is dependent, to varying degrees of influence, on two primary factors:

a) Optical factors (i.e., pupil size and shape; corneal shape, lenticular shape, and overall ocular shape changes over time) have been identified as a likely predominant influence responsible for affecting contrast acuity,\textsuperscript{30} and/or:

b) Neural, adaptational factors as the predominant influence responsible for affecting contrast acuity (i.e., when image presentation is under low contrast, and retinal illumination is decremented).\textsuperscript{30}

In general, recent investigators, after reviewing their own data, as well as that of others, have concluded that factors other than refraction are of primary influence in visual acuity resolution under mesopic, low-contrast/low-luminance conditions.\textsuperscript{31,32} The mesopic range of illumination, roughly described as a mid-level illumination reduction, is characterized by concurrent rod and cone function, which could theoretically occur in-concert, but not necessarily as a constant, due to theories of photoreceptor competition. Basic and applied research into this quasi-joint realm of conflicting photoreceptor functions is critical to understanding DVE sensitivities. Further, decreases in visual resolution occurring at low-contrast and decreased light levels appear to be subordinate or secondary to internal central nervous system-based (cns) neural factors, and not from optical blur or spherical aberration secondary to increased pupil size.\textsuperscript{33,34} The same investigators also recognized that microfluctuations within the accommodative system, occurring once again under a setting of low-contrast and decreased illumination, will contribute to decreased visual resolution.\textsuperscript{35} Similarly, eye movement variability, which also increases in the dark, directly contributes to increased fixational instability and decreased visual resolution.\textsuperscript{36} The payoff in providing a medical evidence-based array of understanding threshold-level visual function influences would be a physical performance standard, against which selection of ideal candidates for specialized duty could be based. This end result would be an optimized force, capable of making use of every level of technology that is, and will be available.

**CONCLUSION**

Thus, identifying those individuals with superior contrast acuity resolution capabilities is perhaps the most effective solution in achieving an effective combat operations capability under DVE conditions (in both ground and air operations). The increased response range under low-contrast, low-luminance conditions highlights the need for further investigation. At issue are other well-established sources for potential variation
in low-illumination conditions. Theories regarding possible benefit from increased macular pigment deposition, which varies considerably across our population, serve to increase the signal-to-noise ratio due to stray light absorbance. Iris color variation is suggested to reflect varied stray light absorbance, as well. As earlier alluded to, there are a myriad of optical and neuronal activation issues potentially in operation. Lastly, mid-level lighting intensity (mesopic lighting), when both rods and cones are both activated, can be variably be influenced via photoreceptor competition issues, pitting one type against the other. Ideally, only a comprehensive visual resolution research program, in collaboration with all the Services programs, will provide a thorough understanding of the key pertinent processes governing contrast acuity. Only then can the selection and accession of superior aviation candidates be guaranteed.

ACKNOWLEDGMENTS

The author has indicated no financial relationship relevant to the manufacturers of the tests, equipment, or contents of this article to disclose. This work was supported by the U.S. Army Aeromedical Research Laboratory, Fort Rucker, AL 36362.

REFERENCES

Self-Reported Visual Quality of Life After Combat Ocular Trauma

LCDR Andrew McLaughlin, MC USN*; LTC Marcus H. Colyer, MC USA*; Denise S. Ryan, MS†; Rose K. Sia, MD‡; Eric D. Weichel, MD‡; Andrew Spiegelman, PhD§; S. G. Eaddy ¶; COL Robert A. Mazzoli, MC USA (Ret.)§; COL Kraig S. Bower, MC USA (Ret.)¶

ABSTRACT Objectives: To describe the visual outlook and quality of life of service members after combat ocular trauma. Methods: In a single-center, prospective observational study of service members sustaining ocular trauma, participants underwent a series of ocular examinations and noninvasive tests, including the National Eye Institute Visual Functioning Questionnaire (VFQ-25). Results: Of the 165 enrolled participants, 137 completed the VFQ-25. The mean VFQ-25 composite score was 74.4 ± 20.7 (range: 1.4–100). Among 118 participants with visual acuity assessment, 92% had best corrected visual acuity (BCVA) of 20/20 or better in at least one eye. Among participants with severe vision loss (BCVA ≤20/200), there was no statistically significant difference in self-reported general health compared to those without severe vision loss (p = 0.17). However, there was a significantly lower visual quality of life reported in the composite score and all of the 11 subscales of the VFQ-25. Conclusions: While this study provides evidence that combat ocular trauma is associated with a lower visual quality of life, limitations include the relatively small sample size and the limited documentation of the use of eye protection at time of injury among participants.

INTRODUCTION

Trauma is one of the leading causes of unilateral vision loss worldwide, resulting in an estimated 18 million people with unilateral blindness or low vision. Over 1 million people worldwide have lost vision in both eyes because of trauma.1,2 Although the eyes occupy only 0.1 percent of the total body surface,3 ocular trauma accounts for three percent of emergency department visits in the United States.3,4

Among military service members injured during combat before the 20th century, eye injury rates were less than four times the expected percentage based on body surface area. This rate has steadily increased during the past century to a times the expected percentage based on body surface area. The increase is thought to be due to improved diagnostic testing, due to mental or physical conditions. The views expressed in this presentation are those of the author and do not reflect the official policy of the Department of Defense or U.S. Government.

doi: 10.7205/MILMED-D-16-00139

METHODS

This is a single-center, prospective observational study of consecutive service members sustaining ocular trauma during Operation Iraqi Freedom and Operation Enduring Freedom who were evacuated to Walter Reed Army Medical Center or subsequently Walter Reed National Military Medical Center between December 6, 2006, and February 13, 2013. Approval was obtained from the Walter Reed Army Medical Center and the Walter Reed National Military Medical Center Institutional Review Boards before study implementation. All participants or their legally authorized representative gave written informed consent. Participants were included if they were at least 18 years old and had a history of a combat injury to head, face, or neck with ocular injury. Participants were excluded if they were unwilling or unable to undergo vision testing and/or special sensory or diagnostic testing, due to mental or physical conditions. When feasible, participants received a comprehensive baseline ocular examination and a standard battery of noninvasive
tests, including the National Eye Institute Visual Functioning Questionnaire (VFQ-25).13,14 Participants were excluded from data analysis if they failed to complete the VFQ-25.

The VFQ-25 is a 25-item questionnaire that was developed to assess the effect of eye disorders on a patient’s visual quality of life. It includes 24 items grouped into 11 subscales and a general health item. The questionnaire was self-administered by participants and scored on a scale of 0 to 100 in accordance with the VFQ-25 scoring algorithm, with a score of 100 indicating the best visual quality of life. Normative data have previously been described for cohorts with chronic eye disease and individuals free of known eye disease.13–23

Injury and clinical data were captured from records or by self-report at the time of enrollment and included: demographics, theater of injury, duration of follow-up, mechanism of injury, documented use of eye protection, best corrected visual acuity (BCVA), and slit-lamp exam. Due to the smaller sample size, the VFQ-25 composite score is presented and compared. Subscale scores are presented descriptively to present functional ability in this population. Statistical analysis was performed with a p value of 0.05 considered statistically significant. A Student’s t test was used to test hypotheses regarding VFQ-25 composite score, sex, age, injury cause, injury location, and number of eyes injured.

**RESULTS**

Of the 165 enrolled participants, 137 completed a VFQ-25 (83% of total enrolled). Descriptions of the injuries of the study population are described in Table I. The mean age was 26.8 years, with more than half of the participants between 20 and 29 years (range: 18–51 years). The majority of participants were male (n = 134; 98%). Among the 118 participants who completed a visual acuity assessment, 105 (92%) had BCVA of 20/20 or better in at least one eye. There were nine (8%) participants with mild or moderate vision loss (BCVA <20/20 and >20/200), and four (3%) with vision loss less than 20/200.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age (Years)</strong></td>
<td></td>
</tr>
<tr>
<td>&lt;20</td>
<td>5 (3.9)</td>
</tr>
<tr>
<td>20–29</td>
<td>89 (69.5)</td>
</tr>
<tr>
<td>30–39</td>
<td>23 (18.0)</td>
</tr>
<tr>
<td>40–49</td>
<td>10 (7.8)</td>
</tr>
<tr>
<td>50–59</td>
<td>1 (0.8)</td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>26.8 (7.0)</td>
</tr>
<tr>
<td><strong>Documented Use of Eye Protection at Time of Injury</strong></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>48 (35.0)</td>
</tr>
<tr>
<td>No</td>
<td>18 (13.1)</td>
</tr>
<tr>
<td>Not Documented</td>
<td>71 (51.8)</td>
</tr>
<tr>
<td><strong>Eye Injured</strong></td>
<td></td>
</tr>
<tr>
<td>Right Eye Only</td>
<td>57 (41.6)</td>
</tr>
<tr>
<td>Left Eye Only</td>
<td>55 (40.1)</td>
</tr>
<tr>
<td>Bilateral</td>
<td>25 (18.2)</td>
</tr>
<tr>
<td><strong>Location of Injury</strong></td>
<td></td>
</tr>
<tr>
<td>Iraq</td>
<td>87 (63.5)</td>
</tr>
<tr>
<td>Afghanistan</td>
<td>40 (29.2)</td>
</tr>
<tr>
<td>Not Documented</td>
<td>10 (7.3)</td>
</tr>
<tr>
<td><strong>Mechanism of Injury</strong></td>
<td></td>
</tr>
<tr>
<td>Blast</td>
<td>87 (63.5)</td>
</tr>
<tr>
<td>Gunshot</td>
<td>6 (4.4)</td>
</tr>
<tr>
<td>Other (Motor Vehicle Accidents, Blunt, Falls, etc.)</td>
<td>10 (7.3)</td>
</tr>
<tr>
<td>Not Documented</td>
<td>34 (24.8)</td>
</tr>
<tr>
<td><strong>BCVA at Time of VFQ-25</strong></td>
<td></td>
</tr>
<tr>
<td>Better Eye</td>
<td></td>
</tr>
<tr>
<td>20/10–20/15</td>
<td>60 (43.8)</td>
</tr>
<tr>
<td>20/20</td>
<td>45 (32.8)</td>
</tr>
<tr>
<td>20/25–20/40</td>
<td>6 (4.4)</td>
</tr>
<tr>
<td>20/50–20/150</td>
<td>3 (2.2)</td>
</tr>
<tr>
<td>20/200 or Worse</td>
<td>4 (2.9)</td>
</tr>
<tr>
<td>Worse Eye</td>
<td></td>
</tr>
<tr>
<td>20/15</td>
<td>17 (12.4)</td>
</tr>
<tr>
<td>20/20</td>
<td>26 (19.0)</td>
</tr>
<tr>
<td>20/25–20/40</td>
<td>21 (15.3)</td>
</tr>
<tr>
<td>20/50–20/150</td>
<td>12 (8.8)</td>
</tr>
<tr>
<td>20/200 or Worse</td>
<td>42 (30.7)</td>
</tr>
<tr>
<td>Not Documented</td>
<td>19 (13.9)</td>
</tr>
</tbody>
</table>

**TABLE II.** VFQ-25 Composite and Subscale Data for All Participants and Stratified by BCVA in Better Eye

<table>
<thead>
<tr>
<th>VFQ-25 Subscales</th>
<th>Overall (n = 137)</th>
<th>BCVA ≤20/200 (n = 4)</th>
<th>BCVA &gt;20/200 (n = 114)</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean ± SD</td>
<td>Mean ± SD</td>
<td>Mean ± SD</td>
<td></td>
</tr>
<tr>
<td>General Health</td>
<td>65.8 ± 22.1</td>
<td>81.3 ± 12.5</td>
<td>65.5 ± 22.7</td>
<td>0.17</td>
</tr>
<tr>
<td>General Vision</td>
<td>64.1 ± 22.3</td>
<td>15.0 ± 10.0</td>
<td>67.1 ± 20.4</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Ocular Pain</td>
<td>74.0 ± 20.3</td>
<td>53.1 ± 18.8</td>
<td>74.5 ± 19.2</td>
<td>0.03</td>
</tr>
<tr>
<td>Near Activities</td>
<td>72.7 ± 26.6</td>
<td>18.8 ± 17.2</td>
<td>74.3 ± 24.7</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Distance Activities</td>
<td>75.6 ± 24.0</td>
<td>31.3 ± 12.5</td>
<td>76.7 ± 22.4</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Social Function</td>
<td>84.2 ± 22.8</td>
<td>31.3 ± 16.1</td>
<td>86.1 ± 19.9</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Mental Health</td>
<td>71.2 ± 26.0</td>
<td>29.7 ± 21.9</td>
<td>72.6 ± 24.5</td>
<td>0.001</td>
</tr>
<tr>
<td>Role Difficulties</td>
<td>72.8 ± 28.6</td>
<td>25.0 ± 35.4</td>
<td>73.9 ± 27.0</td>
<td>0.001</td>
</tr>
<tr>
<td>Dependency</td>
<td>81.4 ± 26.9</td>
<td>31.3 ± 37.5</td>
<td>83.2 ± 24.6</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Driving</td>
<td>74.3 ± 27.1</td>
<td>0.0 ± 0.0</td>
<td>73.1 ± 26.4</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Color Vision</td>
<td>91.4 ± 20.7</td>
<td>25.0 ± 35.4</td>
<td>93.3 ± 16.1</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Peripheral Vision</td>
<td>61.0 ± 29.5</td>
<td>25.0 ± 28.9</td>
<td>61.3 ± 27.9</td>
<td>0.01</td>
</tr>
<tr>
<td>Composite</td>
<td>74.4 ± 20.7</td>
<td>27.5 ± 11.1</td>
<td>75.8 ± 18.3</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>
TABLE III. Mean Composite Scores of the VFQ-25 by Injury Characteristics.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Mean ± SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Location of Injury</td>
<td></td>
</tr>
<tr>
<td>Iraq</td>
<td>73.9 ± 20.1</td>
</tr>
<tr>
<td>Afghanistan</td>
<td>71.5 ± 23.9</td>
</tr>
<tr>
<td>p Value</td>
<td>0.604</td>
</tr>
<tr>
<td>Documented Use of Eye Protection at Time of Injury</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>80.0 ± 22.5</td>
</tr>
<tr>
<td>No</td>
<td>72.6 ± 22.2</td>
</tr>
<tr>
<td>p Value</td>
<td>0.294</td>
</tr>
<tr>
<td>Eye Injured</td>
<td></td>
</tr>
<tr>
<td>Unilateral Injury</td>
<td>76.4 ± 19.1</td>
</tr>
<tr>
<td>Bilateral Injury</td>
<td>66.2 ± 21.2</td>
</tr>
<tr>
<td>p Value</td>
<td>0.139</td>
</tr>
</tbody>
</table>

participants with severe vision loss (BCVA ≤20/200) in the better eye.

VFQ-25 subscale data are presented in Table II. The mean overall composite score was 74.4 (SD = 20.7; range: 1.4–100). The mean number of days between injury and VFQ-25 completion was 339 days (SD = 537; range: 3–2,753). Among participants with severe vision loss (BCVA ≤20/200), there was no statistically significant difference in self-reported general health compared to those without severe vision loss (p = 0.17). However, there was a statistically significant lower visual quality of life reported in the composite and all of the 11 subscales of the VFQ-25 (Table II). Table III shows the VFQ-25 composite scores by injury characteristic. There were no statistically significant differences in VFQ-25 composite scores when comparing: location of injury (p = 0.61), use of eye protection (p = 0.29), or unilateral vs bilateral injury (p = 0.14).

Among participants with documented use of eye protection at the time of injury, there was a trend for higher visual quality of life reported in the composite and all of the subscales of the VFQ-25 except color vision. However, this trend was not statistically significant for the composite or any of the subscales (Table IV).

DISCUSSION

This study is the first to use the VFQ-25 to describe the visual quality of life of service members following combat ocular trauma. Previous studies have characterized the visual quality of life of individuals with eye disease and free of known eye disease. A 4- to 6-point change in VFQ-25 score has previously been shown to represent a clinically meaningful change corresponding to a 15-letter change in BCVA. The VFQ-25 composite score for service members following combat ocular trauma is lower than the original reference cohort of 122 individuals free of known eye disease (74 ± 21 in our cohort compared to 92 ± 21 as reported by Noble et al.). The VFQ-25 composite score is also lower in this study population than that among patients previously reported with glaucoma (mean VFQ-25 composite score 88 ± 12), age-related macular degeneration (88 ± 10), multiple sclerosis (83 ± 18), idiopathic intracranial hypertension (82 ± 15), or patients who have undergone retinal detachment surgery (80 ± 15). The VFQ-25 composite scores were similar to those found in patient groups previously described in the literature with cataract (76 ± 21), keratoconus (75 ± 17), macular telangiectasia type 2 (77 ± 13), or patients after Boston keratoprosthesis implantation (72). Participants with ocular trauma resulting in severe vision loss (BCVA ≤20/200 in the better eye) reported poorer visual quality of life across all subscales. Subscale scores were most depressed for general vision, near vision, driving, and color vision. Difficulty with near vision is not typical in the age group of the study population and warrants additional investigation.

TABLE IV. VFQ-25 Subscale Data Comparison Stratified by Documented Use of Eye Protection

<table>
<thead>
<tr>
<th>VFQ-25 Subscales</th>
<th>With Eye Protection (n = 37)</th>
<th>Without Eye Protection (n = 14)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>General Health</td>
<td>66.9 ± 21.3</td>
<td>64.8 ± 22.7</td>
<td>0.36</td>
</tr>
<tr>
<td>General Vision</td>
<td>68.6 ± 21.8</td>
<td>60.7 ± 23.6</td>
<td>0.42</td>
</tr>
<tr>
<td>Ocular Pain</td>
<td>79.1 ± 24.3</td>
<td>70.7 ± 19.2</td>
<td>0.42</td>
</tr>
<tr>
<td>Near Activities</td>
<td>80.6 ± 24.5</td>
<td>67.5 ± 28.0</td>
<td>0.20</td>
</tr>
<tr>
<td>Distance Activities</td>
<td>81.1 ± 25.7</td>
<td>70.2 ± 23.3</td>
<td>0.86</td>
</tr>
<tr>
<td>Social Function</td>
<td>85.5 ± 27.2</td>
<td>81.6 ± 20.8</td>
<td>0.98</td>
</tr>
<tr>
<td>Mental Health</td>
<td>79.1 ± 25.9</td>
<td>66.1 ± 25.7</td>
<td>0.17</td>
</tr>
<tr>
<td>Role Difficulties</td>
<td>81.4 ± 26.0</td>
<td>67.0 ± 28.0</td>
<td>0.17</td>
</tr>
<tr>
<td>Dependency</td>
<td>86.6 ± 24.5</td>
<td>78.0 ± 29.2</td>
<td>0.32</td>
</tr>
<tr>
<td>Driving</td>
<td>86.3 ± 18.2</td>
<td>68.7 ± 32.6</td>
<td>0.38</td>
</tr>
<tr>
<td>Color Vision</td>
<td>89.6 ± 24.9</td>
<td>92.1 ± 18.7</td>
<td>0.92</td>
</tr>
<tr>
<td>Peripheral Vision</td>
<td>71.5 ± 30.6</td>
<td>52.3 ± 31.5</td>
<td>0.11</td>
</tr>
<tr>
<td>Composite</td>
<td>80.0 ± 22.5</td>
<td>70.1 ± 20.9</td>
<td>0.29</td>
</tr>
</tbody>
</table>
While this study provides evidence that combat ocular trauma is associated with a lower visual quality of life, limitations include the relatively small sample size and the lack of documentation of the use of eye protection at time of injury among participants. It is important to note that confounders were not taken into consideration when comparing the VFQ-25 composite scores to those previously reported in other ocular conditions. These results may not be generalizable to service members who sustain combat ocular trauma in future conflicts as changes in warfare and the combat environment may result in changes to injury patterns in the future. Further research and tracking of visual quality of life outcomes in service members with a history of combat ocular trauma are warranted. Future research should also assess the impact of anatomic location of combat injuries to the eye on the visual quality of life outcomes as this may aid in the design of future eye protection and in the management of combat ocular trauma.

REFERENCES
Understanding Post-Deployment Reintegration Concerns Among En Route Care Nurses: A Mixed-Methods Approach

LTC Felecia M. Rivers, AN USA; Lt Col Susan Dukes, USAF NC; Lt Col Jennifer Hatzfeld, USAF NC; COL Linda H. Yoder, AN USA (Ret.); Sandra Gordon, BA; LTC Angela Simmons, AN USA

ABSTRACT The objective of this study was to better understand the post-deployment behavior health symptoms and readjustment/reintegration experienced by military nurses who provided en route care while serving in Operation Enduring Freedom/Operation Iraqi Freedom. Employing an exploratory, concurrent, mixed-methods design with an electronic survey consisting of several valid instruments and single, face-to-face interviews; data were gathered from 119 surveys and 22 interviews. Four qualitative themes aligned with the Post-Deployment Readjustment Inventory items. Findings from interviews support and illuminate the outcomes of the Post-Deployment Readjustment Inventory. Behavioral health usage was high in the quantitative sample. Nearly 74% (n = 88) of respondents indicating they had used Military Behavioral Health services following deployment. Statistically significant differences were noted among all subscales except Intimate Relationship Problems. Combined results indicated en route care nurses encountered difficulties when attempting to return to predeployment roles; behavioral health problems mirrored those of combat warriors. Interventions to assist post-deployment reintegration of en route care nurses should be conducted at the peer, leader, and health care provider levels. Embedding military mental health providers into en route care units is needed. It is imperative to gather lessons learned and identify ways to improve preparation for future conflicts and behavioral health of en route care nurses.

INTRODUCTION

Returning home from the wars in Iraq and Afghanistan has proven for many to be more complicated than first imagined. Individuals who served are discovering they are not the same when they come home, the transition to a noncombat environment is difficult, and it can be challenging returning to previous roles in family, work, and community. Disturbing experiences encountered by combat warriors have contributed to substance abuse and behavioral health issues that persist for a prolonged period of time after returning home. Reintegration is the process of military personnel transitioning back into their personal, organizational, and societal roles after having been deployed. For the purpose of this study, reintegration referred to the military en route care nurses who have left the combat environment and returned home to resume their roles within family, work, and community. Difficulties, both physical and emotional, emerging from wartime experiences typically appear during reintegration.

Background

As with combat warriors, it has been identified that military nurses have similar difficulties returning to predeployment life. Outcomes of several qualitative studies targeting deployment experiences and reintegration indicate nurses are a different people after deployment and that little things are not as important as before deployment. Many mentioned nightmares as a problem, as well as dealing with anxiety and difficulties making decisions. As one author commented “one does not have to be a combatant to be traumatized by war.” Combat nursing has been described as a different kind of nursing that can prove to be personally challenging and extremely demanding, which culminates in emotional exhaustion. Typically, the expected role of military nurses during war is to provide care to service members at specific medical facilities in certain locations. However, there are others who provide extended nursing care during transport from point of injury to definitive care and demonstrate the adaptability and flexibility of nursing care required during wartime. These individuals perform what has become known as “en route care nursing.”

En Route Care Nursing

En route care nurses provide nursing care in all types of vehicles, but traditionally have been assigned to transport patients in airplanes and helicopters. In 1910, after the introduction of transport airplanes in the military, the Army modified an airplane to demonstrate the ability to
accommodate medical patients. However, it was not until 1943 that Army nurses first trained in en route care procedures on airplanes while assigned to secure missions in North America, New Guinea, and India. During World War II, flight nursing expanded, with 20% of patients returning by air evacuation at a rate of nearly 100,000 patients transported per month. During the Korean War, helicopters were introduced to transport wounded from the point of injury to initial medical care. This “MEDEVAC” capability was expanded during the Vietnam War, with the use of a larger helicopter that allowed medical personnel to provide care during flight.

Following the establishment of a separate service, the Air Force was tasked in 1949 to provide long-range aeromedical evacuation (AE) for the entire U.S. military. Staffing consisted of specially trained nurses and medical technicians who provided medical care during flight. Currently a standard AE crew consists of two flight nurses and three AE technicians, although this staffing pattern can be tailored to meet mission requirements. Before 1994, any patient requiring intensive care unit–level care during transport required a dedicated medical attendant from the sending facility with critical care expertise. To eliminate the requirement to send this scarce resource with the patient, a dedicated Critical Care Air Transport Team (CCATT) capability was created with a physician, critical care nurse, and respiratory therapist that could be tasked to augment an AE crew using standard AE equipment. Following a successful demonstration of the capability, a dedicated CCATT training program began in 1999.

Similar to the evolution of CCATT, Army critical care nurses were specifically assigned to provide en route care in Iraq and Afghanistan starting in 2003 and the U.S. Navy recognized the importance of assigning a dedicated en route care nurse to Forward Resuscitative Surgery Systems in Afghanistan to transport critically injured patients. With the demonstrated value of the en route critical care nurse capability, the Joint Enroute Care Course was developed in 2006 by the Army to train nurses from all services to provide intensive care unit–level care aboard helicopters. However, even after the development of the Joint Enroute Care Course, predeployment training varied among services. This became apparent as en route care nurses deployed individually, joining existing units whose training was distinctly different than their own.

Stresses of En Route Care

Serving in an austere combat environment is stressful. From the moment nurses enter the area until they return home, there is no reprieve. Loud noises and long hours compound the weariness of the military nurses. These nurses are present at some of the most searing moments of life and sometimes the final ones. Exposure to polytrauma and death places these nurses at risk for developing behavioral health symptoms (BHS) such as anxiety, depression, or post-traumatic stress disorder (PTSD).

Other stresses multiply the chance of BHS. On helicopters, the en route care nurse and the medic are the ones making critical decisions. The area is cramped. Environmental conditions can include temperatures above 120°F, providing care in the dark and during evasive movement because of enemy fire. Stressors for AE and CCATT nurses include long hours, “barometric pressure changes, hypoxia, noise, vibration, gravitational forces, thermal stress, and dehydration.”

Previous studies have reviewed military nurses overall experiences of deployment and reintegration. Participants in those studies overwhelmingly reported issues with behavioral health and reintegration. No studies were identified that targeted en route care nurses and post-deployment BHS. Therefore, the purpose of this study was to better understand the post-deployment BHS and readjustment/reintegration experienced by military nurses who provided en route care while serving in Operation Iraqi Freedom/Operation Enduring Freedom. This article pertains to the research question: How do military nurses’ narratives extend, refute, or illuminate the quantitative outcomes from the Post-Deployment Readjustment Inventory (PDRI)?

METHODS

Design and Recruitment

This analysis is part of a larger exploratory, concurrent, mixed-methods study consisting of several valid and reliable behavioral health instruments generally used with military populations. Data were gathered through an electronic survey and single, one-on-one interviews with a subset of the sample. A purposive, snowballing method was applied to locate the nurses. Military nurse researchers from the Air Force, Army, and Navy assisted in advertising the study. Flyers explaining the study were disseminated through multiple venues among military hospitals and units where the nurses might be working. Participants were asked to refer others to the study.

Procedures

The study was approved by the Northern Regional Medical Command Institutional Review Board (control no. 392375-1). The electronic survey was approved by the Defense Manpower Data Center (control no. DD-HA(OT)2546). En route care was defined as the health care that military nurses provided for military or civilian personnel as they were transported from lower levels of care (the battlefield) to higher levels of care (combat support hospitals/hospitals outside the area of conflict). Inclusion criteria for the study were (1) age ≥18 years, (2) deployed to Operation Iraqi Freedom/Operation Enduring Freedom for at least 30 days, (3) completed at least one en route care mission while deployed, (4) completed post-deployment re-integration processing, and (5) understood, read, and spoke English. Nurses were
excluded if they had a behavioral health diagnosis before their first deployment.

Completion of the survey instruments via SurveyMonkey (SurveyMonkey, Inc, San Mateo, California) implied informed consent for the quantitative portion of the study. All survey data were anonymous. Upon survey completion, a page provided contact information to volunteer for an interview. Verbal informed consent was used with the interviews. Interviews were conducted face to face, by Skype, or by telephone and digitally recorded. Interview numbers and pseudonyms were used to protect participant confidentiality; all collected data were stored securely.

**Mixed-Methods Analytic Approach**
Quantitative results and qualitative findings were analyzed separately. Statistical results were reported followed by specific exemplars that support, refute, and/or illuminate the quantitative outcomes.

**Quantitative Data**
As part of the electronic survey, participants completed the “Demographic and Personal Military Characteristics Instrument” (28 questions) designed by the research team to collect essential demographic information about the military nurses’ personal characteristics, military deployments, and behavioral health status. Questions about readjustment/reintegration were measured by the “PDRI”, a 36-item instrument comprising five subscales and scored on a 5-point Likert type scale (1 = “Not at all true” to 5 = “Extremely true”) designed to gather information regarding service members’ post-deployment experiences. Scores are obtained by summing the items of the five subscales. The scale and subscale score ranges are as follows: Total Score (36–180), Career Challenges (5–25), Social Difficulties (7–35), Intimate Relationship Problems (5–25), Health Concerns (5–25), Concerns About Deployment (6–30), and PTSD Symptoms (8–40). The PDRI has a Cronbach’s $\alpha$ of 97 and was useful in detecting unique patterns of readjustment depending on war stressors.\(^{17}\)

**Quantitative Data Analysis**
A power analysis using G*Power (Heinrich-Heine University, Dusseldorf, Germany) was conducted with the following parameters: $\alpha = 0.05$, power = 0.8, and an anticipated medium effect size ($r = 0.3$). Following Cohen’s\(^{18}\) recommendations, a medium effect size was used because no prior studies were available. Based on the power analysis, $n = 85$ would be adequate to answer the research questions. SPSS v. 22 (IBM Corporation, Armonk, New York) was used to analyze the data. Descriptive statistics were used to examine sociodemographic information about the participants and data from the PDRI and a 1-way analysis of variance with transformed scores were used to analyze potential differences between the three services. Given the negative skew of PDRI subscale scores and global scores, data were logarithmically transformed to determine significance.

**Qualitative Data**
Participant interviews were derived from an existential, phenomenological design as described by Thomas and Pollio\(^{19}\) based on the philosophical works of Maurice Merleau-Ponty.\(^{20}\) Interviews, lasting 25 to 90 minutes, were conducted in a location selected by the participant. The opening question: “When you think about the en route care you provided while in Iraq and/or Afghanistan, what stands out to you about how it may have impacted your own behavioral health?” was followed by additional probing questions to clarify information provided during the interviews.

**Qualitative Data Analysis**
Two methods were used to complete the analysis process. First, the interviews were entered into NVivo, Version 9 (qualitative software) (QSR International, Inc, Burlington, MA,) and second, members of the research team completed line-by-line analysis of each transcript to identify key words and common phrases within and across transcripts.

**RESULTS**

**Demographic Results**
Tables I and II depict demographic characteristics relating to the electronic survey. A total of 125 nurses attempted the survey, with 119 completed surveys returned. Most of the nurses were Air Force, married, Caucasian, had over 6 years of military service, and possessed at least a bachelor’s degree. The number of males and females was nearly equal. The majority of the nurses deployed as individuals ($n = 75$,

**TABLE I.** Demographic Characteristics of Survey Participants

<table>
<thead>
<tr>
<th>Variable</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>58</td>
<td>48.7</td>
</tr>
<tr>
<td>Male</td>
<td>61</td>
<td>51.3</td>
</tr>
<tr>
<td>Marital Status</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single</td>
<td>11</td>
<td>9.2</td>
</tr>
<tr>
<td>Significant Other</td>
<td>5</td>
<td>4.2</td>
</tr>
<tr>
<td>Married</td>
<td>85</td>
<td>71.4</td>
</tr>
<tr>
<td>Separated</td>
<td>6</td>
<td>5.0</td>
</tr>
<tr>
<td>Divorced</td>
<td>12</td>
<td>10.1</td>
</tr>
<tr>
<td>Spanish, Hispanic, Latino</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>7</td>
<td>5.9</td>
</tr>
<tr>
<td>No</td>
<td>112</td>
<td>94.1</td>
</tr>
<tr>
<td>White</td>
<td>102</td>
<td>85.7</td>
</tr>
<tr>
<td>Black/African American</td>
<td>6</td>
<td>5.0</td>
</tr>
<tr>
<td>Indian/Alaskan</td>
<td>3</td>
<td>2.5</td>
</tr>
<tr>
<td>Asian</td>
<td>8</td>
<td>8.0</td>
</tr>
<tr>
<td>Current Level of Education</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Associates</td>
<td>4</td>
<td>3.4</td>
</tr>
<tr>
<td>Bachelors</td>
<td>63</td>
<td>52.9</td>
</tr>
<tr>
<td>Masters</td>
<td>51</td>
<td>42.9</td>
</tr>
<tr>
<td>Doctoral</td>
<td>1</td>
<td>8</td>
</tr>
</tbody>
</table>
63%) and returned by themselves (n = 78, 65.5%) rather than as part of a military unit. Four demographic questions queried nurses about their need and use of behavioral health assistance after deployment; 39 (32.8%) of respondents actively sought assistance, although 31 (26.1%) actually used military behavioral health services (Table III).

### Interview Results

Thirty-three nurses originally indicated they wanted to participate in the interview. However, participant job constraints and military moves affected the interview process; 22 interviews were completed. Seventeen interviews were conducted face to face and five were conducted by Skype/telephone because of location and/or job constraints (Table IV).

Based on the qualitative analysis, seven themes emerged from the nurses’ narratives. Four of the seven themes are relative to the PDRI subscales. The themes are as follows: “Leadership Matters,” “I Don’t Fit In,” “Here is my Suffering,” and “The Terror of War—You can’t Unsee That.”

### PDRI Results

As summarized in Table V, the mean total score for the PDRI was 76.08 (SD = 27.80); scores ranged from 35 to 154. Statistically significant differences were noted among services in all subscales except Intimate Relationship Problems as well as PDRI global scale scores. In all instances, the total score and subscales for Army respondents were higher than those of the Force and Navy.

**Career Challenges, “Leadership Matters”**

In the PDRI, 81 of the nurses (68%) answered they felt pressured to work, which aligned most closely with the qualitative theme “Leadership Matters.” Not only did leadership matter during deployment, it also influenced the nurses’ return to work. It did not matter that most nurses had been gone 6 or more months or that they had worked extremely longs hours without time off, nursing leaders sent them back to work shortly after returning home.

“You’ve been gone for a year. You’ve dealt with all of this trauma but we are really short . . . so we are going to give you two weeks of leave and send you back to work.”

“I was anticipating coming back in December, [but] I was sent home in November . . . they gave me a four-day pass for Thanksgiving. I came back and just started working . . . they gave me Christmas off and I worked until I took leave in January.”

The nurses also spoke of the perception of how leaders demonstrated uncaring behaviors during their reintegration.

“[When you reintegrate] You have to have all of this stuff filled out. Nobody gives a shit. Nobody wants that piece of paper when you are done. They make you fill out all of this reintegration paperwork and go and get all of these signatures who nobody cares, nobody does shit for you. Why make me do that before I can go take leave and be with my family?”
Understanding Postdeployment Reintegration Concerns Among En Route Care Nurses

<table>
<thead>
<tr>
<th>Social Difficulties, “I Don’t Fit In”</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instrument responses from 46 nurses (39%) indicated they were not fitting in socially; 73 nurses (61%) indicated “I’ve changed or others have changed”; and 61 nurses (51%) stated they were feeling pressure to be back to normal. During the interviews, nurses described the difficulty they had trying to fit back into who they were predeployment, wanting to get back to normal. They began to realize that was not quite possible because they were not the same, they changed during deployment. They also realized that some things were not as important anymore.</td>
</tr>
</tbody>
</table>

| “My family and friends … I just didn’t want to talk to them about it because a lot of times I would think that they wouldn’t understand.” |
| “Work colleagues], “they just don’t understand. They have no clue what you did, what you saw.” |
| Switching from combat roles to family roles was more difficult than the nurses anticipated, especially those families with children. Some of the struggles came from the care provided to the injured children. Some of their obstacles related to the scattered demands of the family. |

| [It took] “three to six months to reengage [with my children] … making a constant effort to cuddle with them … to spend more time with them.” |
| “I was gone for 10 months so she [my daughter] was a couple of months shy of two when I came home and that very night she wouldn’t let me hold her.” |
| “Getting back into the role with family was a bit difficult, dealing with all of the different demands.” |

<table>
<thead>
<tr>
<th>Health Concerns, “Here Is My Suffering”</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instrument items “my body not functioning like it use to” (n = 20, 33%) and “having lots of medical appointments” (n = 16, 14%) were also congruent with the theme “Here is my suffering.” The nurses did comment on physical pain, but their narratives were interwoven with the psychological pain they suffered. One nurse highlighted how his body function was different after his deployment experience while another told of the numerous medical visits encountered over the past year.</td>
</tr>
</tbody>
</table>

| “I will sit and see my kids and I’ll think back to those kids [we cared for during deployment] and I don’t really have control of my emotions anymore.” |
| “I spent six weeks in an intensive outpatient group for PTSD. I’ve seen a psychiatrist monthly and a psychologist at least two times a month. I’ve been in the step-down PTSD group twice a week for the better part of the year now. I’m still dealing with the issues. I take more medication than anybody should have to take.” |

<table>
<thead>
<tr>
<th>Social Concerns</th>
</tr>
</thead>
<tbody>
<tr>
<td>16.05 (6.48)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Intimate Relationship Problems</th>
</tr>
</thead>
<tbody>
<tr>
<td>9.77 (4.46)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Health Concerns</th>
</tr>
</thead>
<tbody>
<tr>
<td>9.35 (4.68)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Concerns About Deployment</th>
</tr>
</thead>
<tbody>
<tr>
<td>12.68 (4.98)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>PTSD Symptoms</th>
</tr>
</thead>
<tbody>
<tr>
<td>18.76 (7.95)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Total PDRI Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>76.08 (27.79)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TABLE V. Post-Deployment Readjustment Inventory Subscale Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Participants (N = 119)</td>
</tr>
<tr>
<td>---------------------------</td>
</tr>
<tr>
<td>Career Concerns</td>
</tr>
<tr>
<td>Social Concerns</td>
</tr>
<tr>
<td>Intimate Relationship Problems</td>
</tr>
<tr>
<td>Health Concerns</td>
</tr>
<tr>
<td>Concerns About Deployment</td>
</tr>
<tr>
<td>PTSD Symptoms</td>
</tr>
<tr>
<td>Total PDRI Score</td>
</tr>
</tbody>
</table>
Concerns About Deployment, “I Don’t Fit In”
Instrument responses from the nurses showed 49 (41%) reported everything seemed trivial since deployment and 59 (50%) were missing the structure and focus of being deployed. Several nurses highlighted how trivial typical things in life seemed to be, how things that were extremely important before deploying did not hold the same significance after coming home.

“It’s almost like things here can be so trivial sometimes after what you see down range.”

“I realized when I got back; I was no longer tolerant of the little stuff.”

Structure and focus, important tools in the combat environment seemed to disappear upon returning home. The nurses missed that part of their experiences and it caused difficulties in their reintegration.

“Order was important ... really important to have on the aircraft because everything is happening ... so you get use to that ... how important it is even though it’s not a big deal that the screwdriver is not in the drawer, it feels like a big deal. Lack of order, I just can’t stand it.”

PTSD Symptoms, “Terror of War—You can’t Unsee That”
Several instrument questions had relevance to this theme. They were having frequent thoughts about deployment (n = 44, 37%), having nightmares or difficulty sleeping (n = 47, 40%), feeling tense, jittery, or anxious (n = 59, 50%), and being easily irritated (n = 62, 52%). The nurses spoke of the difficulties of readjusting after deployment, how things they witnessed continued to play over and over in their minds. Some still did not feel safe at home.

“When you come home and try to readjust that’s the difficult part because the stuff that we’ve seen, the stress that you face when you are put in that position up in the plane or the helicopter in the middle of the night—all that stuff—you just can’t unsee that.”

“I was waking up at night, having constant thoughts of locations, seeing them in my nightmares, still thinking I was in danger.”

“When it is time to shut down and sleep ... your brain takes over ... start replaying everything ... it’s like a constant running reel of stuff in your head ... you think about it all of the time.”

Several of the nurses described the stress of the combat environment, how being in flight responsible for the life of others made you hypervigilant and suspicious of others, never letting your guard down, and how noises were causing problems.

“The stress of being in a hostile environment in flight, it makes you become hyper-vigilant. And you get suspicious of people.”

“When I came back I found I have sensitivity to noise. Loud noises kind of bother me.”

Several examples were provided regarding how easily the nurses became irritated with daily occurrences.

“I was very irritated by people’s small, petty whining.”

“Have someone pull out in front of you, when it’s your turn to go or cut in front of you in a line at the grocery store and you go ballistic. You just lose it.”

DISCUSSION
As noted by the findings of this study, the intricacies of post-deployment cannot be solely portrayed by a quantitative instrument. Qualitative interview data add a richness to understanding the experiences and emotions of the participants. For example, the PDRI Intimate Relationship subscale scores did not reflect serious problems among the nurses. However, the qualitative exemplars provided a robust picture of the difficulties the nurses encountered within family roles after deployment. Within the PTSD Symptoms subscale, several of the nurses scored high on feeling irritated, having nightmares, and/or difficulty sleeping, but without the qualitative interview data, one cannot grasp their experience of “terrors of war.” Unfortunately, these nurses are neither unique nor alone in their suffering. Previous works have revealed struggles similar to those noted in this study. These include hardships of returning to a noncombat environment, realizing they are a different person, missing the structure of deployment, trying to fit into previous relationships, dealing with leadership issues, and coping with behavioral health concerns.

Across the services, both men and women alike, the en route care nurses spoke vigorously about the inadequacy of the reintegration process and support. Their words were mirrored in the findings of the PDRI and support earlier studies.

Despite the differences in predeployment training based on their role and type of aircraft, as well as the level of care provided by en route care nurses in each of the services, there were commonalities in the qualitative responses of the participants. However, when analyzed for differences by service, the quantitative responses indicated statistical significant differences. It has been identified that differences in the diagnosis of PTSD can be attributed to many factors, especially combat exposure during deployment.

It is possible that Army en route care nurses were more likely to be assigned to a transport unit that responded to unsecure point-of-injury transport locations, which would be associated with higher rates of PTSD. Still, potential differences in combat exposure, perceived danger, or even level of leadership support by service were not noted in the qualitative comments.
Overall, the need for behavioral health support was high in the quantitative sample. Nearly 53% (n = 63) of respondents indicated needing to speak to someone after deployment. However, only 26% (n = 31) used Military Behavioral Health services since their return from deployment. This is a lower rate of behavioral health usage than the 35% identified from Army and Marine Corps service members returning from deployments to Iraq earlier in the conflict. Quantitative findings from this study support previous work identifying the stigma of seeking behavioral health assistance in the military. Qualitative exemplars of negative career impact after seeking behavioral health further explain the low use of Military Behavioral Health. These joint findings suggest stigma still exists. However, the appropriate use of behavioral health services in the military system should be viewed by nursing leaders as a healthy approach to dealing with post-deployment reintegration challenges.

**RECOMMENDATIONS**

Interventions to assist post-deployment reintegration of en route care nurses can be conducted at the peer, leader, and system levels. Participation from nursing leaders should be viewed by the military system as a necessity.11 Qualitative exemplars of behavioral health experiences in the military can be leveraged to reduce the stigma of seeking behavioral health care.11

Leaders also play a key role in helping en route care nurses in the reintegration process. It was found that leaders who had previously deployed often provided a more understanding and caring persona. Providing insight and awareness to new leaders or to those who had not been deployed to an en route care environment might provide help in understanding the redeploying en route care nurse’s experience to help facilitate reintegration. With an understanding of the en route care nurses’ experiences, leaders can aim to moderate the impact of long hours, stresses of combat, and other operational issues that can impact behavior health.

Embedding military mental health providers with en route care experience into en route care units would increase communication and awareness of the impact of the operational and combat-related stressors. These providers would not only provide easy access to mental health support for members of the en route care unit but they could also engage in consultation with commanders. A leadership and health care collaboration could foster a culture of health promotion with programs tailored to address healthy habits and physical and mental fitness.25

**LIMITATIONS**

This study was limited by the use of an electronic, self-report survey and the fact that the interview participants could not be linked to their survey responses. Although the study was designed in this manner to allow for greater anonymity, it limited the ability to interpret the narrative responses in the context of the PDRI scores.

**CONCLUSION**

En route care nurses’ struggle with reintegration/readjustment problems that typify those noted among combat warriors and other nurses who have deployed previously. Difficulties of this nature effect morale and military retention. Although there is clearly a need to develop protective measures to handle stressors of war or prolonged interaction with poly-trauma patients, we must contemplate the psychological impact of war and reintegration issues encountered by military nurses and gather the lessons learned to adequately prepare for future conflicts.

**ACKNOWLEDGMENTS**

This research was funded by the TriService Nursing Research Program (HT9404-13-1-TS06 (N13-P06).

**REFERENCES**


24. Hoge CW, Auchterlonie JL, Miliken CS: Mental health problems, use of mental health services, and attrition from military service after returning from deployment to Iraq or Afghanistan. JAMA 2006; 295(9): 1023–32.

Risk Factors for Incident Postdeployment Mental Health Conditions Among U.S. Air Force Medical Service Personnel

Genny M. Maupin, MPH†; Col Anthony P. Tvaryanas, USAF MC*; Edward D. White, PhD†; Heather J. Lysfjord, PhD*

ABSTRACT The prevalence of postdeployment mental health (PDMH) conditions in military health care personnel appears to be on par with that of other military personnel. However, there is no comprehensive analysis of incident PDMH conditions within the overall population of U.S. Air Force Medical Service personnel. This study explored the epidemiology of incident PDMH conditions among Air Force Medical Service personnel returning from deployment. A cohort survival analysis was conducted of 24,409 subjects without preexisting mental health conditions and at least one deployment during 2003–2013. Electronic health record data were used to ascertain the diagnosis of a PDMH condition. The primary outcome measure was an incident PDMH condition defined as a mental health diagnosis on at least two separate clinical encounters. The incidence of PDMH conditions was 59.74 per 1,000 person-years. Adjustment, anxiety, mood, sleep, and post-traumatic stress disorders accounted for 78% diagnoses. Protective factors included officer, surgeon, specific enlisted career fields, Air National Guard or Air Force Reserve, and multiple deployments. Risk factors included nurse, other specific enlisted career fields, female, and unmarried with dependents. Most subjects (73%) were diagnosed within the standard 30-month surveillance time period; median time to diagnosis was 13 months.

INTRODUCTION

Recent military conflicts in Iraq (Operation Iraqi Freedom), Afghanistan (Operation Enduring Freedom), and elsewhere have been associated with adverse postdeployment psychological events in military personnel, including post-traumatic stress disorder (PTSD), anxiety, depression, and alcohol misuse.1–3 Although a significant body of literature has evolved on the impact of recent conflicts on deployed military combat and combat support personnel, there has been scant information addressing the impact on military health care personnel. Recently, several of the authors conducted a retrospective cohort analysis that observed previously deployed U.S. Air Force Medical Service (AFMS) critical care personnel had a similar burden of postdeployment mental health (PDMH) conditions compared to that cited in the literature for personnel serving in combat-specific occupations.4 Moreover, these AFMS critical care personnel had a higher burden of PDMH conditions relative to burdens cited in the literature for other populations of military health care personnel. Within the cohort, women were determined to be at 1.4 times increased risk for incident PDMH conditions, and nurses and medical technicians were at twice the risk of physicians for incident PDMH conditions. Consequently, the purpose of this study was to expand upon the earlier analysis and to identify risk factors associated with an incident PDMH condition among the entire population of AFMS personnel returning from the deployed environment. Additionally, this study used a nested cohort analysis of those with an incident PDMH condition to explore factors associated with diagnosis outside the standard postdeployment medical surveillance period.

METHODS

Study Design

This study was conducted under a human use protocol approved by the 711th Human Performance Wing Institutional Review Board. The protocol included a waiver of informed consent of subjects because it used existing data that were routinely collected for other purposes. The study involved a retrospective cohort comprising the entire population of AFMS personnel with no history of preexisting mental health conditions who had at least one deployment during 2003–2013. Using a “one-shot case study” design,5 this cohort was followed, from the time of the first deployment until censure, for the occurrence of the outcome of interest: an incident PDMH condition.

Subjects

Inclusion Criteria

Air Force Personnel Center (AFPC) data were used to define the study population of AFMS personnel. AFMS personnel who were on active duty, Air National Guard, or Air Force Reserve status between 2003 and 2013 (inclusive) and whose Air Force Specialty Code (AFSC) began with the number 4 were selected (N = 117,844). Of note, some subjects cross-trained either into or out of a medical AFSC during the study period. Subjects who completed a deployment when they carried a medical AFSC were included in the study (N = 32,354).
Risk Factors for Incident PDMH Conditions Among Medical Personnel

Exclusion Criteria
Data from the Military Health System Data Mart (M2), a database of medical encounters from military treatment facilities and civilian network providers of out- and inpatient care, were used to ascertain whether members had a diagnosis of a mental health condition before deployment. A mental health condition was defined as an International Classification of Diseases, 9th Revision, Clinical Modification code between 290 and 319 (inclusive), excluding 305.1 for tobacco use disorder. A single mental health diagnosis was sufficient to exclude a service member as having a preexisting mental health condition (N = 7,067)—i.e., the threshold to screen service members out of the study cohort was set low as no confirmatory diagnosis was required. Additionally, service members without any M2 medical data were excluded (N = 878), for a total study population of 24,409. The date of first deployment was used to set the value of certain independent variables.

Measurements

Independent (Exposure and Risk Factor) Variables
Once the cohort was identified, the following variables were collected from the AFPC data: Social Security number (SSN) for merging purposes, rank, AFSC, service component, gender, date of birth (DOB) to calculate age, marital status, total number of dependents, deployment dates, and deployment locations. These variables were subsequently refined into the following set of potential risk and protective factors:

1) Rank (enlisted or officer)
2) Career field (identified by AFSC)
   a) Officers: aerospace medicine physician, biomedical clinician, biomedical specialist, dentist, health services administrator, medical commander, nonsurgical/ nonaerospace medicine physician, nurse, surgeon, and missing (i.e., no data)
   b) Enlisted: aeromedical, aerospace and operational physiology, bioenvironmental engineering, cardiological laboratory, dental assistant, diagnostic imaging, diet therapy, health services management, medical laboratory, medical service, mental health services, ophthalmic, pharmacy, physical medicine (including orthotics), public health, and missing (i.e., no data)
3) Service component (active duty, Guard, or Reserve)
4) Gender (male or female)
5) Age (transformed into quartiles: ≤24, 25–30, 31–38, and ≥39)
6) Marital status (divorced, single, married, or other)
7) Total number of dependents (0 or ≥1)
8) Total number of deployments (1 or ≥2)
9) Location of first deployment (Afghanistan, Germany, Iraq, Kuwait, Qatar, the United States, other, or unknown/classified)

As constructed, career field was conditional on rank (i.e., individual AFSCs were specific to either enlisted or officer ranks). For those variables that could change over time, the values were ascertained at the time of each subject’s first deployment.

Dependent (Outcome) Variable
M2 data were used to ascertain the outcome of interest: diagnosis of a mental health condition. Relevant International Classification of Diseases, 9th Revision, Clinical Modification codes ranged from 290 to 319 (excluding 305.1). M2 data fields utilized included SSN for merging purposes, service member DOB to confirm data were for the member versus an associated beneficiary, date(s) of care, and diagnosis codes. Once M2 data were merged with AFPC data using SSN and DOB, and age was calculated, SSN and DOB were removed from the dataset. Dates of care were used to ensure diagnoses were subsequent to the first deployment. Subjects needed to have a mental health diagnosis on at least two separate encounters to be classified as having an incident PDMH condition—i.e., an isolated diagnosis was insufficient for classification as the outcome of interest. This rule was implemented to improve the specificity of ascertaining an incident PDMH condition.

Statistical Analysis
All the statistical analyses were performed with JMP Version 11 (SAS Institute Inc., Cary, North Carolina) and IBM SPSS Modeler Version 14 (IBM Corp., Armonk, New York). The Kaplan–Meier method was used to calculate the actuarial rate of incident PDMH conditions. Subjects began contributing person-years of observation at the end of their first deployment. Subjects were censored at the time of the diagnosis of an incident PDMH condition or the time of last available medical data. Contingency table analyses were used to compare response rates across levels within categorical risk factor variables with three or more levels. Levels were collapsed where no statistically significant interlevel differences were observed. A multivariable time-dependent proportional hazard Cox’s model was used to identify individual and occupational factors associated with incident PDMH conditions. A nested analysis of subjects with an incident PDMH condition used a multivariable logistic regression model to identify individual and occupational factors associated with diagnosis beyond the standard 30-month postdeployment medical surveillance period. Potential exploratory variables included rank, career field, service component, gender, age, marital status, total number of dependents, total number of deployments, and location of first deployment. The threshold for statistical significance was set a priori at p = 0.001 given this study’s large N.

RESULTS

Study Population Descriptive Statistics
The characteristics of the overall study population are summarized in Table I. The study population comprised 24,409 subjects
with an aggregate time of follow-up of 92,919 person-years, averaging 3.8 years per subject. A total of 5,551 subjects (22.74%) developed an incident PDMH condition for an overall population incidence rate of 59.74 per 1,000 person-years. Table II summarizes the frequency of observed mental health diagnoses. For any subject, more than one mental health diagnosis may have been assigned during the period of observation covered by this study. Accordingly, a subject may have contributed multiple diagnoses but not multiple counts for the same diagnosis. The diagnosis categories of

### TABLE I. Cohort Demographic Characteristics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Incident PDMH Condition</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Frequency</td>
<td>Percent</td>
</tr>
<tr>
<td>Rank</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Officer</td>
<td>1,704</td>
<td>30.70</td>
</tr>
<tr>
<td>Enlisted</td>
<td>3,847</td>
<td>69.30</td>
</tr>
<tr>
<td>Career Field</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Officer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nonsurgical Physician</td>
<td>284</td>
<td>5.12</td>
</tr>
<tr>
<td>Nurse</td>
<td>838</td>
<td>15.10</td>
</tr>
<tr>
<td>Other</td>
<td>529</td>
<td>9.53</td>
</tr>
<tr>
<td>Surgeon</td>
<td>53</td>
<td>0.95</td>
</tr>
<tr>
<td>Missing</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>Enlisted</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enlisted Group 1</td>
<td>265</td>
<td>4.77</td>
</tr>
<tr>
<td>Enlisted Group 2</td>
<td>489</td>
<td>8.81</td>
</tr>
<tr>
<td>Enlisted Group 3</td>
<td>456</td>
<td>8.21</td>
</tr>
<tr>
<td>Health Services Management</td>
<td>772</td>
<td>13.91</td>
</tr>
<tr>
<td>Medical Service</td>
<td>1,865</td>
<td>33.60</td>
</tr>
<tr>
<td>Missing</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>Service Component</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Active Duty</td>
<td>4,925</td>
<td>88.72</td>
</tr>
<tr>
<td>Air Force Reserve</td>
<td>244</td>
<td>4.40</td>
</tr>
<tr>
<td>Air National Guard</td>
<td>382</td>
<td>6.88</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>2,404</td>
<td>43.31</td>
</tr>
<tr>
<td>Male</td>
<td>3,147</td>
<td>56.69</td>
</tr>
<tr>
<td>Age (Years)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤24</td>
<td>1,474</td>
<td>26.55</td>
</tr>
<tr>
<td>25–30</td>
<td>1,296</td>
<td>23.35</td>
</tr>
<tr>
<td>31–38</td>
<td>1,536</td>
<td>27.67</td>
</tr>
<tr>
<td>≥39</td>
<td>1,245</td>
<td>22.43</td>
</tr>
<tr>
<td>Marital Status</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Divorced</td>
<td>539</td>
<td>9.71</td>
</tr>
<tr>
<td>Married</td>
<td>3,321</td>
<td>59.83</td>
</tr>
<tr>
<td>Single</td>
<td>1,680</td>
<td>30.26</td>
</tr>
<tr>
<td>Other</td>
<td>11</td>
<td>0.20</td>
</tr>
<tr>
<td>Total Dependents</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>2,456</td>
<td>44.24</td>
</tr>
<tr>
<td>1 or more</td>
<td>3,095</td>
<td>55.76</td>
</tr>
<tr>
<td>Total Deployments</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>3,899</td>
<td>70.24</td>
</tr>
<tr>
<td>2 or more</td>
<td>1,652</td>
<td>29.76</td>
</tr>
<tr>
<td>Deployment Location</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Afghanistan</td>
<td>892</td>
<td>16.07</td>
</tr>
<tr>
<td>Germany</td>
<td>364</td>
<td>6.56</td>
</tr>
<tr>
<td>Iraq</td>
<td>1,577</td>
<td>28.41</td>
</tr>
<tr>
<td>Kuwait</td>
<td>289</td>
<td>5.21</td>
</tr>
<tr>
<td>Qatar</td>
<td>477</td>
<td>8.59</td>
</tr>
<tr>
<td>United States</td>
<td>284</td>
<td>5.12</td>
</tr>
<tr>
<td>Other</td>
<td>1,130</td>
<td>20.36</td>
</tr>
<tr>
<td>Unknown or Classified</td>
<td>538</td>
<td>9.69</td>
</tr>
</tbody>
</table>

PDMH, post-deployment mental health. “Enlisted group 1: aeromedical, mental health services, and pharmacy; Enlisted group 2: aerospace and operational physiology, bioenvironmental engineering, dental assistant, diagnostic imaging, and ophthalmic; Enlisted group 3: cardiopulmonary laboratory, diet therapy, medical laboratory, physical medicine (including orthotics), and public health.
adjustment, anxiety, mood, sleep, and post-traumatic stress disorders accounted for 78.45% of the total observed diagnoses.

**Risk Factors for Incident Mental Health Conditions**

The Kaplan–Meier estimate of the survival function from first deployment return date until diagnosis of an incident PDMH condition is shown in Figure 1. Table III displays the results of the Cox regression analysis. Significant protective factors included being an officer; a surgeon; in one of the enlisted career fields of aerospace and operational physiology, bioenvironmental engineering, dental assistant, diagnostic imaging, or ophthalmic; a member of the Guard or Reserve component; and having greater than one deployment. Significant risk factors included being a nurse; in one of the enlisted career fields of aeromedical, mental health services, pharmacy, or medical service; female; and unmarried with dependents.

**Time to Diagnosis of Mental Health Condition**

This analysis was a nested cross-sectional study of the 5,551 subjects who were diagnosed with an incident PDMH condition. The boxplot of the time interval from the end of the most recent deployment preceding the diagnosis of an incident PDMH diagnosis to assignment of the diagnosis is shown in Figure 2. The distribution was right skewed with a median time to diagnosis of an incident PDMH condition of 13 months (interquartile range 4–32 months). There were 4,051 (72.98%) subjects diagnosed during the standard 30-month postdeployment medical surveillance period and 1,500 (27.02%) subjects diagnosed outside that period. The results of the final fitted logistic regression model for the outcome of diagnosis beyond the standard 30-month postdeployment medical surveillance period are summarized in Table IV; the area under the receiver-operating characteristic curve measuring the accuracy of the model predicting the response was 0.633. Subjects in the Air National Guard had a greater likelihood of presenting with an incident PDHM condition outside the standard medical surveillance period. Subjects who were female, less than 25 years of age, with more than one deployment, or with an initial deployment to Afghanistan, Germany, or Iraq, were more likely to present within the standard medical surveillance period. Clinically relevant factors, based on large effect sizes,
included initial deployment to Afghanistan (odds ratio 0.34) and being in the Guard (odds ratio 1.82).

DISCUSSION

This study is the first comprehensive analysis of incident PDMH conditions in AFMS personnel returning from the deployed environment. Unlike other studies addressing mental health conditions in veteran populations using self-administered questionnaires, this study used actual diagnoses as derived from electronic health records (EHRs). Additionally, ascertainment of both exposures and outcomes through data mining allowed this study to include all AFMS personnel who could be identified from existing record systems.

The incidence of diagnosed PDMH conditions in the population of previously deployed AFMS personnel was 59.74 per 1,000 person-years. This incidence is 2.4 times that observed by the authors in their prior study of a sample of AFMS critical care personnel.4 The prevalence of diagnosed PDMH conditions in the study population was 22.74%, which is approximately two-thirds the prevalence reported by Seal et al6 for Iraq and Afghanistan veterans. In contrast, Hoge et al1 and Milliken et al2 reported a prevalence of positive screens for PDMH conditions among various subpopulations of military personnel returning from deployment to Iraq or Afghanistan of 11 to 19% and 20 to 42%, respectively. Points of comparison with other populations of military health care personnel were limited. In a small sample (n = 102) of health care personnel within one military hospital, Kolkow et al7 estimated the prevalence of PTSD at 9% and depression at 5% among those who had previously deployed. Jones et al8 reported a prevalence of psychological distress of 25% in the U.K. armed forces medical personnel who served in the 2003 Iraq war. Thus, it was concluded that this study population did not have an excessive burden of mental health conditions relative to the larger population of military personnel.

Five conditions composed three-quarters of the PDMH diagnoses in the study population: adjustment, anxiety (excluding PTSD), mood, and sleep disorders and PTSD. This pattern was similar to that observed in the authors’ prior study of AFMS critical care personnel.4 Contrary to other studies of various subpopulations of military personnel returning from deployment to Iraq or Afghanistan,2,6,8,9 this study did not find a high frequency of alcohol-related disorders.

The survival analysis identified potential protective and risk factors for incident PDMH conditions in AFMS personnel returning from deployment to Iraq or Afghanistan of 11 to 19% and 20 to 42%, respectively. Points of comparison with other populations of military health care personnel were limited. In a small sample (n = 102) of health care personnel within one military hospital, Kolkow et al7 estimated the prevalence of PTSD at 9% and depression at 5% among those who had previously deployed. Jones et al8 reported a prevalence of psychological distress of 25% in the U.K. armed forces medical personnel who served in the 2003 Iraq war. Thus, it was concluded that this study population did not have an excessive burden of mental health conditions relative to the larger population of military personnel.

Five conditions composed three-quarters of the PDMH diagnoses in the study population: adjustment, anxiety (excluding PTSD), mood, and sleep disorders and PTSD. This pattern was similar to that observed in the authors’ prior study of AFMS critical care personnel.4 Contrary to other studies of various subpopulations of military personnel returning from deployment to Iraq or Afghanistan,2,6,8,9 this study did not find a high frequency of alcohol-related disorders.

The survival analysis identified potential protective and risk factors for incident PDMH conditions in AFMS personnel returning from deployment to Iraq or Afghanistan of 11 to 19% and 20 to 42%, respectively. Points of comparison with other populations of military health care personnel were limited. In a small sample (n = 102) of health care personnel within one military hospital, Kolkow et al7 estimated the prevalence of PTSD at 9% and depression at 5% among those who had previously deployed. Jones et al8 reported a prevalence of psychological distress of 25% in the U.K. armed forces medical personnel who served in the 2003 Iraq war. Thus, it was concluded that this study population did not have an excessive burden of mental health conditions relative to the larger population of military personnel.

### TABLE III. Cox Regression Model for the Hazard of an Incident Post-Deployment Mental Health Condition

<table>
<thead>
<tr>
<th>Variable</th>
<th>Regression Coefficient</th>
<th>Standard Error</th>
<th>Hazard Ratio</th>
<th>95% CI for Hazard Ratio</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rank</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Officer</td>
<td>−0.429</td>
<td>0.0454</td>
<td>0.651</td>
<td>0.595</td>
<td>0.711</td>
</tr>
<tr>
<td>Career Field&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enlisted Group 1</td>
<td>0.310</td>
<td>0.0677</td>
<td>1.363</td>
<td>1.191</td>
<td>1.553</td>
</tr>
<tr>
<td>Enlisted Group 2</td>
<td>−0.220</td>
<td>0.0535</td>
<td>0.803</td>
<td>0.722</td>
<td>0.891</td>
</tr>
<tr>
<td>Enlisted Medical Service</td>
<td>0.317</td>
<td>0.0371</td>
<td>1.373</td>
<td>1.277</td>
<td>1.477</td>
</tr>
<tr>
<td>Nurse</td>
<td>0.510</td>
<td>0.0506</td>
<td>1.665</td>
<td>1.508</td>
<td>1.839</td>
</tr>
<tr>
<td>Surgeon</td>
<td>−0.580</td>
<td>0.1418</td>
<td>0.560</td>
<td>0.419</td>
<td>0.731</td>
</tr>
<tr>
<td>Service Component</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Guard</td>
<td>−0.665</td>
<td>0.0536</td>
<td>0.514</td>
<td>0.462</td>
<td>0.570</td>
</tr>
<tr>
<td>Reserve</td>
<td>−0.600</td>
<td>0.0663</td>
<td>0.549</td>
<td>0.481</td>
<td>0.624</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>0.201</td>
<td>0.0281</td>
<td>1.222</td>
<td>1.157</td>
<td>1.291</td>
</tr>
<tr>
<td>Total Deployments</td>
<td>−0.843</td>
<td>0.0301</td>
<td>0.430</td>
<td>0.406</td>
<td>0.456</td>
</tr>
<tr>
<td>Unmarried With Dependents</td>
<td>0.219</td>
<td>0.0507</td>
<td>1.245</td>
<td>1.126</td>
<td>1.373</td>
</tr>
</tbody>
</table>

CI, confidence interval. <sup>a</sup>Enlisted group 1: aeromedical, mental health services, and pharmacy; Enlisted group 2: aerospace and operational physiology, bioenvironmental engineering, dental assistant, diagnostic imaging, and ophthalmic.
Risk Factors for Incident PDMH Conditions Among Medical Personnel

TABLE IV. Logistic Regression Model for Diagnosis Beyond the Standard Postdeployment Medical Surveillance Period

<table>
<thead>
<tr>
<th>Variable</th>
<th>Regression Coefficient</th>
<th>Standard Error</th>
<th>Odds Ratio</th>
<th>95% CI for Odds Ratio</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Service Component</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Guard</td>
<td>0.596</td>
<td>0.112</td>
<td>1.816</td>
<td>1.455 2.261</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>-0.291</td>
<td>0.064</td>
<td>0.748</td>
<td>0.659 0.848</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Age ≤24 Years</td>
<td>-0.466</td>
<td>0.076</td>
<td>0.628</td>
<td>0.540 0.728</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Total Deployments</td>
<td>-0.471</td>
<td>0.071</td>
<td>0.625</td>
<td>0.543 0.717</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Deployment Location</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Afghanistan</td>
<td>-1.076</td>
<td>0.106</td>
<td>0.341</td>
<td>0.276 0.418</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Germany</td>
<td>-0.451</td>
<td>0.134</td>
<td>0.637</td>
<td>0.487 0.825</td>
<td>0.0008</td>
</tr>
<tr>
<td>Iraq</td>
<td>-0.291</td>
<td>0.071</td>
<td>0.748</td>
<td>0.649 0.860</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

CI, confidence interval.

personnel returning from the deployed environment. Protective factors included being an officer; a surgeon; in one of the enlisted career fields of aerospace and operational physiology, bioenvironmental engineering, dental assistant, diagnostic imaging, or ophthalmic; a member of the Guard or Reserve component; and having multiple deployments. Risk factors included being a nurse; in one of the enlisted career fields of aeromedical, mental health services, pharmacy, or medical service; female; and unmarried with dependents. Clinically relevant factors, based on large effect sizes, included being a nurse (relative risk 1.67) and having greater than one deployment (relative risk 0.43).

The finding that Guard and Reserve status were protective factors was contrary to that of the authors’ prior study of AFMS critical care personnel as well as other studies of various subpopulations of military personnel returning from deployment to Iraq or Afghanistan with the study by Jones et al specifically examining military health care personnel. This finding may represent a differential misclassification or detection bias because Guard and Reserve personnel were more likely to be erroneously classified as not having an incident PDMH condition relative to active duty personnel. The potential for this bias exists because Guard and Reserve personnel receive health care outside the AFMS where diagnosis of an incident PDMH condition is not reliably captured in the EHRs used in this study.

The observation that personnel with multiple deployments were at decreased risk for incident PDMH conditions contrasted with the findings from the authors’ prior study, a cohort study in the U.K. armed forces—specifically examining military health care personnel—and the study of veterans by Seal et al. This finding may reflect the healthy worker effect where individuals in good health remain in the workplace and are at lower risk of illness than the larger population as a whole. As applied to this study, those personnel who were more vulnerable to the occupational stress of deployment might attrite from the AFMS after their first deployment, thereby leaving a cohort that was more resilient and who deployed multiple times as a result of their longevity in the workforce.

The observation that nurses were at increased risk for mental health conditions was consistent with the results of the authors’ prior study and that of Ben-Ezra et al. The latter conducted a cross-sectional study of post-traumatic symptoms among nurses and physicians in a general hospital targeted by missiles during the war between Lebanon and Israel in 2006. They found 23% of the hospital staff sampled exhibited clinically significant post-traumatic stress symptoms; nurses were five times more likely to manifest symptoms compared to physicians.

This study also explored the time from end of the most recent deployment to first diagnosis of an incident PDMH condition, providing valuable data for individuals designing postdeployment medical surveillance programs. It is worth noting that the optimal timing for when postdeployment health assessments should be administered has not been empirically determined. Data analysis from this study revealed the distribution of time from end of the most recent deployment to first diagnosis of an incident PDMH condition was positively skewed, with a median of 13 months and 27% of cases presenting beyond the 30-month medical surveillance period.

The observation of a lengthy delay from the end of the last deployment until initiation of mental health care is consistent with the study by Maguen et al, which found a median lag of 2.1 years in a cohort of veterans. The logistic regression analysis revealed that Guard status, gender, age, number of deployments, and deployment location were predictive of lag time, but the model was not robust enough to personalize medical surveillance requirements.

Strengths and Limitations

Both a strength and a weakness of this study was the ascertainment of the clinical end point of psychiatric effect using data extracted from EHRs. The EHR data yielded psychiatric diagnoses that were not constrained to those conditions for
which survey instruments exist, thereby providing a more comprehensive picture than that offered solely by examination of survey instruments administered in postdeployment health assessments. Additionally, EHR data allowed detection of cases outside the standard postdeployment surveillance period and assessment of time from exposure to diagnosis as a continuous variable, thereby yielding distributional data. However, it is known that documented diagnoses represent only visible cases—i.e., the iceberg phenomenon where people with diagnosed disease are a smaller subset of all people with the disease.11 Accordingly, this study’s estimate of the incidence of PDMH conditions likely underestimates the true incidence of PDMH conditions.

Another weakness of depending on EHR diagnoses is the potential lack of an objective correlation to the symptoms a service member is reporting. A clinician’s decision to enter a mental health diagnosis into the EHR is a complex one. For example, clinicians are aware of the potential implications of mental health diagnoses and may select diagnostic labels that allow them to address a service member’s presenting problem without triggering other downstream effects. The diagnosis of a mental health condition also relies heavily on reported symptoms, which may be exaggerated or fabricated for secondary gain or minimized or not reported at all. Overall, a mental health diagnosis is an imprecise assessment of a service member’s condition as it appears, and it is difficult to ascertain how this imprecision may have biased the study results. Also related to this issue of imprecision is the limitation that an antecedent deployment may have been unrelated to an incident PDMH condition, which should temper the interpretation of the study results. Diagnosis data alone provided insufficient information to assign attribution beyond temporal association.

As previously mentioned, a limitation of this study was the potential for a differential misclassification bias for Guard and Reserve personnel who sought health care outside the AFMS using their civilian health insurance. It was also possible, although less likely, that some active duty personnel sought off network care at their own expense to avoid potential scrutiny resulting from receiving care within a military treatment facility. In either situation, diagnoses resulting from off network care would not be reliably captured in the EHR data used in this study.

**Future Research**

Future research should determine the relative contribution of career field-unique exposures versus exposure to the deployed health care environment in general. These analyses should account for geographic information, including both deployed and nondeployed work locations. Future studies should also more directly ascertain the nature of wartime exposures (patient contact versus personal threat of injury or death) to assess the relative contribution of different exposures to the development of PDMH conditions. Additionally, postdeployment health assessment and reassessment data should be correlated with EHR data to explore the utility of current assessment instruments in predicting objective outcomes. Finally, future research should involve collaboration with researchers in the other military services to evaluate health care personnel across the four branches.

**CONCLUSIONS**

The present study has demonstrated that AFMS personnel returning from the deployed environment do not appear to have an excessive burden of mental health conditions relative to the larger population of military personnel serving in other occupations. While multiple protective and risk factors were identified, clinically significant differences in risk were observed for the nurse career field and multiple deployments.

**ACKNOWLEDGMENTS**

This research study was supported by Defense Health Program Operations and Maintenance funding.

**REFERENCES**

1. Hoge CW, Auckerlonie JL, Miliken CS: Mental health problems, use of mental health services, and attrition from military service after returning from deployment to Iraq or Afghanistan. JAMA 2006; 295(9): 1023–32.

MILITARY MEDICINE, Vol. 182, March/April Supplement 2017 257
Post-Traumatic Stress Symptoms in United States Air Force Aeromedical Evacuation Nurses and Technicians

Julie M. Swearingen, PhD*; Tanya M. Goodman, MS*; Wayne L. Chappelle, PsyD, ABPP†; William T. Thompson, MA*

ABSTRACT Critical Care Air Transport Teams (CCATT) are specialized military medical personnel who provide high-acuity care in an aeromedical environment. The rate of post-traumatic stress disorder (PTSD) symptoms was assessed in CCATT personnel and their rates were compared to general aeromedical evacuation (AE) personnel. As part of a computer-based occupational stress survey, 188 crew members (138 AE nurses and technicians, 50 CCATT nurses and respiratory therapists) completed the PTSD Checklist – Military Version. A categorical MEET/DOES NOT MEET CRITERIA variable was created, and a Fisher’s exact test was computed to identify differences between groups. Contingency table analyses were used to assess associations between demographic and occupational variables with meeting criteria. χ² or Fisher’s exact test results, relative risks, and 95% confidence intervals were obtained, with 4.35% of AE and 14.00% of CCATT crew members meeting PTSD symptom criteria. The CCATT crew members were 3.22 times (95% confidence interval = 1.14–9.12) more likely to meet symptom criteria than AE, and for CCATT meeting criteria, the most commonly endorsed symptoms were arousal and avoidance. The demographic and occupational factors assessed in this study were not associated with meeting PTSD symptom criteria. Current findings are discussed in relation to current research on post-traumatic stress in ground-based critical care personnel.

INTRODUCTION

Since the 1940s, the aeromedical evacuation (AE) system of the United States Air Force (USAF) has been a vital component of the military’s medical evacuation mission, efficiently transporting wounded, ill, and injured military members to and between medical facilities throughout various parts of the globe. According to Air Force Instruction 10-2912, the mission and capabilities of the AE system require these assets to provide “expedient evacuation of patients to save life, limb, and eyesight, prevent undue suffering, and preserve military strength.” These teams, traditionally consisting of two flight nurses and three aeromedical technicians, provide timesensitive en route care of up to 113 regulated, stabilized casualties using USAF organic and/or contracted fixed-wing aircraft with medical aircrew trained explicitly for this mission. AE forces can operate as far forward as fixed-wing aircraft are able to conduct air operations, across the full range of military operations, and in all operating environments.

The advent of the Critical Care Air Transport Team (CCATT) program in 1994 allowed for the transport of critically ill and injured patients in addition to the medically stabilized patients transported by general AE crews. When needed, CCATTs augment a standard AE team and comprise a critical care physician, critical care nurse, and respiratory therapist with supplies and equipment necessary to provide a critical care environment that would move with the patient during evacuation. The scope of CCATT care is designed to match that of an intensive care unit (ICU) in the field, and they are trained to actively manage high-acuity patients, including “multisystem trauma, burns, shock, and respiratory failure.”

The aerospace environment introduces unique demands and stressors both on providers and patients. AE and CCATT crew members must understand and compensate for factors such as dynamic changes in gravity and acceleration, fluctuating barometric pressures, excessive vibrations, low light, high levels of noise, limited medical supplies and resources, and crossing through multiple time zones. Not only do these providers require a working knowledge of the physiologic impacts of these factors, they must also understand the impact on their equipment and respond in real time to their constantly changing environment.

In addition to the demands encountered by nurses and technicians in ground-based hospitals, AE and CCATT crew members are exposed to a number of operational stressors. Although ground-based personnel must, in theory, be prepared to deploy at all times, for the past 8 years, the Air Force has planned deployment taskings according to a range of “deploy to dwell” ratios in order to add predictability for personnel. However, as limited, rapidly deployable assets, members of the AE and CCATT programs, are not constrained by these ratios and must truly be ready to deploy at all times. During an actual deployment experience, stressors include being deployed for months at a time in austere environments away from loved ones and in close quarters with other crew members, high operational workload and manning...
challenges (e.g., long work hours), organizational management issues (e.g., relational conflicts with coworkers or supervisors), and managing work–life balance (e.g., difficulty sustaining healthy sleep, diet, and exercise habits). In addition, maintaining aircrew training and continuing education relevant to patient care in the aerospace environment are challenging.1,8

AE and CCATT duties also require flexibility, cross-disciplinary knowledge, and teamwork to manage higher than typical patient to provider ratios in a resource-limited, dynamic environment. CCATT nurses are expected to oversee three to six patients simultaneously, compared to the one or two patients expected of ICU nurses.3,8 In contrast, physicians and respiratory therapists in CCATTs experience smaller patient-to-provider ratios than in traditional medical treatment facilities and ICUs and, therefore, are more directly involved in the care and monitoring of the patients and assist nurses as needed.3

For CCATT members, exposure to casualties of war and natural disasters beyond the scope of what is seen in traditional medical treatment facilities and ICUs is not uncommon. In a recent review of the epidemiology of patients requiring critical care air transport, explosive blasts were the most frequent mechanism of injury, and bilateral lower extremity amputation was the most common injury.6 In addition to managing high-acute, traumatically injured patients, the transport of acutely injured patients often takes place farther forward, requiring these medical personnel to carry weapons.5

Due to these unique environments, experiences, and expectations, many CCATT crew members are exposed to military-related, potentially traumatic events. However, little evidence regarding post-traumatic symptoms has been reported for critical care providers in general, or for CCATT personnel specifically. A recent study of CCATT personnel conducted by Tvaryanas and Maupin reported that 7.89% of CCATT crew members currently meet criteria for a diagnosis of post-traumatic stress disorder (PTSD). In the same study, only 5.71% of a comparison group consisting of physicians, nurses, and respiratory therapists without CCATT experience were diagnosed with PTSD.8

Differing methodologies between these prior studies prevent quantitative comparisons of their findings, but when compared to current point prevalence estimates for the civilian population (1–3%),9 the limited data suggest that medical personnel may be at elevated risk, especially those managing high-acute patients in the aerospace environment.

Although prior studies have included physician members of CCATT in their analyses, the AE program does not contain an equivalent physician member. In comparing these two specific aerospace medical environments, and in an attempt to eliminate potential confounding variables in terms of education and training experiences, CCATT physician data were not included in the analyses for the current study. The primary objectives of this study were to (a) measure the frequency and severity of self-reported PTSD symptoms between AE and CCATT nurses and medical technicians, (b) examine patterns and compare the rate of Diagnostic and Statistical Manual of Mental Disorders, fourth Edition (DSM-IV) PTSD symptom criteria10 between AE and CCATT personnel, and (c) assess for demographic and operational variables associated with an increased risk for meeting PTSD symptom criteria.

METHODS

Participants

A total of 188 active duty AE and CCATT crew members participated in the study. Of these, 138 (73.40%) were AE crew members (i.e., nurses and medical technicians) and 50 (26.60%) were CCATT crew members (i.e., nurses and respiratory therapists). PTSD symptoms were assessed as part of a larger occupational health survey for which the response rate was 40% (95% confidence interval [CI] = 32–48%) for AE and 37% (95% CI = 24–50%) for CCATT based on manning data from USAF operational leadership. The CCATT physicians were also included in the occupational health survey sample, but their responses were not included in the current analyses because AE crews do not have a physician counterpart.

Both AE and CCATT personnel completed similarly adapted formats of a web-based occupational health screening survey. This study was reviewed and approved by the Air Force Research Laboratory Institutional Review Board (approval for AE protocol: FWR20120183E; approval for CCATT protocol: FWR20130027E).

Instruments

Demographics Questionnaire

Respondents were asked to complete a personal and occupational demographics questionnaire that assessed medical profession, gender, age range, marital status, and average time spent in current duties. For AE crew members, those who indicated nurse as their medical profession were categorized as officers for rank range, and medical technicians were categorized as enlisted. For CCATT crew members, those who indicated nurse as their medical profession were categorized as officers for rank range, and respiratory therapists were categorized as enlisted. The survey did not request personally identifiable information from participants (e.g., Social Security number, name, or date of birth) to ensure anonymity and encourage candid responses. Demographics are shown in Table I.

PTSD Checklist – Military Version

The PTSD Checklist – Military Version (PCL-M) is a standardized questionnaire that assesses military-related PTSD symptomatology within the past month. The 17-item survey was developed using the PTSD criteria outlined in the DSM-IV.10,11 Respondents rate each item on a scale from 1 (“not at all”) to 5 (“extremely”). All items sum to yield a
**TABLE I.** Demographics for AE and CCATT Crew Members

<table>
<thead>
<tr>
<th>Demographics</th>
<th>AE (n = 132)</th>
<th>CCATT (n = 50)</th>
<th>Total (n = 188)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>75</td>
<td>32</td>
<td>107</td>
</tr>
<tr>
<td>Female</td>
<td>61</td>
<td>18</td>
<td>79</td>
</tr>
<tr>
<td>Age Range (Year)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18–25</td>
<td>16</td>
<td>11</td>
<td>27</td>
</tr>
<tr>
<td>26–30</td>
<td>38</td>
<td>6</td>
<td>44</td>
</tr>
<tr>
<td>31–35</td>
<td>32</td>
<td>6</td>
<td>38</td>
</tr>
<tr>
<td>36–40</td>
<td>19</td>
<td>11</td>
<td>30</td>
</tr>
<tr>
<td>41+</td>
<td>32</td>
<td>20</td>
<td>52</td>
</tr>
<tr>
<td>Marital Status</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single</td>
<td>60</td>
<td>17</td>
<td>77</td>
</tr>
<tr>
<td>Married</td>
<td>78</td>
<td>33</td>
<td>111</td>
</tr>
<tr>
<td>Rank Range</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enlisted</td>
<td>67</td>
<td>16</td>
<td>83</td>
</tr>
<tr>
<td>Officer</td>
<td>71</td>
<td>34</td>
<td>105</td>
</tr>
<tr>
<td>Time in Duties</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;1 Year</td>
<td>40</td>
<td>5</td>
<td>45</td>
</tr>
<tr>
<td>(\geq 1) Year</td>
<td>97</td>
<td>90</td>
<td>187</td>
</tr>
</tbody>
</table>

Valid percentages are reported. \(^a n = 138, \, ^b n = 50, \, ^c n = 188\).

Procedures

Although participation was advocated by military unit leadership via e-mails to USAF military accounts, the survey link directed participants to a nonmilitary Web site. The USAF leadership stated that the study was an occupational health screening that would be used to influence resiliency and morale and that participation was voluntary and anonymous. The AE survey was open to all AE personnel for a 13-week period and the CCATT survey was open to all CCATT personnel for a 9-week period. The surveys were made available to all AE and CCATT personnel.

It took an average of 25 to 30 minutes to complete the survey, and participants who completed the survey were instructed on how and when to obtain the summarized results of the study.

**TABLE II.** PCL-M Total Score and DSM-IV PTSD Symptom Criteria Endorsement for AE and CCATT Crew Members

<table>
<thead>
<tr>
<th>PCL-M Total Scores</th>
<th>AE Meeting DSM-IV PTSD Symptom Criteria (n) (% of AE overall)</th>
<th>CCATT Meeting DSM-IV PTSD Symptom Criteria (n) (% of CCATT overall)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low (17–36)</td>
<td>Does Not Meet (126) (91.30) Meets (0) (0.00)</td>
<td>Does Not Meet (37) (74.00) Meets (0) (0.00)</td>
</tr>
<tr>
<td>Mod (37–49)</td>
<td>Does Not Meet (6) (4.35) Meets (4) (2.90)</td>
<td>Does Not Meet (6) (12.00) Meets (3) (6.00)</td>
</tr>
<tr>
<td>High (50–85)</td>
<td>Does Not Meet (0) (0.00) Meets (2) (1.45)</td>
<td>Does Not Meet (0) (0.00) Meets (4) (8.00)</td>
</tr>
<tr>
<td>Total</td>
<td>Does Not Meet (132) (95.65) Meets (6) (4.35)</td>
<td>Does Not Meet (43) (86.00) Meets (7) (14.00)</td>
</tr>
</tbody>
</table>

Data Analysis

Frequencies and proportions for demographic and operational variables were obtained. Total score for the PCL-M was obtained by summing all-item responses. Total scores were separated into three categories—low (17–36), moderate (37–49), and high (50 or more)—to provide more information regarding score distributions in the sample. Fifty was chosen as the cutoff for the “high” category because a score of 50 or more has shown to have high specificity, sensitivity, and accuracy with correctly identifying those at a high risk for PTSD.\(^{12,14–17}\) The cutoff for the moderate category (37–49) was based on the cutoff used in a previous study by Chappelle et al\(^{18}\) and was based on the lowest observed PCL-M score for a participant meeting DSM-IV criteria for PTSD in that study.

Categorical variables were created for DSM-IV PTSD criteria B (re-experiencing symptoms), C (avoidance symptoms), and D (arousal symptoms). Respondents were then categorized into either a “Meets” or “Does Not Meet” group based on their responses within these categories. Individuals were considered in the “Meets” category if they endorsed one or more re-experiencing symptoms, three or more avoidance symptoms, and two or more arousal symptoms with a severity rating of 3 (moderately) to 5 (extremely). This is an accepted alternate methodology for scoring the PCL-M.\(^{19}\)

Descriptive statistics were obtained by the PCL-M total score ranges for individuals meeting these criteria. Percentages for PCL-M total score categories by the DSM-IV PTSD symptom criteria variable are shown in Table II. A Fisher’s exact test was computed to identify differences between groups on proportions of those meeting DSM-IV PTSD symptom criteria. See Table III for the frequency with which each PCL-M item was endorsed with a severity rating of moderately or higher for all respondents.

Contingency table analyses were conducted to assess associations between demographic and occupational variables (i.e., gender, age range, marital status, rank range, and time on station) with meeting the DSM-IV PTSD symptom criteria. \(\chi^2\) or Fisher’s exact test results, relative risks, and 95% CIs were obtained. Fisher’s exact test results were calculated in lieu of \(\chi^2\) results in instances where an expected cell count was five or less. Relative risks were obtained to...
TABLE III. Percentage of AE and CCATT Crew Members Self-Reporting PTSD Symptoms of Moderate to Extreme Severity

<table>
<thead>
<tr>
<th>PCL-M Items</th>
<th>DSM-IV Symptom Cluster</th>
<th>AE (%) not meeting DSM-IV criteria (n = 132) (95% CI)</th>
<th>% (95% CI) Endorsing “Moderately” or Higher</th>
<th>CCATT (%) not meeting DSM-IV criteria (n = 6) (95% CI)</th>
<th>CCATT (%) meeting DSM-IV criteria (n = 43) (95% CI)</th>
<th>CCATT (%) meeting DSM-IV criteria (n = 7) (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Repeated, disturbing memories, thoughts, or images of a stressful military experience?</td>
<td>B</td>
<td>5.30 (1.48–9.13)</td>
<td>83.33 (53.51–100.00)</td>
<td>16.28 (5.24–27.31)</td>
<td>71.43 (37.96–100.00)</td>
<td></td>
</tr>
<tr>
<td>2. Repeated, disturbing dreams of stressful military experiences?</td>
<td>B</td>
<td>6.06 (1.99–10.13)</td>
<td>66.67 (28.95–100.00)</td>
<td>9.30 (0.62–17.98)</td>
<td>42.86 (6.20–79.52)</td>
<td></td>
</tr>
<tr>
<td>3. Suddenly, acting or feeling as if a stressful military experience were happening again (as if you were reliving it)?</td>
<td>B</td>
<td>2.27 (0.00–4.82)</td>
<td>33.33 (0.00–71.05)</td>
<td>6.98 (0.00–14.59)</td>
<td>28.57 (0.00–62.04)</td>
<td></td>
</tr>
<tr>
<td>4. Feeling very upset when something reminded you of a stressful military experience?</td>
<td>B</td>
<td>6.06 (1.99–10.13)</td>
<td>83.33 (53.51–100.00)</td>
<td>13.95 (3.60–24.31)</td>
<td>57.14 (20.48–93.80)</td>
<td></td>
</tr>
<tr>
<td>5. Having physical reactions (e.g. heart pounding, trouble breathing, sweating) when something reminded you of a stressful military experience?</td>
<td>B</td>
<td>4.55 (0.99–8.10)</td>
<td>16.67 (0.00–46.49)</td>
<td>9.30 (0.62–17.98)</td>
<td>71.43 (37.96–100.00)</td>
<td></td>
</tr>
<tr>
<td>6. Avoid thinking about or talking about a stressful military experience or avoiding having feelings related to it?</td>
<td>C</td>
<td>6.06 (1.99–10.13)</td>
<td>66.67 (28.95–100.00)</td>
<td>20.93 (8.77–33.09)</td>
<td>71.43 (37.96–100.00)</td>
<td></td>
</tr>
<tr>
<td>7. Avoiding activities or situations because they remind you of a stressful military experience?</td>
<td>C</td>
<td>3.03 (0.11–5.95)</td>
<td>50.00 (9.99–90.01)</td>
<td>9.30 (0.62–17.98)</td>
<td>57.14 (20.48–93.80)</td>
<td></td>
</tr>
<tr>
<td>8. Trouble remembering important parts of a stressful military experience?</td>
<td>C</td>
<td>0.00</td>
<td>16.67 (0.00–46.49)</td>
<td>4.65 (0.00–10.95)</td>
<td>14.29 (0.00–40.21)</td>
<td></td>
</tr>
<tr>
<td>9. Loss of interest in activities that you used to enjoy?</td>
<td>C</td>
<td>3.03 (0.11–5.95)</td>
<td>66.67 (28.95–100.00)</td>
<td>2.33 (0.00–6.83)</td>
<td>57.14 (20.48–93.80)</td>
<td></td>
</tr>
<tr>
<td>10. Feeling distant or cutoff from other people?</td>
<td>C</td>
<td>9.09 (4.19–14.00)</td>
<td>83.33 (53.51–100.00)</td>
<td>20.93 (8.77–33.09)</td>
<td>57.14 (20.48–93.80)</td>
<td></td>
</tr>
<tr>
<td>11. Feeling emotionally numb or being unable to have loving feelings for those close to you?</td>
<td>C</td>
<td>6.06 (1.99–10.13)</td>
<td>50.00 (9.99–90.01)</td>
<td>11.63 (2.05–21.21)</td>
<td>85.71 (59.79–100.00)</td>
<td></td>
</tr>
<tr>
<td>12. Feeling as if your future will somehow be cut short?</td>
<td>C</td>
<td>0.76 (0.00–2.24)</td>
<td>66.67 (28.95–100.00)</td>
<td>2.33 (0.00–6.83)</td>
<td>57.14 (20.48–93.80)</td>
<td></td>
</tr>
<tr>
<td>13. Trouble falling or staying asleep?</td>
<td>D</td>
<td>12.12 (6.55–17.69)</td>
<td>66.67 (28.95–100.00)</td>
<td>27.91 (14.50–41.31)</td>
<td>100.00</td>
<td></td>
</tr>
<tr>
<td>14. Feeling irritable or having angry outbursts?</td>
<td>D</td>
<td>4.55 (0.99–8.10)</td>
<td>66.67 (28.95–100.00)</td>
<td>11.63 (2.05–21.21)</td>
<td>100.00</td>
<td></td>
</tr>
<tr>
<td>15. Having difficulty concentrating?</td>
<td>D</td>
<td>2.27 (0.00–4.82)</td>
<td>66.67 (28.95–100.00)</td>
<td>6.98 (0.00–14.59)</td>
<td>85.71 (59.79–100.00)</td>
<td></td>
</tr>
<tr>
<td>16. Feeling jumpy or easily startled?</td>
<td>D</td>
<td>3.03 (0.11–5.95)</td>
<td>66.67 (28.95–100.00)</td>
<td>18.60 (6.97–30.24)</td>
<td>71.43 (37.96–100.00)</td>
<td></td>
</tr>
<tr>
<td>17. Being “superalert” or watchful or on guard?</td>
<td>D</td>
<td>6.82 (2.52–11.12)</td>
<td>66.67 (28.95–100.00)</td>
<td>9.30 (0.62–17.98)</td>
<td>71.43 (37.96–100.00)</td>
<td></td>
</tr>
</tbody>
</table>

B, intrusion; C, avoidance; D, arousal.

compare the probabilities of demographic and occupational variables with meeting DSM-IV PTSD symptom criteria. Relative risks could not be computed if a contingency table had an observed cell count of zero.

RESULTS

**PCL-M Total Score and DSM-IV PTSD Criteria Prevalence**

In total, 13 out of 188 (6.91%; 95% CI = 4.08–11.46) AE and CCATT crew members met the DSM-IV PTSD symptom criteria. PCL-M total scores ranged from 41 to 74 for AE personnel meeting the DSM-IV PTSD symptom criteria and 42 to 68 for CCATT personnel meeting the DSM-IV PTSD symptom criteria. In total, 6 out of 138 (4.35%; 95% CI = 0.95–7.75) AE crew members and 7 out of 50 CCATT crew members (14.00%; 95% CI = 4.38–23.62) met the DSM-IV PTSD symptom criteria. A Fisher’s exact test identified that CCATT crew members were 3.22 times (95% CI = 1.14–9.12) more likely than AE crew members to meet the DSM-IV PTSD symptom criteria (p < 0.05). The frequencies and percentages of AE and CCATT crew members
meeting the DSM-IV PTSD symptom criteria by PCL-M total score range are shown in Table II.

**Most Commonly Endorsed PTSD Symptoms**

Because none of the respondents in the “low” range of PCL-M scores met criteria for PTSD, Table III reports only the percentages of AE and CCATT respondents endorsing “moderately” or “extremely” for each of the PCL-M items. For AE crew members meeting the DSM-IV PTSD symptom criteria, the most commonly endorsed items for this subset were “repeated disturbing memories, thoughts, or images of a stressful military experience” (re-experiencing), “feeling very upset when something reminded you of a stressful military experience” (re-experiencing), and “feeling distant or cutoff from other people” (avoidance). For AE crew members that did not meet the symptom criteria, the most commonly endorsed items were “trouble falling or staying asleep” (arousal), “feeling distant or cutoff from other people” (avoidance), and “feeling ‘super alert’ or watchful or on guard” (arousal).

For CCATT crew members meeting the DSM-IV PTSD symptom criteria, the most commonly endorsed items were “trouble falling or staying asleep” (arousal), “feeling irrigable or having angry outbursts” (arousal), “having difficulty concentrating” (arousal), and “feeling emotionally numb or being unable to have loving feelings for those close to you” (avoidance). For CCATT crew members that did not meet the symptom criteria, the most commonly endorsed items were “trouble falling or staying asleep” (arousal), feeling distant or cutoff from other people” (avoidance), and “avoid thinking about or talking about a stressful military experience or avoiding having feelings related to it” (avoidance).

**DSM-IV PTSD Criteria Associations**

Contingency table analyses were conducted to identify associations between demographic (gender, age range, and marital status) and occupational factors (rank range and time on station) with meeting the DSM-IV PTSD symptom criteria. Subsequent Pearson \( \chi^2 \) analyses or Fisher’s exact tests did not identify associations that met the a priori significance level of \( p < 0.05 \).

**DISCUSSION**

USAF AE crews and CCATTs represent highly skilled and highly trained military medical assets who perform lifesaving care in austere, dynamic, and resource-limited environments. The current study sought to identify the prevalence of post-traumatic symptoms within a sample of these unique medical personnel.

In an anonymous, online, occupational health survey, AE crews (nurses and aeromedical technicians) and CCATT members (critical care nurses and respiratory therapists) were asked about current symptoms of post-traumatic stress. Overall, these results suggest that a proportion of CCATT members endorsed symptoms consistent with PTSD (14%) and that CCATT personnel were more than 3 times as likely to do so than their general AE counterparts (relative risk = 3.22; 95% CI = 1.14–9.12).

Consistent with this finding, recent research conducted with civilian, ground-based nurses suggests greater PTSD symptomatology in ICU nurses than in their general nursing counterparts. \(^3\) Although they did not report percentages meeting clinical diagnostic criteria for PTSD, Mealer et al’s survey of ICU and general nurses from a metropolitan area indicated that between 24% and 29% of ICU nurses reported elevated post-traumatic symptoms related to the work environment as compared to 14% of general nurses.

Examining the current findings in the context of other recent PTSD prevalence studies gives some added perspective to these data. The percentage of AE personnel meeting the symptom criteria for PTSD in this study (\( n = 6/138; 4.35\% \)) is consistent with PTSD point prevalence estimates for remotely piloted aircraft (RPA) operators (4.34%) and imagery analysts (4.22%). \(^18,22\) These are military populations similar to AE in that they are not personally engaged in combat but are exposed to some degree of war-related traumatic events. The prevalence of PTSD among AE in this sample is slightly elevated as compared to point prevalence estimates for PTSD in the general population (1–3%), and it is at the low end of recent general military point prevalence estimates, which range from 4% to 17%. \(^23\) The AE findings from this survey suggest that, for this sample, the percentage meeting clinical criteria for PTSD is consistent with other military service members who have been deployed or “deployed on station” but not exposed to direct combat or the immediate physical threats of combat.

In contrast to the AE sample, the rate of CCATT members endorsing clinically significant PTSD symptoms in the current study (\( n = 7/150; 14\% \)) is more consistent with point prevalence estimates for deployed military service members exposed directly to combat (7.6–8.7%), and it is also consistent with recent point prevalence estimates for civilian, ground-based ICU nurses (18%), both of which are elevated as compared to the general population (1–3%). \(^9\) Although the PTSD estimates from the current study are tentative given our small sample size, the findings are consistent with the findings from prior studies with much larger samples, and suggest that the highly selective, highly trained critical care personnel who make up CCATTs are not immune to the stresses inherent to either military service or critical care medicine.

The percentage of CCATT personnel meeting clinical criteria for PTSD suggests the need for a deeper analysis of the causes and correlates of this issue in order to establish a plan for prevention and intervention. Differences in patient acuity between AE and CCATT missions could account for a portion of the difference in PTSD symptoms. Although the current survey did not assess exposure, recent analyses of the epidemiology of patients requiring CCATT transport.
report high frequencies of burns, musculoskeletal injuries, and amputations related to explosive devices,\textsuperscript{5,26} which suggest CCATT personnel are exposed to and managing young (median age in 2011 was 25 years) patients with disfiguring, painful, war-related injuries.

Although the types of presenting injuries in the CCATT environment cannot be controlled, the current status of the PTSD literature suggests that both group and individual factors may be important keys to intervening with high-risk populations.\textsuperscript{27–29} Subject-matter experts within the CCATT community have indicated that the training and deployment practices for these teams are a potential risk factor for increased stress and decreased resiliency. The CCATT personnel do not train or deploy as a team. Instead, individual members receive their requisite training and are deployed from their home station to join a CCATT with whom they may have never worked. In contrast, AE teams typically train and deploy as a unit, which means they have an opportunity to build cohesiveness and understand each other’s strengths and weaknesses before deploying. This concern is consistent with recent research on PTSD in deployed military populations that suggests that low unit cohesion may be an important, modifiable risk factor for PTSD.\textsuperscript{27}

An individual factor that may be involved in increased risk for PTSD for critical care personnel is a lack of resiliency. Research conducted by Mealer and colleagues over the past decade has indicated that high levels of resiliency are associated with lower levels of PTSD, anxiety, depression, and burnout in civilian ICU nurses,\textsuperscript{28–30} and that an education/intervention program that builds resilience can potentially protect against and mitigate the symptoms of PTSD.\textsuperscript{31} Specific cognitive and behavioral skills, including mindfulness-based stress reduction techniques, cognitive restructuring, aerobic exercise, and written exposure therapy, taught over a 12-week intervention, was shown to be associated with increased resiliency and reduced PTSD symptom scores. Additionally, this intervention was shown to have high satisfaction and feasibility ratings, suggesting it was well tolerated by these busy professionals.\textsuperscript{31} This line of research by Mealer and colleagues has not only begun to shed light on the factors associated with high levels of resiliency in civilian ICU nurses, but has also begun to establish the connection between resiliency building and PTSD symptom reduction in this population. Future research investigating the feasibility and acceptability of a similar type of resilience training for CCATT personnel may help to reduce the prevalence of post-traumatic symptoms in this high-risk population.

Notable differences in the self-reported symptom patterns of AE and CCATT personnel who met criteria for PTSD were seen. For AE respondents who met the DSM-IV PTSD symptom criteria, the most often endorsed items were re-experiencing symptoms, whereas CCATT personnel who met PTSD symptom criteria most frequently endorsed arousal symptoms. Although additional research is necessary to substantiate, differences in training and operational experiences may contribute to this finding. Even though CCATT personnel endorsed higher levels of PTSD symptomatology overall, the extensive trauma training that CCATT personnel receive may enable them to better compartmentalize (and therefore be less likely to re-experience) potentially traumatic experiences than their general AE counterparts. Additionally, a recent study suggests that individual factors such as poor cognitive control (i.e., the ability to regulate the content of one’s cognition by keeping desired/needed information active while inhibiting unneeded information) may be specifically associated with re-experiencing/intrusion symptoms of PTSD.\textsuperscript{32} Although additional research is needed, group and individual differences between CCATT and AE personnel could be associated with their symptom presentation and be used to guide prevention and intervention.

None of the assessed demographic or occupational factors were identified as being associated with the PTSD symptom criteria, but this lack of finding may be explained by the low number of respondents in this sample who met the PTSD symptom criteria ($n = 13$) and by limitations in the questionnaire. Occupational factors that have been shown to be associated with PTSD symptomatology in prior research on military populations (e.g., shift schedule, shift rotation, hours worked per week)\textsuperscript{18,33} were not assessed by both the AE and CCATT questionnaires and therefore were not included in this study. Length of time spent in current duties also emerged as a significant association with PTSD symptomatology in prior research with other military populations (e.g., RPA operators)\textsuperscript{33} and with research on civilian ICU nurses.\textsuperscript{34} Although this variable was assessed in the current surveys, the association between RPA operators meeting the DSM-IV PTSD criteria and time in current duties utilized a dichotomous variable of less than 24 months and greater than or equal to 24 months.\textsuperscript{33} The study with civilian ICU nurses used time as a continuous variable that was associated with PTSD.\textsuperscript{34} In the current study, time in AE and CCATT duties was assessed using a dichotomous variable of less than 12 months and greater than or equal to 12 months. The difference in length of time used to dichotomize the respondents may have impacted the ability to detect meaningful differences in the current sample.

Although this study helps shed light on the rate of PTSD symptomatology reported by AE and CCATT crew members, there are some limitations. First, the sample size included in this study is small compared to the prior research cited that examined DSM-IV PTSD symptom criteria. The small sample size limits the generalizability of the results to the population as a whole and may have adversely affected the ability of the analyses to detect the effect of significant differences among demographic and occupational factors. Although sample size was small, the estimated response rates (40% and 37% of AE and CCATT crew members, respectively) are within the acceptable range and are considered good for web-based surveys.\textsuperscript{35,36}
The results of this study were from self-report surveys, which rely on genuine responses from participants and may also affect the generalization of results. In the military, as well as in the general U.S. population, there is a hesitancy to disclose mental health issues, even when anonymity is assured. Despite the anonymous nature of the surveys, the careful design of the study, and reassurances from leadership regarding anonymity, it may be possible that some participants did not trust that responses were completely anonymous and were suspicious as to how survey results were to be used. Additionally, refusal to participate in a survey regarding post-traumatic stress may be the result of avoidance, which is a symptom of PTSD in its own right. As a result, the current results could underestimate PTSD prevalence in this population. However, this is a limitation for any volunteer-based, self-report assessment of post-traumatic stress.

The measure of PTSD used in the study, the PCL-M, assesses symptoms of PTSD based on the criteria listed in the DSM-IV. In 2013, after the data collection period for this study closed, the American Psychiatric Association published the DSM-V. Although additional research using the criteria stipulated in the most current version of the DSM is warranted, initial point prevalence comparisons between the previous and current criteria suggest that the new criteria will not significantly change PTSD prevalence data, and comparisons to earlier research using DSM-IV criteria remain valid. Lastly, the PCL-M is designed as a screening tool not as a diagnostic tool and therefore actual diagnoses of PTSD would need to be verified by a clinical interview.

A potentially relevant occupational factor, namely the number of times a respondent deployed as an AE or CCATT member, was not included in the current analyses. The two surveys assessed this factor in slightly different ways making comparisons between the groups on this variable unclear at best. As the number of deployments likely correlates highly with exposure to traumatic experiences, not being able to assess for differences between the two groups on this variable is a shortcoming of this study. However, a similar variable, time in duties, was assessed. This variable should correlate highly with number of deployments (i.e., the longer one is in the job, the more times they are deployed), and this factor was not found to be associated with PTSD symptomatology in these samples.

Finally, the results of this study are descriptive in nature and do not allow for cause-effect conclusions to be drawn. Despite these limitations, the current study provides a descriptive representation of the PTSD symptomatology experienced among a sample of AE and CCATT crew members.

CONCLUSIONS

The results demonstrate the importance of assessing and monitoring PTSD symptoms among AE and CCATT crew members. Similar to comparisons between civilian ground-based ICU nurses and their general nursing counterparts, the aeromedical critical care personnel in this study reported PTSD symptoms at a higher rate than their general aero-counterparts. The results of this study also highlight the need for different intervention practices for AE and CCATTs. Although both groups have to contend with military operational stresses as well as the demanding aerospace environment, the management of trauma under these conditions appears to put CCATT nurses and respiratory therapists at greater risk for post-traumatic symptomatology. Based on the findings of the current study and prior research on modifiable factors, such as resiliency, that are associated with reductions in PTSD symptomatology, future research should focus on replicating the current findings, identifying additional factors that are associated with both increased risk and resiliency, and developing operationally feasible prevention and interventions for Air Force CCATT and AE members.

ACKNOWLEDGMENT

This research was supported, in part, by Defense Health Program research funds from the U.S. Air Force School of Aerospace Medicine.

REFERENCES


Finding Their Way Back In: Family Reintegration Following Guard Deployment

Col Deborah C. Messeccar, NC ANG (Ret.)

ABSTRACT  Objective: The aim of this study was to describe deployed National Guard members’ and their families’ perceptions of their experience with family reintegration, and the causes and conditions of challenges reintegration presents after deployment. Methods: A total of 26 National Guard members and 19 family members participated in individual (n = 22), couples (n = 6), or focus group (n = 17) interviews. In-depth interviews were used to assess needs and maximize input from military families regarding deployment-related experiences and reintegration issues. Qualitative coding and analysis of data were completed using NVivo. Results: Finding their way back in is the key process that the military members must complete to successfully reestablish their desired social connections with the family and reclaim their place within the family. Several conditions shape the degree of challenges with reintegration that veterans and their family will encounter. These include preparation for deployment, length and type of deployment, communication during deployment, and finally, awareness of how deployment changes the military member and the family. Conclusion: Support resources dedicated to providing National Guard members and their families with assistance in preparing for deployments and educating them about the importance of communication during deployment should be maintained and expanded. Broader educational efforts that increase awareness of what to expect regarding how deployment changes the military member and the family are needed.

INTRODUCTION
The wars in Iraq and Afghanistan have resulted in unprecedented number of deployment separations among the members and families of the National Guard and Reserve. Although the consequences of service-induced separation are usually thought of in terms of only active duty, members of the National Guard and Reserve and their families make similar sacrifices. In 2013, approximately 1.1 million men and women were serving in the National Guard and Reserve versus the 1.37 million serving on active duty.1 The ratio of dependents of the National Guard and the Reserve nearly equal those of active duty at 1:1.3 versus 1:1.4.1,2 The length and frequency of deployments for the National Guard and Reserve have also been increasing. Since 2001, Army National Guard members have been deployed on an average of 2.2 times for a duration of 12 months per deployment.3

Because National Guard and Reserve personnel are older, more likely to be married, and are parents as compared to their active duty military counterparts, the impact of more frequent and longer deployments creates more challenges for those members to reintege back into civilian family life.3 In fact, some experts contend that reintegration difficulties account for much of the distress these service members experience after deployment.4

Family reintegration is the process of reentering—the family unit and returning to previous roles. Though return from deployment can be a happy occasion, homecoming can turn into a stressful event for those who are not alert to the impact of changes that occurred during separation. Further, the individual returning from deployment may still be experiencing the stressful effects of deployment and these problems may start to emerge only after the honeymoon of reunification has passed. Indeed, for some veterans the acute responses to stress they developed while deployed, such as exaggerated startle responses, the need to control the immediate environment and being alert to danger, or aggressive behavior that was adaptive or necessary in the setting, may persist and their symptoms become chronic. Historically, unidentified and untreated post-traumatic stress disorder presents special risks for family reintegration that may put the military service members and their families at higher danger for maladaptive responses to stress such as alcoholism and family violence, particularly if these problems existed before deployment.5-7 Women, National Guard members, and reservists seem particularly vulnerable to higher rates of negative outcomes from deployment.8,9 Among the National Guard and Reserve, problems following deployment seem to emerge later rather than immediately after deployment. One study found that as much as 40% Army National Guard soldiers returning from deployment reported having mental health problems 3 months after returning home.10 A longitudinal assessment of mental health problems among Army soldiers returning from Iraq found that the National Guard and Reserve members reported higher rates of both mental health and general health problems and were referred at substantially higher rates 3 to 6 months after return from deployment than their active counterparts, suggesting that problems develop on a different time trajectory.11 Why differences were found is
unknown, possible explanations offered by the researchers included lack of access to service, lack of access to support from military peers, and problems with reentering civilian life. Yet data on the causes and conditions associated with challenges with National Guard and Reserve family reintegration after deployment are limited.

**Background**

Many of the studies to assess the impact of wartime separation and subsequent family reintegration have concentrated on either active duty alone\(^5,12,13\) or samples in which findings are combined with those of active duty.\(^14\) Thus, the unique experiences of the National Guard and Reserve members and their families were not detailed. Studies concentrated solely on National Guard and Reserve populations have focused on the emergence of psychological symptoms months after return that appear to be related to reintegration distress,\(^10,11\) the barriers to seeking help and the role of peer support with reintegration,\(^15\) and the piloting/proposing self-help psychoeducational pilot interventions.\(^3,4,16\) These studies documented the distress many National Guard and Reserve members and their families have with reintegration, highlighted stigma barriers to receiving help by the military member, and piloted possible interventions to deal with member and family distress. However, the ability to fully understand the linkages of the deployment experience with poor reintegration outcomes and therefore tailor interventions could be further enhanced. For example, the pilot study\(^16\) of a self-directed program of using integrative therapies to promote reintegration, though well received by participants, did not explore the existence of reintegration issues among the dyads that would indicate who might benefit most from this intervention approach.

Studies specifically exploring the reintegration experience and/or communication difficulties have contributed greatly to our appreciation of how difficult reintegration can be, but further study could add important dimensions to our understanding. One study that focused substantially on the experience of National Guard and Reserve women after deployment was conducted without asking participants to reflect on the overall deployment experience,\(^17\) making it difficult to link possible causes and conditions of challenges with reintegration. Similarly, interviews conducted by Rivers et al\(^13\) with Army active duty nurses also focused only on the post-deployment experience with reintegration. An integrative review of communication during deployment that included studies on both active duty and the Guard and Reserve\(^18\) could postulate only associations with subsequent possible reintegration outcomes. Hinojosa et al\(^19\) conducted in-depth interviews with 20 Guard and Reserve male deployed members. Though participants were queried about all phases of the deployment, communication issues were the sole focus of the interviews and other causes and conditions of reintegration difficulties were not explored. In the study by Lapp et al\(^20\) of 18 spouses of Guard members in Wisconsin who were in various stages of the deployment process, only a small proportion of interviewees had been already through the deployment process and were working on postdeployment adjustment. For these spouses, they described a process of finding a new normal after deployment; however, the conditions and context of how that occurs were not described. A qualitative descriptive study\(^21\) of returned Reserve members indicated that those who were simultaneously coping with major life events such as divorce, marriage, illness, or changing jobs struggled, as did those who closed off their communications whereas deployed to shield the family. In a qualitative study of risk factors for suicide in returning veterans,\(^22\) participants reported thoughts of suicide in response to the sense of failed belongingness (or the inability to establish social connections) they endured because of reintegration problems. These studies show that deployment separation is difficult, and reintegration can be hard to achieve, though all of the conditions and context that explain variable responses to reintegration could be further described.

**METHODS**

**Research Design**

This was a qualitative descriptive study that used intensive interviewing as the method of data collection. Participants were offered individual, couple, or focus group interviews. Individual and couple interviews were offered in addition to focus group interviews because of concerns that participants may be less willing to discuss sensitive aspects of their reintegration experience in group settings. Focus group interviews among the military members allowed exploration of issues that many shared in common. Individual and couple interviews permitted a far more in-depth exploration of the topic and the experiences of an individual family experience and proved to be a very useful technique for this interpretive inquiry. As Charmaz has said, “...the in-depth nature of an intensive interview fosters eliciting each participant’s interpretation of his or her experience.”\(^23\)

**Sample and Setting**

The purpose of this study was to describe deployed National Guard members’ and their families’ perceptions of their experience with family reintegration and the causes and conditions of challenges reintegration presents after deployment. Participants included both National Guard members and/or family members of guardsmen deployed since 2001 after the September 11, 2001 attacks. A total of 26 Guard members and 19 family members participated in either individual (n = 22), couple (n = 6), or focus group (n = 17) interviews. Four couples participated in the interviews, three were interviewed as a couple whereas one couple participated in separate individual interviews. Otherwise, the family member participants and the deployed member participants were not related. Only military members participated in focus group interviews; no family members chose this interview option. To be eligible for study participation, participants had to have been deployed
and returned home within the last 3 months to 5 years, or be the family member of a deployed and returned National Guard member. A family member could participate and be interviewed even if the National Guard member they were related to was not interviewed and vice versa. Participants had to be 18 years or older, be able to speak English, and consent to be audiotaped. A summary of the key characteristics of the participants is presented in Table I.

Recruitment
Following university and military institutional review board approval, recruitment was conducted with the cooperation of the Family Support and Assistance Programs of the Army and Air National Guard in Oregon. Recruitment strategies included developing a study website with links for contacting project staff, making presentations at family support meetings and gatherings; placing ads in a newsletter for the Oregon National Guard, asking interviewees to share the study flyer with other potential participants, and traveling to armories around the state and asking them to distribute flyers about the study. Interested participants contacted study personnel, who then scheduled times to review consent procedures and conduct interviews. As an incentive for participation, a gift certificate of $20 was offered.

Procedures
Interviews lasted from 1 to 2.5 hours and were conducted by either the researcher or the study project director and were taped with participants’ permission. The consent information sheet was covered before the start of any interview. Participants were given a copy of the interview consent information sheet but were not asked to sign the document as we wanted to be able to assure military members that their comments could not be tied to their individual participation. Protocols for individual and couple interviews and focus groups were identical except for minor changes to wording (see Table II). The interviewer asked each participant to describe and reflect upon his or her experiences related to the deployment and their experience with reintegration. According to Charmaz,23 a few broad, open-ended questions started the interview process. Participants were probed to invite detailed descriptions of the topics. By using open-ended, nonjudgmental questions, participants were encouraged to provide unanticipated statements and stories. A concerted effort was made to go beneath the surface of ordinary conversation and examine the participants’ feelings about the events they were describing. A demographic instrument was also completed by all participants at the end of the interview. All of the interviews were transcribed verbatim by a professional transcriptionist who had signed a confidentiality agreement. Data were collected and analyzed between 2009 and 2011.

Data Analysis
Initial analysis by researcher and the study project director involved detailed, slow, and reflective exploration of all the transcribed interviews—by doing line-by-line coding, reading between the lines, identifying concepts, and thinking about all of each segment of text’s possible meanings which were then recorded in codes along with preliminary definitions in NVivo 10 (QSR International, Burlington, Massachusetts). To maintain coding consistency during open coding, codes and code definitions were reviewed and discussed after each interview was coded. The next stage of the analysis was to refine the open coding system while simultaneously adding new data by coding more interviews. During axial coding, the properties of the codes were examined to establish links among the

<table>
<thead>
<tr>
<th>TABLE I. Sample Demographics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characteristic</td>
</tr>
<tr>
<td>Age (Years)</td>
</tr>
<tr>
<td>20s</td>
</tr>
<tr>
<td>30s</td>
</tr>
<tr>
<td>40s</td>
</tr>
<tr>
<td>50s</td>
</tr>
<tr>
<td>60s</td>
</tr>
<tr>
<td>Marital Status (Years)</td>
</tr>
<tr>
<td>&lt;5</td>
</tr>
<tr>
<td>&gt;5</td>
</tr>
<tr>
<td>Not married</td>
</tr>
<tr>
<td>Children</td>
</tr>
<tr>
<td>Infant and/or Preschool</td>
</tr>
<tr>
<td>School Age</td>
</tr>
<tr>
<td>Teen Age</td>
</tr>
<tr>
<td>Grown</td>
</tr>
<tr>
<td>No Children</td>
</tr>
<tr>
<td>Gender</td>
</tr>
<tr>
<td>Male</td>
</tr>
<tr>
<td>Female</td>
</tr>
<tr>
<td>Education</td>
</tr>
<tr>
<td>High School</td>
</tr>
<tr>
<td>Associate Degree</td>
</tr>
<tr>
<td>Bachelors or Higher</td>
</tr>
<tr>
<td>Branch</td>
</tr>
<tr>
<td>Air Guard</td>
</tr>
<tr>
<td>Army Guard</td>
</tr>
<tr>
<td>Rank</td>
</tr>
<tr>
<td>E4 and Below</td>
</tr>
<tr>
<td>E5 and Above</td>
</tr>
<tr>
<td>O3 and Below</td>
</tr>
<tr>
<td>O4 and Above</td>
</tr>
<tr>
<td>Deployment Location</td>
</tr>
<tr>
<td>Iraq/Afghanistan</td>
</tr>
<tr>
<td>Other Southwest Asia Country</td>
</tr>
<tr>
<td>(Kuwait; Qatar, Turkey)</td>
</tr>
<tr>
<td>Germany; Spain</td>
</tr>
<tr>
<td>Multiple Deployments</td>
</tr>
<tr>
<td>Yes</td>
</tr>
<tr>
<td>No</td>
</tr>
<tr>
<td>Deployment Duration (Year)</td>
</tr>
<tr>
<td>&lt;1</td>
</tr>
<tr>
<td>&gt;1</td>
</tr>
<tr>
<td>Time Home (Year)</td>
</tr>
<tr>
<td>&lt;1</td>
</tr>
<tr>
<td>&gt;1</td>
</tr>
</tbody>
</table>
categories. At various stages in the open and axial coding process, the researcher met with a nationally recognized qualitative research consultant experienced with the NVivo software to review the analysis. Queries (a powerful feature of the software) suggested by the consultant were used to explore associations of various conceptual categories with one another and to determine what might be unique about a given category. Hypotheses were then generated about categories of data and their relationships and interrelationships. Patterns of relationships were tested and developed with the data. The final stage of the analysis occurred as the analysis was written and evidence was drawn from the data to illustrate the categories and subcategories.

RESULTS
Finding their way back in is the key process that the military member must complete to successfully reestablish their desired social connections with family members and reclaim their place within the family. An Army soldier summarized this basic need to find a way back in:

When we first got back the mind-set of a soldier is: I’m gonna come back, and everything is going to be normal,” it’s gonna be right where I left off. And that’s probably the biggest misconception that we could have . . . Because when I came back, I was right where I left off, but my family had moved on for a year and a half. So now I’m playing catch-up . . . I mean that was a big conflict right there, but they learned for a year and a half how to live without me. So they didn’t need me anymore. And that was really tough to get back in that element.

Few military members had anticipated the degree of adjustments that would be required:

When you have been in a foreign country, coming back is a culture shock. Well it was: I’m really back, but things had changed here. There was no question. Finding my way back in the family was really rocky and I think probably the hardest thing.

One Army wife described the yearning her husband felt for trying to get back his life:

. . . they want so badly to come home. They want so much to have a normal life again. And they work very hard to get it . . . So we need to be taught . . . how do we learn the new life instead of trying to get back what we had, because that’s forever gone.

Finding their way back in requires quite a bit of self-discovery and thoughtful reflection on the part of the military members and their family about the changes that have occurred during the separation. They must identify and deal with an array of problems from changes in the self and family relationships that have piled up during the separation. For some, the changes are minimal, and the reintegration process is uneventful. For many, the veteran has been changed subtly by his or her service in ways that might not be apparent to them and is only recognized by family and friends. These changes are a normal response to stress that have been carried over from the deployed environment and include the actual symptoms of distress along with ways of coping that worked while deployed environment but may interfere with the reintegration process. In addition, families face issues with renegotiating roles and relationships that may have subtly changed during the military member’s absence.

Military members and families pointed to several causes and conditions that contributed to creating challenges with reintegration. These included preparation for deployment, length and type of deployment, communication during deployment, and finally awareness of how deployment changes the military member and the family.
An airman said
readjusting is as one soldier stated about his children:

“...”

The longer the member is gone, the more difficulties holding on to ways of coping that worked well in the deployed setting, but not in the home setting.

Multiple deployments are an additional problem. As one airman shared:

We do shorter durations; however, I was working with people that were on their ninth trip. So 4–6 month deployment times nine. So the Army does an 18-month shot, and they’re done for a while. Whereas these guys were ... going through their AEF buckets, and every single time they’d come up, they deploy.

Experiencing more stresses during the deployment seemed to lead to reporting more difficulties holding on to ways of coping that worked well in the deployed setting, but not in the home setting.

Communication During Deployment
Communication—how it occurred, with who it occurred, and the quality of that connection—was an important part of trying to maintain the National Guard members’ place in their families during deployment. Many strategies and methods were used. As one wife said:

I would get in bed and turn on the computer, and we’d just email and almost like instant messaging, back and forth ... .

One parent used texting:

One of the things that really helped out ... my daughter text messages all of her friends. And I was able to text message and send her ... texting. “How’s your day going?” And when I started doing that, it seemed to help a lot.

However, contact may not work to maintain the kind of communication that facilitates true sharing in one another’s lives. During separation, the ability to be able to talk about family problems as they come up can help to keep the deployed family member in the middle of family life. However, given that they might not be able to do anything about a problem at home, it actually can add more stress. One guard member said:

No worse feeling than feeling useless in a situation that you can do nothing about across the world. You know, I’m halfway around the world. There’s not a thing I can do about it. I felt useless. What can I do? What can I do? So that’s a bad place to be.

So in many cases either the family or the military tried to protect one another from the stresses they were experiencing by not sharing what was happening. Communicating while deployed can be both a positive and a negative. Conversely for family members, knowing about some of the stresses and strains the Guard member was experiencing in the environment could be frightening. During one wife’s call, her
Finding Their Way Back In: Family Reintegration Following Guard Deployment

Awareness of How Deployment Changes You

Deployment often changes military members and their families in ways that might not be apparent until the family is reunited. Conditions that affect the National Guard member’s response to reintegration include the difficult symptoms that they may be experiencing postdeployment and their residual use of ways of coping that worked well in the deployed setting, but interfere with family life. Guard members in dangerous settings got used to large amounts of adrenaline every day, which was hard to cope with once home and that adrenaline was no longer pumping.

If you are human at all, when you lock and load and head out that gate, your adrenaline is going... And it’s that way until you come back in... and you [clear] your weapon. And... and then you’re... you think you’re safe and you’re... and mortar rounds come in. And then here’s the adrenaline again. So you live on this adrenaline rush so long, when you get back here, you’re just waiting for that to snap in.

Deployed ways of coping are ways of coping in the deployed environment that do not work well or interfere with communications about what was really going on.

One Air Guard member had changed his normal easygoing style and this was intimidating his family:

... when I got back, little by little things were going wrong, and I think part of what we do as soldiers is block things out rather than deal with them. If we don’t have an answer to overcome that objective right away, we kinda block it out.

An Army wife observed that her husband also tried to ignore or avoid problematic emotions:

Because some of these other issues that came up with my husband, he didn’t want to address. So he either ignored them or avoided them.

For the family (which the National Guard member is now having to act as a reintroduced member), there is the work of renegotiating roles and relationships, as well as residual family conflict issues that must be addressed. The work of renegotiating roles and relationships can be difficult. As one Army wife related:

We’re totally different people when they come back. My children grew so much in that year plus that he was gone, and it’s so hard for them to come back and realize... the kids are not talking to him the same way, they’re teenagers now. They don’t want to hang out with you....

One Army wife noted her husband’s frustration with trying to spell his wife from childcare:

He wanted to get the kids... it started off okay, and then he felt like his whole summer was spent carting kids around... it made for some rough times, you know, when he first came back.

In situations where the work of reintegration has become too much them—the barrier of the military member’s refusal to recognize they need help or to seek it can lead to not being able to manage the task of finding their way back in to the family. As one Air Guard wife put it:

... my husband is a former Marine, and has it altogether and has the pride up to here, and there’s never anything wrong with them. “No, no, I don’t have that problem.”... “Guess what, honey. I’ve got news for you.”

DISCUSSION

Previous studies described problems with family reintegra-

tion following deployment5–9 and some researchers have reported that National Guard and Reserve members may be at higher risk for these difficulties.10,11 This research built upon that earlier work by adding to what is known about reintegration by interviewing both National Guard and their family members and exploring the conditions that shape the degree of challenges with reintegration. If change in the family and the member has been great enough, they must deal with resetting expectations and renegotiating roles and relationships. Among military members interviewed...
Finding Their Way Back In: Family Reintegration Following Guard Deployment

postdeployment in other recent and historical studies, many have reported that their place in the family seemed to have changed. In one study, returnees reported feeling like a guest in their own house, children acting afraid of or not being warm toward them, or being unsure or expressing doubts about what their family role is now. These findings are consistent with the findings in this study, where some participants expressed quite a bit of existential angst over not fitting in anymore when they got home. The primary task the family faces is renegotiating roles and relationships that may have subtly changed during the military member’s absence. Lapp et al labeled this work finding a new normal.

Findings from this study are also consistent with many other studies that have found that those struggling with postdeployment sequelae can become significantly alienated from their families after deployment. In this study, indicators that the National Guard member may become alienated from the family and the community included a persistent and growing separateness from those they love captured by the feeling of not fitting in. Being unwilling to seek help can perpetuate this feeling and create a negative feedback loop where these feelings are increased.

Other studies have identified similar “finding their way back in” challenges during reintegration. Boundary ambiguity defined as a state in which family members are uncertain in their perception about who is in or out of the family and who is performing which roles and tasks within the family is consistent with the main finding in this study that families who have endured the greatest changes in their members, struggle more with reintegration. Those who had experienced a great deal of stress in the deployed environment, and retained difficult problematic ways of coping, had more difficulty working their way back in. In families where members have changed substantially and where there is role conflict upon the member’s return, also struggle.

Limitations

The most important limitation of this study was the use of retrospective interviews that capture only a one-time picture of the participant’s experience. Prior research suggests that individuals will alter accounts of the past to create a coherent picture of the present. Although this may be viewed as a limitation, it may also be seen as a strength in that participants stories are “the most internally consistent interpretation of presently understood past.” Therefore the narrative data shared by participants represent their present construction of events concerning the family separations they endured and their lives since that time.

CONCLUSIONS

Recommendations from the findings in this study overlap and in some cases extend the recommendations from the 2013 Rand report on family reintegration. As suggested by the Rand report, support resources dedicated to provid-


ABSTRACT

The number of overweight and obese service members has tripled since the beginning of Overseas Contingency Operations.1 Overwhelming evidence suggesting links between obesity and increased risk for conditions such as musculoskeletal injuries, type II diabetes, and cardiovascular disease,3 poses a threat to the strength and performance of our current and future fighting force. Objective: The purpose of the study was to test nurse health coaching (NHC) and/or herbal supplementation for weight reduction in Soldiers during a 12-week intervention. Methods: Body composition, biomarkers, adherence, and motivation were measured at three time points. The NHC provided a weekly scripted interaction. Change scores were compared across study groups using general linear models. Results: Sample demographics (N = 435): mean age 30 ± 8.2 years, 73.4% men, predominantly white (70.1%) and non-Hispanic (80%), 71% married, and 91% enlisted. Results represent the 3 NHC groups compared to control group. Beneficial intervention effects were observed for heel bone mineral density (d = 0.3), 25-Hydroxyvitamin D (d = 0.43), and fasting blood sugar (d = –0.4), but were not significant following application of a 10% false discovery rate. There were no significant findings for any other comparisons. Conclusion: Weight loss proved difficult for all groups; there was no advantage of NHC over an herbal supplement as adjuncts to Army MOVE! for weight reduction. Highly motivated Soldiers were unable to sustain weight loss or body composition changes.

INTRODUCTION

Overweight and obesity continue to be a major public health concern in U.S. adults; the statistics are staggering with over 72 million adults categorized as obese.1,2 According to the 2012 National Health and Nutrition Examination Survey, the prevalence estimate for overweight (body mass index [BMI] 25–29.9) and obesity (BMI over 30) combined in adults of non-Hispanic white origin aged 20 years or older is 68.5% (95% confidence interval, 64.30%–72.5%) with little change in rates since initial publication in 2003–2004.1,3 Today’s Soldiers are not exempt from the influences of poor dietary choices, less physical activity, genetics, and environmental issues. In fact, the number of overweight and obese Soldiers tripled between 1998 and 2010 with the greatest increase noted since the beginning of Overseas Contingency Operations in 2003.4 The 2011 Executive Summary to the “Health Related Behaviors Survey of Active Duty Military Personnel” found that just over one-third (35.7%) of active duty (AD) personnel reported a healthy weight; across all ages, over 50% of men and 34% of women were classified by BMI as overweight with 12.4% meeting criteria for obesity. In addition, 10% of AD service members were required to lose weight before joining the military.4 Based on survey results, dietary behaviors may have contributed substantially to the weight issues, with fruit, vegetable, and whole grain consumption far below national nutritional standards at 11.2%, 12.9%, and 12.7%, respectively.4 The overwhelming evidence for links between obesity and increased risk for conditions including musculoskeletal injuries, diabetes mellitus, cardiovascular disease, and certain cancers,2 poses a real threat to the performance and strength of our current and future fighting force.

For the last few years, the U.S. Army Medical Command has emphasized Soldier medical readiness as the number one priority, followed by implementation of health promotion and risk prevention programs. Medical readiness implies physical and psychological fitness. The condition of being overweight puts additional stress on the body, particularly on joints, and may lead to degradation in performance.5 In addition, unit operational readiness is impacted by lengthy medical evaluations, time away from duty for treatment while on medical profile, or deploying Soldiers who are not combat ready because of excess weight.5 U.S. Army Regulation 600–96 indicates that failure to meet physical fitness or weight standards can negatively impact a Soldier’s career and is grounds for an Army discharge. In the last 5 years, the Army introduced MOVE! as its standardized weight management program adopted from the Veterans Affairs National Center for Health Promotion and Disease Prevention.7 A complete description of the program is beyond the scope of this article but can be found in the Army Public Health Command weight management guide.7 The Army MOVE! approach to weight management includes self-help, individual counseling, and facilitated group support and focuses on behavior modification, physical activity, and diet. In addition, the program emphasizes self-monitoring (food and activity logs) and maintenance of weight loss. Army MOVE! is a comprehensive program that requires intensive face-to-face...
The following speci
To address the overall purpose of this study, we developed
Study Design and Participants
To address the overall purpose of this study, we developed the following specific aims and hypotheses:

1. Evaluate the impact of a nurse coaching intervention, with and without an herbal supplement, on weight loss, body composition, lipid profile, bone density, adherence, and motivation.

H1a: A nurse coaching intervention without an herbal supplement will lead to greater success in meeting weight loss goals as compared to the control group (CG).

H1b: A nurse coaching intervention with an herbal supplement will lead to greater success in meeting weight loss goals as compared to the CG.

2. Describe the influence of motivation on the primary outcome of weight loss and the secondary outcome of adherence at 6 weeks and 12 weeks.

H2: Motivation, as measured by the Self-Motivation Inventory, will positively influence weight loss and adherence to MOVE!

This prospective randomized controlled trial included five arms and two interventions for comparison. The study received funding from the TriService Nursing Research Program (No. HT9404-12-1-TS03, N12-007) and approval by the Madigan Army Medical Center and Uniformed Services University of Health Sciences Institutional Review Boards.

To qualify for the study, participants had to be over 18 years of age, fluent in English, not deploying for 3 months, not previously command referred to Army MOVE!, and generally healthy by self-report. We excluded women who were less than 6 months postpartum or breastfeeding and anyone with an endocrine abnormality, eating disorder, or taking medications contraindicated with Garcinia Cambogia (G. cambogia). The study herbal supplement. The selected preparation also contained calcium and potassium that was felt to be important for bone health. We ordered the supplement for this trial directly from a manufacturer who also prepared the placebo product and shipped the necessary quantity of supplement and placebo with a lot number and a certificate of authenticity.

The study team recruited participants at the first Army MOVE! session of each new cycle in the hospital nutrition clinic following a briefing about the study which described the purpose, voluntary nature of participation, and methods used to ensure confidentiality and anonymity. Each participant provided written informed consent. Immediately after the consent process, enrolled subjects were randomized to 1 of 4 groups; control (group 1), NHC (group 2), NHC + supplement (group 3), or NHC + placebo (group 4). Group assignments occurred with equal probability through a random number generator. The Project Director utilized concealed allocation with opaque envelopes numbered sequentially and group assignment enclosed. The team discussed the details of involvement for the group at the time of assignment. Subjects were encouraged to keep a food journal and set behavioral goals for weight loss which was a component of Army MOVE!. Based on subject preference, the nurse coach called or e-mailed those in groups 2, 3, and 4 once a week for 12 weeks. If they failed to respond to e-mails for more than a week, the nurse coach attempted contact via phone. Coaching included modified motivational interviewing, goal setting, and reminders of upcoming MOVE! sessions, taking the study supplement, and scheduling an individual registered dietitian (RD) appointment.

We reserved the fifth arm for self-referred Soldiers who expressed concern over changes in their weight and ability to perform well on the Army Physical Fitness Test. These subjects attended 7 sessions dedicated to healthy living and behavior change to achieve weight and fitness goals. A RD led these sessions and provided one individual appointment. Group 5 subjects had weight loss, body composition, lipid profile, motivation, and adherence documented at baseline, 6 weeks, and 12 weeks; they did not undergo a dual-energy X-ray absorptiometry (DXA) scan as they did not receive the supplement.

The manufacturer prepared the product to our specifications with a 30-day supply per bottle. The research pharmacist created the computer generated randomization table, determined which number would represent the supplement group and which the placebo group, and kept the research team and the volunteers blinded to this activity. The pharmacist also maintained product dispensing and return records. She labeled and dispensed the product with instructions to take two pills, three times a day, 30 to 60 minutes before meals for maximum effectiveness. We provided subjects with an inexpensive pill container that held exactly six pills and fit easily in the military uniform pocket. Research team members queried subjects at each visit whether they thought they were taking the supplement or the placebo. Subjects returned for refills each month for a total of 3 months, to complete the 12-week intervention period.

Measures
A 17-item demographic tool captured relevant personal and family history as well as age, gender, ethnicity, military occupational specialty, weight history and status since
Motivation Inventory measured motivation at the three data collection points. This 40-item valid and reliable tool measures an individual’s tendency to persevere independent of situational reinforcement. This tool has predicted successful weight loss and may correlate with number of sessions attended in weight loss programs.16

Statistical Analysis

Missing data were imputed using machine-learning techniques (10 imputations). We compared change scores across study groups using general linear models, adjusted for covariates imbalanced at baseline. Conclusions were extracted using effect sizes and significance tests, at a false discovery rate (FDR) of 10%.17 Statistical significance was set at the \( p \leq 0.05 \) level.

RESULTS

Over the course of 30 months, 435 AD Soldiers enrolled in the study, with sample demographics reported in Table I. When comparing the 3 NHC groups to the CG, from baseline to follow-up (Weeks 6 + 12), beneficial intervention effects were observed for heel BMD (\( d = 0.3 \)), 25(OH)D (\( d = 0.43 \)), and FBS (\( d = -0.4 \)); the cutoff for relevance of effect size was Cohen’s \( d \) of \( \sim0.3 \).18 Despite replicated findings on 25(OH)D and FBS in the multiple imputation analysis (\( d = 0.28 \), and \( d = -0.32 \), respectively), once the FDR was applied at follow-up, no significant differences in change from baseline on any outcome remained (Table II). For the NHC + supplement group, there were no significant differences from baseline to follow-up on any outcome measure when compared to all other study groups (Table III). Overall attrition rate was highest between week 6 and 12 at 40%. An exploratory analysis found subjects under 28 years old and assigned to the NHC + supplement group were more likely to drop out during this time.

Adherence scores ranged from 8.92% (±4.19) for the CG to 53.78% (±22.56) for the self-referred group. The

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Mean (SD) or N (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (Years)</td>
<td>30 ± 8.2</td>
</tr>
<tr>
<td>Men</td>
<td>319 (73.4)</td>
</tr>
<tr>
<td>Race</td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>305 (70.1)</td>
</tr>
<tr>
<td>Black</td>
<td>83 (19)</td>
</tr>
<tr>
<td>Other</td>
<td>47 (10.8)</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
</tr>
<tr>
<td>Non-Hispanic or Latino</td>
<td>348 (80)</td>
</tr>
<tr>
<td>Hispanic or Latino</td>
<td>87 (20)</td>
</tr>
<tr>
<td>Married</td>
<td>309 (71)</td>
</tr>
<tr>
<td>Married With Children</td>
<td>178 (41)</td>
</tr>
<tr>
<td>Enlisted</td>
<td>396 (91)</td>
</tr>
<tr>
<td>Service Component</td>
<td></td>
</tr>
<tr>
<td>Army Active Duty</td>
<td>435 (100)</td>
</tr>
<tr>
<td>History Overweight</td>
<td>263 (61)</td>
</tr>
<tr>
<td>Lost Weight to Enter Army</td>
<td>166 (38)</td>
</tr>
</tbody>
</table>

TABLE I. Sample Demographics
## DISCUSSION

The primary outcome of weight loss proved difficult for all groups; there were no significant differences for primary or secondary outcomes across the four randomized groups of 335 Soldiers referred to Army MOVE!, who also volunteered to participate in this 12-week study. There were no significant results for 100 Soldiers who self-referred to our study because of concerns about their weight or meeting AR 600-9 standards, either. The hypothesis that a nurse health coaching intervention, delivered weekly by phone or e-mail, leads to greater success in meeting weight loss goals compared to the CG was not supported. A review of the published literature since 2012, reveals similar results from many studies, although interventions and populations are varied.\(^\text{19–23}\) One comprehensive review of nutrition interventions for weight loss in adult males found that delivery of quantitative information on diet and the use of self-monitoring and tailored feedback seem to be associated with better outcomes.\(^\text{24}\) However, limited evidence is available to describe Army efforts to address weight management using lifestyle interventions for Soldiers.\(^\text{25}\) The current AR 600-9 Body Composition Program\(^6\) lists an acceptable upper limit of body fat for the age group 21 to 27 years of 22% for men and 32% for women. The relatively minimal effect of Army MOVE! plus NHC from this study leaves a large pool of overweight and overfat Soldiers; at 12 weeks the average BMI was 31.4 kg/m\(^2\), (categorized as obese)\(^\text{10}\) and body fat was 29.9%. The disappointing adherence scores for all

### TABLE II. Baseline Measures by Group Assignment

<table>
<thead>
<tr>
<th>Measure</th>
<th>Control N = 86</th>
<th>Coach N = 81</th>
<th>Supplement N = 83</th>
<th>Placebo N = 85</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weight (lbs)</td>
<td>217.06 (33.32)</td>
<td>217.27 (38.86)</td>
<td>218.87 (30.54)</td>
<td>214.12 (30.25)</td>
</tr>
<tr>
<td>Height (in)</td>
<td>68.42 (3.85)</td>
<td>68.38 (3.62)</td>
<td>68.97 (2.93)</td>
<td>67.85 (3.47)</td>
</tr>
<tr>
<td>BMI</td>
<td>32.45 (3.19)</td>
<td>32.44 (3.65)</td>
<td>32.25 (3.2)</td>
<td>32.63 (3.02)</td>
</tr>
<tr>
<td>WC (in)</td>
<td>38.53 (3.87)</td>
<td>38.22 (4.16)</td>
<td>38.65 (3.4)</td>
<td>38.13 (3.48)</td>
</tr>
<tr>
<td>% Fat</td>
<td>31.43 (5.77)</td>
<td>31.36 (5.65)</td>
<td>31.48 (5.66)</td>
<td>32.73 (6.56)</td>
</tr>
<tr>
<td>25(OH)D ng/mL</td>
<td>21.9 (8.46)</td>
<td>21.35 (7.7)</td>
<td>22.69 (6.64)</td>
<td>22.92 (7.34)</td>
</tr>
<tr>
<td>FRS in mg/dL</td>
<td>108.72 (60.01)</td>
<td>109.52 (66.26)</td>
<td>112.23 (69.77)</td>
<td>114.23 (58.99)</td>
</tr>
<tr>
<td>HDL-CHOL mg/dL</td>
<td>40.87 (33.33)</td>
<td>41.91 (32.02)</td>
<td>49.29 (36.75)</td>
<td>50.29 (39.11)</td>
</tr>
<tr>
<td>LDL mg/dL</td>
<td>196.92 (57.37)</td>
<td>179.41 (28.91)</td>
<td>183.23 (34.6)</td>
<td>191.82 (24.6)</td>
</tr>
<tr>
<td>Trig mg/dL</td>
<td>96.33 (55.41)</td>
<td>102.99 (69.31)</td>
<td>109.03 (73.29)</td>
<td>111.03 (75.17)</td>
</tr>
<tr>
<td>HDL:CHOL</td>
<td>0.29 (0.02)</td>
<td>0.27 (0.05)</td>
<td>0.27 (0.05)</td>
<td>0.27 (0.05)</td>
</tr>
<tr>
<td>LDL:CHOL</td>
<td>1.31 (0.12)</td>
<td>1.32 (0.12)</td>
<td>1.32 (0.12)</td>
<td>1.32 (0.12)</td>
</tr>
<tr>
<td>HDL:LDL</td>
<td>0.29 (0.02)</td>
<td>0.27 (0.05)</td>
<td>0.27 (0.05)</td>
<td>0.27 (0.05)</td>
</tr>
<tr>
<td>FBS mg/dL</td>
<td>90.87 (8.47)</td>
<td>91.31 (9.02)</td>
<td>90.29 (8.64)</td>
<td>90.87 (9.02)</td>
</tr>
<tr>
<td>TChol mg/dL</td>
<td>184.99 (33.33)</td>
<td>186.92 (35.7)</td>
<td>179.41 (28.91)</td>
<td>183.23 (34.6)</td>
</tr>
<tr>
<td>DEXA BMD (g/cm²)</td>
<td>153.33 (21.19)</td>
<td>153.51 (20.18)</td>
<td>152.36 (26.58)</td>
<td>148.64 (18.86)</td>
</tr>
</tbody>
</table>

---

**Note:**
- BMI was 31.4 kg/m\(^2\), (categorized as obese)\(^\text{10}\) and body fat was 29.9%.
- The disappointing adherence scores for all...
groups reflect poor attendance at Army MOVE! and engagement in other study activities. This presents a confounding factor for success of the study as Soldiers cannot be expected to apply principles of weight management and healthy behaviors they have not had a chance to learn by missing critical program components. In addition to encouraging use of the buddy system to promote accountability and meet personal goals, both outpatient nutrition clinic and research staff recommended journaling, smartphone applications (e.g., MyFitnessPal), and mindfulness therapy.

The hypothesis that a nurse coaching intervention with an herbal supplement provided to Soldiers for 12 weeks would achieve more success with weight loss goals compared to other groups was not supported. Participants randomized to the supplement or placebo group were initially satisfied with a pill that might accelerate their weight loss efforts but they soon lost interest in taking the pill because they reported it “did not make them feel any different”. The supplement group experienced the highest dropout rate with 26 (31.3%) at 6 weeks and 47 (56.6%) at 12 weeks. The placebo group had the next highest dropout with 29 (34.1%) at 6 weeks and 42 (49.4%) at 12 weeks. This finding was unexpected because the 2011 Health Related Behaviors Survey results did not lead to any different. The supplement group had the highest mean level of 25(OH)D at that time and this level remained unchanged at the final measurement point.

Numerous investigators identified lower levels of physical activity, along with more sedentary behaviors, as a cause of overweight and obesity in the military. In a previous study, a similar pattern emerged when many Soldiers returning from overseas deployments reported less work and leisure activities during the year of deployment which manifested in a higher percent of body fat upon return. In the current study, approximately 50% of overweight Soldiers were on a profile limiting their physical training that likely contributed to a lack of engagement in physically demanding activities as part of their self-directed weight loss. The physical deconditioning that occurs from a lack of physical training makes young Soldiers vulnerable to repeat injuries, particularly stress fractures. We did not track incidence of stress fractures or other musculoskeletal injuries in this cohort over time. Soldiers on a profile are vulnerable for stress injuries once they return to former levels of activity, and low vitamin D stores compound the risk. Vitamin D remains a major area of research with many questions regarding links to overweight and obesity. Current literature suggests, overweight individuals commonly experience low vitamin D status secondary to adipose tissue deposition, leaving less in circulating body pools. The statistics for low vitamin D status in Soldiers are similar to the general population with a significant number of young adults categorized as insufficient (< 30 ng/mL) or deficient (< 20 ng/mL). We did note that as some participants lost weight, 25(OH)D increased. In contrast, the CG had least change in BMI and the least rise in 25(OH)D. We did not see any correlation between higher 25(OH)D and increased BMD. The rise in 25(OH)D may also be the result of the education received as part of the course content in Army MOVE!, as well as the reinforcement from the RD and nurse coach on a regular basis. Even the CG experienced an overall increase in the percent of Soldiers (15.9% to 21.4%) with a normal 25(OH)D level. One other observation is that the self-referred group which had the lowest BMI upon entry into the study, also had the highest mean level of 25(OH)D at that time and this trend remained unchanged at the final measurement point. One must keep in mind that this group received weekly educational sessions with an RD and RN that included content regarding links to overweight and obesity. Current literature suggests, overweight individuals commonly experience low vitamin D status secondary to adipose tissue deposition, leaving less in circulating body pools.

**TABLE III.** Control vs. Combined Coaching Groups, Week 12 Change Scores

<table>
<thead>
<tr>
<th>Measure (lbs)</th>
<th>Baseline Measures, Control (M(SD); n)</th>
<th>Control</th>
<th>Coach</th>
<th>Difference</th>
<th>Effect Size Cohen’s d</th>
<th>p</th>
<th>FDRa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weight</td>
<td>217.08 (33.32); 86</td>
<td>−2.22 (0.56)</td>
<td>−4.39 (0.66)</td>
<td>−2.17 (0.86)</td>
<td>−0.07</td>
<td>0.012</td>
<td>0.061</td>
</tr>
<tr>
<td>BMI</td>
<td>32.45 (3.19); 86</td>
<td>−0.32 (0.08)</td>
<td>−0.65 (0.1)</td>
<td>−0.33 (0.13)</td>
<td>−0.10</td>
<td>0.011</td>
<td>0.062</td>
</tr>
<tr>
<td>WC (in)</td>
<td>38.53 (3.87); 86</td>
<td>−0.31 (0.1)</td>
<td>−0.56 (0.12)</td>
<td>−0.25 (0.16)</td>
<td>−0.06</td>
<td>0.118</td>
<td>0.236</td>
</tr>
<tr>
<td>% Fat</td>
<td>31.43 (5.77); 86</td>
<td>−0.40 (0.18)</td>
<td>−1.02 (0.21)</td>
<td>−0.62 (0.28)</td>
<td>−0.11</td>
<td>0.026</td>
<td>0.071</td>
</tr>
<tr>
<td>25(OH) Dng/mL</td>
<td>21.9 (8.54); 63</td>
<td>1.32 (0.53)</td>
<td>3.67 (0.55)</td>
<td>2.35 (0.76)</td>
<td>0.28</td>
<td>0.002</td>
<td>0.032</td>
</tr>
<tr>
<td>FBS mg/dL</td>
<td>90.87 (8.47); 78</td>
<td>2.09 (0.76)</td>
<td>−0.65 (0.9)</td>
<td>−2.74 (1.18)</td>
<td>−0.32</td>
<td>0.020</td>
<td>0.065</td>
</tr>
<tr>
<td>TChol mg/dL</td>
<td>184.99 (33.3); 79</td>
<td>1.01 (2.21)</td>
<td>−1.5 (2.25)</td>
<td>−2.51 (3.15)</td>
<td>−0.08</td>
<td>0.426</td>
<td>0.523</td>
</tr>
<tr>
<td>Trig mg/dL</td>
<td>108.72 (61.01); 79</td>
<td>9.6 (4.74)</td>
<td>6.86 (4.98)</td>
<td>−2.74 (6.88)</td>
<td>−0.04</td>
<td>0.690</td>
<td>0.736</td>
</tr>
<tr>
<td>HDL: CHOL</td>
<td>0.29 (0.08); 79</td>
<td>−0.01 (0)</td>
<td>0 (0)</td>
<td>0.01 (0.01)</td>
<td>0.13</td>
<td>0.317</td>
<td>0.423</td>
</tr>
<tr>
<td>LDL mg/dL</td>
<td>116.89 (29.02); 79</td>
<td>1.23 (1.75)</td>
<td>−1.03 (2.09)</td>
<td>−3.16 (2.73)</td>
<td>−0.11</td>
<td>0.247</td>
<td>0.417</td>
</tr>
<tr>
<td>SMI Score</td>
<td>153.33 (21.19); 86</td>
<td>2 (1.85)</td>
<td>0.57 (1.41)</td>
<td>−1.43 (2.33)</td>
<td>−0.07</td>
<td>0.539</td>
<td>0.616</td>
</tr>
</tbody>
</table>

FBS, fasting blood sugar; HDL:CHOL, high-density lipoprotein:cholesterol ratio; LDL, low-density lipoprotein; TChol, total cholesterol; Trig, triglyceride. aFalse discovery rate 10%. bSMI, Self-motivation Inventory.
are not uncommon for military research based on the research team’s previous experience. However, it is unusual that the CG, which received no direct benefit from the research interventions beyond learning their body composition details, had the lowest dropout rate. Although all the reasons for the attrition are unknown, the younger enlisted Soldiers would often inform the study team that their Commander no longer supported their participation and they would not be returning to Army MOVE! or the study. Also, Soldiers experienced unexpected deployments and even service discharge after enrollment. We do not know how many of these Soldiers were discharged because of failure to meet AR 600-9. It is possible that basing our interventions around planned attendance at Army MOVE! diminished our ability to be successful when so many Soldiers did not take this responsibility seriously and often dropped out before receiving any benefit of participation. Participant compliance was examined as a whole and was not significantly related to any outcomes. Separate analysis of each component of participant compliance may identify whether journaling, goal setting, individual RD appointment, or class attendance was associated with improved outcomes. In addition, an exit survey to determine which component of the program was most helpful to participants might have been a valuable addition to the study protocol.

CONCLUSIONS

The results are congruent with current literature suggesting healthy lifestyle interventions often meet with disappointing results in research conducted with adult populations, to include the military. We have extended previous work on bone health and vitamin D with the correlation that overweight individuals may experience low vitamin D status. Full-scale implementation of AR 600-9 Body Composition Program (revised in July 2013) may result in greater support from Unit Leaders for future attendance at Army MOVE! The education provided to this group, of mostly young adults, about self-directed health behaviors may diminish chronic disease risk and the related socioeconomic burden leading to a productive and healthy career in the military. A recommendation stemming from this study is to identify new accessions with a history of overweight or of losing weight to enlist early and follow them closely for healthy weight maintenance as they begin their military career.

Brigade nurses are in an influential position at the Soldier level and need a situational awareness of the high rates of overweight and obese Soldiers so they can inform Unit leaders and establish metrics as part of the Unit wellness and readiness assessment. Assisting soldiers to attain/maintain healthy weight with education about good diet choices and staying physically active, and making referrals to nutrition experts as needed, may support normal vitamin D status and reduce risk of chronic disease.

ACKNOWLEDGMENTS

This research was sponsored by the TriService Nursing Research Program, 4301 Jones Bridge Road, Bethesda, Maryland 20814 (Grant No. HT9404-12-1-TS03,N12-007); Study period: March 2012–February 2015.

REFERENCES

23. Tate DF, Jackvony EH, Wing RR: A randomized trial comparing human e-mail counseling, computer-animated tailored counseling, and no counseling in an Internet weight loss program. JAMA 2006; 166(15): 1620–5.
Physical Screening Predictors for Success in Completing Air Force Phase II Air Liaison Officer Aptitude Assessment

Lt Col John Christopher McGee, USAF BSC*; Lt Col Eric Wilson, USAF BSC†; 2d Lt Haley Barela, USAF*; C1C Sharon Blum, USAF*

ABSTRACT  Objectives: Air Liaison Officer Aptitude Assessment (AAA) attrition is often associated with a lack of candidate physical preparation. The Functional Movement Screen, Tactical Fitness Assessment, and fitness metrics were collected (n = 29 candidates) to determine what physical factors could predict a candidate’s success in completing AAA. Methods: Between-group comparisons were made between candidates completing AAA versus those who did not (p < 0.05). Upper 50% thresholds were established for all variables with R^2 < 0.8 and the data were converted to a binary form (0 = did not attain threshold, 1 = attained threshold). Odds-ratios, pre/post-test probabilities and positive likelihood ratios were computed and logistic regression applied to explain model variance. Results: The following variables provided the most predictive value for AAA completion: Pull-ups (p = 0.01), Sit-ups (p = 0.002), Relative Powerball Toss (p = 0.017), and Pull-ups × Sit-ups interaction (p = 0.016). Conclusion: Minimum recommended guidelines for AAA screening are Pull-ups (10 maximum), Sit-ups (76/2 minutes), and a Relative Powerball Toss of 0.6980 ft × lb/BW. Associated benefits could be higher graduation rates, and a cost-savings associated from temporary duty and possible injury care for nonselected candidates. Recommended guidelines should be validated in future class cycles.

INTRODUCTION

According to the Air Force Career Field Education and Training Plan, “the Air Liaison Officer (ALO) specialty (13LXX) is the senior Tactical Air Control Party (TACP) member who functions as the primary advisor to the ground commander on Airpower. The ALO leads, plans, organizes, and supervises day-to-day TACP and Air Support Operations Center operations and personnel both in garrison and while deployed. The ALO represents the Joint/Combined Forces Air Component Commander as a supporting member of the Army Commander’s staff at the Battalion through Corps echelons in a coalition, joint, or interagency force. The ALO provides subject matter expertise to lead, plan, and execute Command-and-Control and terminal control of Air, Space, and Cyber operations in direct support of land component forces from Battalion through Corps, or as required, as part of a coalition, joint, or interagency force. The ALO may engage enemy forces utilizing advanced technologies and weapon systems to direct lethal and nonlethal fires and effects in close proximity to friendly forces as a Joint Terminal Attack Controller.”

Members of this elite career field must complete a multiphase rigorous training pipeline process with an overall selection rate less than 30%. Specifically, the Phase II ALO Aptitude Assessment (AAA) Course is a 5 day selection process to select those who will be eligible to enter Initial Qualification Training with the ultimate goal to eventually join the 13L career field. AAA is a physically and mentally demanding course, executed by the 93rd Air-Ground Operations Wing (AGOW), Moody Air Force Base, Georgia, and its associated tenant units. The AAA is designed to weed out candidates that do not meet the physical standards, and select those who prove they have the physical fitness, leadership capability, and mental toughness to eventually succeed in the career field.

Of the officer candidates selected to attend AAA, only a few are selected to continue through the rest of training pipeline process. In fact, many fail to simply complete the course. Several reasons for such a high attrition rate during this phase exist, however, a large number of candidates regularly do not complete training either due to an injury or inadequate physical preparation. For example, a retrospective analysis between the time frame of August 2012 to March 2014 included a total of six AAA Phase II cycles with 210 candidates attempting selection. Of those, 145 (69%) did not complete their cycles either due an illness/injury (n = 87/60%) or due to self-elimination (n = 58/40%) from lack of physical preparation. The remaining 65 (31%) candidates physically completed their cycle, but not all of those individuals were selected by the cadre from a holistic career field perspective (including leadership capability and mental toughness) to progress to the next phase. On review, it appeared that elimination factors due to injury and inadequate physical preparation could potentially be ameliorated through re-evaluation of minimum initial participation requirements and through improved screening of candidates. First, this may improve the value of the candidate pool by reducing attrition and improving the supply chain for such
high-demand human resource assets. A subsequent benefit could be the associated impact on finite fiscal resources. A more appropriate applicant pool could further translate into significant cost-savings associated with direct temporary duty/training costs and indirect cost savings associated with managing injuries incurred during the selection process.

Over the past 4 years, operational commanders within the Air Force Special Tactics, Battlefield Airman, and Base Defense communities has recognized the value of embedding Air Force Physical Therapists within their units to improve access to care, to ensure combat capability, to enhance member lifetime health and sustainability, and to optimize human performance. Since 2011, a total of 12 line officer billets throughout the Air Force have been converted to embed therapists in this new model, with further expansion planned and programmed. In many cases, embedded therapists incur additional duties beyond their clinical roles as Directors of Human Performance (DHP). As a result, they are charged to investigate, interpret, and implement interventions to counter issues such as improving career field pipeline viability as described in this investigation.

In 2012, the 93rd AGOW DHP began using two additional multimodal instruments as a more comprehensive physical screening mechanism in addition to the current TACP Physical Ability Stamina Test (PAST) to better understand and access low AAA course completion rates: (1) The Functional Movement Screen (FMS) and (2) The Tactical Fitness Assessment (TFA). The present PAST includes the 1½ mile timed run, 1 minute timed push-ups, 2 minute timed sit-ups, and maximum pull-up repetitions. This portion of the test was maintained and is the current standard for AAA acceptance.

The FMS is a validated method of predicting injury risk and is composed of seven different movement tests, each of which assesses a different functional body movement. Components include the Deep Squat, Hurdle Step, In-line Lunge, Shoulder Mobility, Active Straight Leg Raise, Trunk Stability Push-up, and Rotary Stability components (Fig. 1). Overall, the movements are designed to reveal impairments of movement and/or strength for which an individual may be compensating, and would not be noticeable otherwise. The deep squat is an evaluation of total body mechanics and assesses bilateral, symmetrical, functional mobility of the hips, knees, and ankles. The hurdle step evaluates stride mechanics while performing a stepping motion, and assesses bilateral function mobility and stability of the hips, knees, and ankles. The third component, the in-line lunge, assesses the degree of stability while performing rotational, decelerating, and lateral type movements. It challenges the trunk and extremities to resist rotation, while also assessing hip and ankle mobility and stability, quadriceps flexibility, and knee stability. The active straight leg raise assesses the ability to disassociate the lower extremity from the trunk while maintaining stability in the torso. It tests active hamstring and gastrocnemius flexibility while maintaining a stable pelvis and core. The trunk stability push-up evaluates the stability of the core and spine in an anterior and posterior plane and assesses trunk stability in the sagittal plane. The shoulder mobility test evaluates bilateral and reciprocal shoulder range of motion and combines internal rotation with adduction of one shoulder and external rotation with abduction of the other. This motion requires scapular mobility with thoracic spine extension. The rotary stability test requires neuromuscular coordination and energy transfer from one segment of the body to the other through the torso and tests multiplanar trunk stability during a combined upper and lower extremity motion.

The scoring of each FMS component is predicated on the evaluator’s subjective assessment of efficiency of movement (0–3 ordinal scale) with which an individual can perform a motion and the amount of pain they may experience. A component score of 0 indicates the movement caused pain, 1 indicates the movement could not be performed, 2 indicates it could be performed only with associated compensation, and a score of 3 indicates the movement was performed perfectly. Even if a movement cannot be performed, this earns a score of 1 so long as the subject does not experience any pain.

FIGURE 1. 7-functional movement screen components.
While some exercises are performed on both sides (e.g., an in-line lunge with the left leg, and one with the right), only the lower score contributes toward the total. The maximum composite score that can be achieved on the FMS is 21 points. For the FMS, a cumulative score of 14 or less has been associated with increased susceptibility to injury as opposed to those who scored higher. This has been substantiated by experiments conducted on intercollegiate athletes, professional football players, and other military branches such as the Marines. For example, one investigation conducted on 38 female intercollegiate athletes who competed in soccer, volleyball, and basketball with FMS scores ≤14 had a significant injury association (p = 0.0496) with 69% sustaining an injury during their respective season. Similarly, in separate studies conducted on 847 Marine officer candidates and 46 professional football players, individuals demonstrating a composite FMS score ≤14 were 1.5 to 1.9 times more likely to incur a training injury. Because of its demonstrated success in predicting injury, the 93rd AGOW DHP chose to use the FMS as a proxy for predicting successful completion of AAA.

The TFA by contrast is a novel battery of five physical tests that were selected by the DHP based on moderate evidence in performance-based literature and their ability to mimic the performance requirements of ALOs. Selected components for the TFA consisted of the Pro-Agility Shuttle Run, Broad Jump, a 2 to 3 kg Powerball Toss for distance, Maximum Core (front/back plank core ratio) Endurance, and a 300-yard Shuttle Run and Repeat exercise. The (Pro) agility test (Fig. 2) is also commonly referred to as the 5-10-5 timed shuttle. It is a test of speed, explosion, body control, and agility. The individual runs 5 yards, touches the line with their hand, returns to the starting line, and repeats this action with a 10-yard and again a 5-yard interval. The standing broad jump evaluates explosive lower body power and is performed when an individual jumps as far as they can with both feet and lands without falling backward. Lockie et al measures the relative distance of the jump with the following formula: (relative standing broad jump = jump distance × body mass). This is a measure of the work done by the subject when executing the jump, and can be used when comparing results between subjects of different weight. In addition, there is a moderate correlation between the standing broad jump and the pro-agility shuttle: r = (0.580–0.784) and performance in the standing broad jump is predictive of performance in the pro-agility shuttle. The power ball toss tests upper body power. It is performed when the individual pushes a 2 to 3 kg weighted ball up and out from a kneeling position. The toss has also been correlated by Sharrock et al to core strength as measured by the traditional double leg lowering test (p = 0.023). Next, the front and back plank endurance test assesses core endurance and is presented as a ratio to compare the two. Finally, the 300-yard shuttle assesses aerobic conditioning and recovery. The purpose of this analysis was to obtain a baseline for understanding the potential predictive ability of the PAST, FMS, and TFA components for candidates completing the Phase II AAA. In addition, this analysis (based on one AAA cycle) should provide preliminary evidence to identify screening parameters that could ultimately prove effective in reducing attrition in the AAA pipeline. Validation of these and other parameters alone or in combination will require ongoing assessment of future AAA cycles alongside cost-analysis before any strategic policy changes should be fully supported. Nonetheless, given the dynamic and fluid operational military environment, our findings could be translated and monitored locally over subsequent course cycles as a performance initiative nested as part of a larger and more longitudinal based prospective investigation.

METHODS

PAST, FMS, TFA, and ruck march data for this analysis were collected as routine surveillance for one AAA course cycle by the 93rd AGOW DHP upon Phase II candidate arrival (n = 29). During this process, the PAST data were collected by experienced AAA cadre as had been performed in all previous course cycles. Additionally, the DHP trained
the cadre how to administer the FMS and TFA, and was always on location to monitor. Furthermore, this was the fifth AAA cycle that each cadre had participated in the FMS/TFA screening process and collection was always preceded by refresher training the week before. The DHP was also certified by Functional Movement System to administer the FMS. On arrival on the day preceding the course, all candidates were administered the FMS. The following morning the PAST was administered in the following sequence: pull-ups, sit-ups, push-ups, and 1½ mile run. A 5-minute break was allowed between each component of the PAST. On completion of the PAST, a 15-minute hydration break was allowed and the TFA was subsequently administered in three phases. This first phase included the agility/power components of the 5-10-5 shuttle run, the powerball toss, and the broad jump. Candidates were allowed two attempts for each event with the second attempt performed immediately after the first. A 3-minute rest was allowed between each component and the average of those scores was used for analysis. Candidates then proceeded to the second phase of exercises focusing on the core. Fifty percent of the candidates completed the front plank core component while the other half completed the back plank core component. Only one attempt was allowed for each component. After a 5-minute rest, the candidates switched and completed the opposite component. Finally, the candidates completed the third phase of the TFA which was the anaerobic portion of the assessment. This phase consisted of the 300-yard shuttle run. Two attempts were allowed with a 5-minute break between each, and the average of those scores was used for analysis. This method was chosen overall to ensure efficiency in data collection as to not impinge on the existing AAA curriculum, while also maximizing proper recovery. It should also be noted that the DHP and cadre made additional changes to this entire course cycle by standardizing and allowing greater rest time intervals between each assessment criteria. The rationale was to permit greater recovery, and permit full hydration which may be associated with injury risk reduction. Furthermore, candidates were required to eat all portions of their meals-ready to eat for breakfast, lunch, and dinner in order to ensure enough calories were consumed for the intense physical demands of the course, and to promote tissue rebuilding.

Next, all measures of central tendency and dispersion were computed on baseline demographics of age, gender, height, weight, and PAST results. Comparisons were then made on performance component items between candidates completing AAA ($n = 13$) versus those who self-eliminated ($n = 8/27.5\%$) due to physical unpreparedness or injury. Of note, out of the 13 (44\%) who physically completed the course; only 9 (31\%) individuals were selected to continue to the next phase. 8 (27.5\%) other individuals did not complete the training for reasons other than self-elimination due to physical unpreparedness or injury and were not included in this analysis. For the comparisons, $χ^2$ tests ($p < 0.05$) were used for discreet variables and $T$-tests ($p < 0.05$) for continuous variables.

Next, correlation coefficients were computed between groups for FMS composite scores and all TFA components. Variables with $R^2 < 0.8$ were further analyzed due to their low correlation or apparent differences. Of these, thresholds were established in the upper 50\% of the variable range and the data were converted to binary form for counts (0 = did not attain upper 50\% range, 1 = did attain upper 50\% range). From there, we computed positive likelihood ratios +LRs on the variables that were significant ($p < 0.05$) between groups for course completion. Finally, logistic regression was used to explain variance and to determine predictive ability between these variables.

Also, per DHP request, we presented normative data expressed as a ratio of weight thrown in the Powerball Toss to total body weight and compared between groups. Similarly, we presented normative data and compared distance jumped in broad jump as a percentage of height between groups (Relative Broad Jump).

**RESULTS**

Overall, there were no significant differences observed between groups for age, gender, height, and weight. However, the individuals who ultimately completed AAA had significantly better run times and performed better in push-ups, sit-ups, and pull-ups at baseline. Significant findings are listed in Table I. On average, those who completed the course ran the 1½ mile approximately 45 seconds (7.5\%) faster. In addition, they also completed an average of 13 more push-ups (>24\% more), and 20 more sit-ups (>22.5\%)

<table>
<thead>
<tr>
<th>Event</th>
<th>Status</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Confidence Interval (±)</th>
<th>$p$ Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Run (Seconds)</td>
<td>Completed</td>
<td>544.8</td>
<td>35.24</td>
<td>18.77</td>
<td>0.0193</td>
</tr>
<tr>
<td></td>
<td>Noncompleted</td>
<td>588.6</td>
<td>42.66</td>
<td>39.46</td>
<td></td>
</tr>
<tr>
<td>Push-Ups</td>
<td>Completed</td>
<td>54.6</td>
<td>9.04</td>
<td>4.81</td>
<td>0.245</td>
</tr>
<tr>
<td></td>
<td>Noncompleted</td>
<td>41.7</td>
<td>11.116</td>
<td>10.28</td>
<td></td>
</tr>
<tr>
<td>Sit-Ups</td>
<td>Completed</td>
<td>90.3</td>
<td>12.00</td>
<td>6.40</td>
<td>0.0012</td>
</tr>
<tr>
<td></td>
<td>Noncompleted</td>
<td>70</td>
<td>9.933</td>
<td>9.19</td>
<td></td>
</tr>
<tr>
<td>Pull-Ups</td>
<td>Completed</td>
<td>12.87</td>
<td>3.81</td>
<td>2.03</td>
<td>0.0039</td>
</tr>
<tr>
<td></td>
<td>Noncompleted</td>
<td>7.4</td>
<td>2.637</td>
<td>2.44</td>
<td></td>
</tr>
</tbody>
</table>

**TABLE I.** Significant Baseline Demographics (Completed Course Versus Noncompletion)
more) in the time allotted. Finally, individuals who completed AAA also performed an average of 5 more (42% more) repetitions in the maximum pull-ups section than those who self-eliminated.

Overall 13 variables to include composite FMS score, Pro-Agility, Broad Jump, Powerball Toss, 300 Shuttle, Front Core, Back Core, Run, Push-Ups, Sit-Ups, Pull-Ups, Relative Powerball Toss, and Relative Broad Jump were not highly correlated ($R^2 < 0.8$) between the groups suggesting that there were inherently different. At this stage we wanted to be most inclusive before final comparisons and stepwise regressions were performed. Composite FMS scores were found significant between groups with $\chi^2 = 0.03$. Upper 50% thresholds were then established for the remaining 12 variables and each individual was assigned a binary count score ($0 = \text{did not attain threshold}, 1 = \text{did attain threshold}$). $\chi^2$ tests completed on those frequency counts between groups demonstrated significance ($p < 0.05$) for all of the variables except for the Back Core ($p = 0.49$) and Front Core ($p = 0.07$) exercises. $+LR$s for course completion were then calculated for achieving above the 50% threshold on each variable. The following six variables all ended up having $+LR$s $> 1.2$ and were therefore entered into the final logistic regression along with the FMS composite score: Relative Powerball Toss, Relative Broad Jump, 1½ Mile Run, Pull-Ups, Push-Ups, and Sit-Ups with $+LR$s ranging from 3.47 to 5.18.

Table II demonstrates the six variables with the highest likelihood for AAA completion. Out of those, final logistic regression modeling demonstrated the following as best physical predictors of successfully completing AAA. “Pull-ups > 11.8 reps max, Sit-ups > 85 reps/2 min, Relative Powerball Toss of 0.7843 ft x lb/BW and there was a demonstrated interaction effect between sit-ups and pull-ups. The final logistic regression model can be observed in Figure 3.

It should be noted that the Composite FMS score was eliminated from the final model on the first step with $\text{Prob} > \text{ltl} = 0.09$.

| Component         | Upper 50% Cut-Off | $+LR$s | Prob $> |t|$ |
|-------------------|-------------------|--------|---------|-----|
| Relative          | 0.7843            | 4.81   | 0.0174* |
| Powerball Toss    | Ft x lb/BW        |        |         |     |
| Relative Broad Jump| 1,479.5 ft x lb  | 4.81   |         |     |
| Run               | 9:19 Minutes      | 3.47   |         |     |
| Pull-Ups          | 11.80 Repetition  | 4.44   | 0.0100* |
| Push-Ups          | 51.12 Repetition  | 4.71   |         |     |
| Sit-Ups           | 85 Repetition     | 5.18   | 0.0022* |

DISCUSSION

Our results should provide early justification to inform a much larger prospective validation study examining potential minimum cutlines for acceptance to AAA specifically for Pull-ups, Sit-ups, and Relative Powerball Toss as criteria to admit Phase II AAA candidates. $+LR$s found (~3–5) for these variables improve the post-test probability of completion 20 to 30% assuming candidates had a coin-flip chance (50% pretest probability) of completing AAA at the start of the cycle. If these results were to remain stable, a reasonable new minimum could be established as within one standard error of measurement below the current upper 50% thresholds from this investigation. Under these guidelines, minimum entrance requirements would therefore be 10.5 (rounded to 10) max Pull-ups, 75.7 Sit-ups/2 minutes (rounded to 76), and a Relative Powerball Toss of 0.6980 ft x lb/BW. For Pull-ups, the recommended 10 max Pull-ups is much higher than the minimum of 6 that is required by the current TACP PAST in order to be accepted into AAA Phase II. Similarly, the recommended cut-off of 76 Sit-ups is also much higher than the 48 Sit-up minimum required by the current PAST standards. The Powerball Toss is currently not a metric included in the prescreening of candidates but could be added to the current PAST requirements without detracting too much in terms of assessment timing and footprint.

Overall, a validated minimum requirement adjustment might substantially lower attrition and save the Air Force significant costs associated with temporary duty and possible injury care for nonselected or self-eliminated candidates while also permitting the cadre to concentrate more on candidates that are likely to complete the course. At a minimum, knowledge of the adjustments might enable candidates to better prepare physically for AAA beforehand. Furthermore, the training pipeline value could be improved through significantly increasing the viable supply chain numbers associated with these high-demand human resource assets. For example, all individuals ($n = 8$) that self-eliminated in this particular course cycle fell below the newly recommended guidelines in several components which may have been translated into considerable direct and indirect savings/benefits if their slots were filled instead with individuals who were better prepared for the physical rigors of the course. On the other hand, only three of the selected candidates demonstrated deficiencies with a single component (Pull-ups) as compared to the newly recommended guidelines. Two of those candidates missed the new threshold for this component by one repetition and the
incorporating any of the components from the current TFA. 1½ mile run time requirement is 10:42 minutes, and the consistency of these results over time. The current minimum standards for those events provided consistency, and it would therefore be advised to maintain the 1½ mile run within the current PAST at this time, particularly those in charge of squadron fitness and optimizing human performance. Finally, the substantial reduction of injuries associated with this cycle may also provide valuable introspective on the changes that were made in the course regarding rest breaks, hydration, and nutrition, particularly since these were the only differences made to the course from previous cycles. To reiterate, however, validation of this analysis requires a more prospective and robust sample size inclusive of subsequent graduating cycles to see if these recommendations and prevention interventions bear consistency, along with an associated cost analysis.

Additionally, it was determined that the FMS is not a useful proxy for predicting course completion as hypothesized. Nonetheless, it should be noted that performance prediction was never the intent of the FMS developers and that this finding in no way detracts from the FMS’s value as an injury prediction screen. Therefore it should be maintained as part of the DHP repertoire as an injury screen. Furthermore, additional work could be added to the body of literature assessing the FMS’s capability in predicting injury in AAA candidates and other tactical athletes. Finally, these findings should not fully preclude the use of the push-ups or the 1½ mile run within the current PAST at this time, particularly given the low sample size associated with this investigation, and it would therefore be advised to maintain the current minimum standards for those events provided consistency of these results over time. The current minimum 1½ mile run time requirement is 10:42 minutes, and the current push-up requirement is 40 repetitions/minute. Furthermore, the results from this investigation do not support incorporating any of the components from the current TFA at this time except for the Powerball Toss, but this leaves room for the evolution of other metrics if high attrition due to physical self-elimination or injury continues as a normative trend despite adoption of these recommended changes.13

Finally, guidelines suggested from this investigation, and if validated, would only serve to predict which candidates are most physically capable of completing AAA Phase II selection, and cannot predict who will be selected. Selection into the ALO career field is also based on mental fortitude, ability to act under stress, leadership, and other traits that cannot be solely measured by a physical performance or fitness test.

REFERENCES


Patient Litter System Response in a Full-Scale CH-46 Crash Test

Charles A. Weisenbach, MS*†; Tyler Rooks, MS*; Troy Bowman, BS*†; Vince Fralish, BS*; B. Joseph McEntire, MS*

ABSTRACT  U.S. Military aeromedical patient litter systems are currently required to meet minimal static strength performance requirements at the component level. Operationally, these components must function as a system and are subjected to the dynamics of turbulent flight and potentially crash events. The first of two full-scale CH-46 crash tests was conducted at NASA’s Langley Research Center and included an experiment to assess patient and litter system response during a severe but survivable crash event. A three-tiered strap and pole litter system was mounted into the airframe and occupied by three anthropomorphic test devices (ATDs). During the crash event, the litter system failed to maintain structural integrity and collapsed. Component structural failures were recorded from the litter support system and the litters. The upper ATD was displaced laterally into the cabin, while the middle ATD was displaced longitudinally into the cabin. Acceleration, force, and bending moment data from the instrumented middle ATD were analyzed using available injury criteria. Results indicated that a patient might sustain a neck injury. The current test standard fails to account for higher impact accelerations, environments are considered the simple approach of a static test standard and does not maintain restraint control of its patients. It is unknown if a modern litter system, with components tested to the same static criteria, would perform differently. A systems level dynamic performance requirement needs to be developed so that patients can be provided with protection levels equivalent to that provided to seated aircraft occupants.

INTRODUCTION

In the early 1950s, the Office of the Surgeon General of the U.S. Army recognized the importance of air mobility in medical evacuation and the benefit of internal transport of patients in helicopters.1 Over the course of the last half century, significant improvements have been made in military rotorcraft design to improve occupant protection and safety. These improvements have resulted in improved test standards for occupant protection. Despite these vehicle level and subsystem level improvements and test standards, the performance requirements for litter systems have received little modernization. Knowledge gained from the implementation of aircraft level crew and seated passenger occupant protection strategies have yet to be incorporated in the design of occupant protection for litter-borne patients.

U.S. Military aeromedical patient litter systems commonly consist of a litter secured in a litter support structure mounted inside an aircraft. Currently, litters and the litter support structure are statically tested to standard force levels outlined in MIL-A-8865B: Airplane Strength and Rigidity Miscellaneous Loads.2 Standard force levels are derived using Newton’s Second Law (force = mass x acceleration) by multiplying litter and occupant mass (133.4 kg) by an acceleration load factor of 8 G forward, 1.5 G lateral, 4.5 G vertically down, and 2.0 G vertically up.2 Litters themselves are tested to two static criteria described in the MIL-L-49511B standard.3 The first criterion states that for a litter supported at four points (0.95 m from the center of each pole) with a distributed load of 889.64 N there will be no more than 3.18 cm downward deflection of the poles.3 The second criterion requires the litter to be resting on its stirrups with a 7,117.15 N load applied for 5 minutes.3 This load is applied through a 45.72 cm wide board positioned crosswise on the poles at the center of the litter. Upon removal of the load, the litter shall not exhibit a permanent deformation in excess of 1.59 cm.3 From an operational perspective the litter system is subjected to the dynamics of turbulent flight, hard landings, and potentially crash events. When dynamic environments are considered the simple approach of a static test standard fails to account for higher impact accelerations, rate of acceleration onset, pulse duration, and material rate dependence.

Only two previous full-scale dynamic crash tests investigating aeromedical patient litter systems and patient response have been identified. In 1963, Schamadan et al investigated the crashworthiness of a three-tiered strap and...
pole litter system during a controlled dynamic crash test of a U.S. Army H-21 helicopter. Patients were simulated during the test using two anthropomorphic test devices (ATDs), which occupied the top and middle litter positions. Results of this test revealed that the litter system “failed completely,” providing inadequate litter occupant restraint under severe but survivable crash conditions. During the crash test, the two ATDs failed to remain completely restrained, which resulted in the ATDs displacing forward in the airframe. Schamadan et al recommended that both static and dynamic testing of the litter system be conducted at energy levels consistent with conditions expected in moderately severe but potentially survivable accidents. A follow-on crash test of a second H-21, modified to accept a litter system designed for the UH-1D helicopter, yielded similar results. A combined analysis of both H-21 crash tests and an assessment of litter system specifications concluded that litter systems were inadequate to withstand moderate crash accelerations. The static strength performance requirements for the legacy litter systems used in the H-21 crash tests have not been updated in the current MIL-A-8865B. Although legacy and present day litter systems are similar in design specifications, no recent studies have investigated their performance during full-scale dynamic tests.

In 2013, the National Aeronautics and Space Administration (NASA) conducted a full-scale crash test of a retired CH-46 helicopter as part of their Transport Rotorcraft Airframe Crash Testbed (TRACT) research program. The test, known as TRACT 1, was the first of two CH-46 crash tests planned and included collaboration with the U.S. Army Aeromedical Research Laboratory (USAARL), Naval Air Systems Command, and the Federal Aviation Administration, as well as industry partners. For the crash test, USAARL designed an experiment to assess crashworthiness of litter systems and their components in order to investigate litter-borne occupant safety during a severe but potentially survivable crash scenario.

METHODS

A retired CH-46E Sea Knight fuselage was provided to NASA by the Navy CH-46 Program Office (PMA-226) for this experiment. The subject airframe is a tandem rotor helicopter capable of carrying five crew members and 25 troops. In preparation for this test, the stub wings, landing gear, engine, rotors, transmission, and vertical tail were removed, resulting in a bare fuselage weight of 1,133.98 kg. The fuselage was fitted with the necessary equipment and hardware for eight onboard experiments, including the USAARL litter system experiment. When fully instrumented and outfitted, the total weight of the fuselage was 4,672.00 kg.

The CH-46 crash test was conducted at the Landing and Impact Research Facility at NASA Langley Research Center. The impact surface was flat soil (a mixture of sand and clay) 36.58 m long, 6.10 m wide, and 0.61 m deep. The impact surface composition was chosen based on military and civil-
Each litter position was occupied with an ATD for the test event (Fig. 2). The middle litter was populated with an instrumented 50th percentile Hybrid III pedestrian ATD (Humanetics, Plymouth, Michigan). Load cells were located within the ATD’s neck, lumbar spine, and both femurs. Accelerometers were located in the head, chest, and pelvis of the ATD. All ATD sensors were arranged so that the positive x axis was in the ATD’s posterior to anterior direction, positive z axis was in the ATD’s superior to inferior direction, and positive y axis was in the left to right direction. The lower and upper litter positions were outfitted with Grumman–Alderson Research Dummy (GARD) ATDs (Humanetics), each with a mass of approximately 91.72 kg. The GARD ATDs were not instrumented, but provided inertial mass to the litter system simulating realistic occupant loading. The lower and middle ATDs were positioned with the head toward the airframe nose, whereas the upper ATD was positioned with the feet toward the airframe nose (representing a potential nonstandard operational loading condition anecdotaly reported by flight medics).

Patients are typically secured with two straps that encircle the litter and patient, one around the patient’s chest and the second around the thigh. Strap placement can vary depending on the injury location(s) of the patient. For this test, the straps were located at the optimum chest and thigh positions. The bottom and middle litters both used standard straps. The Talon II litter is configured with two integrated patient restraint straps, which were utilized to restrain the upper litter occupant. A webbing load cell was attached to each occupant restraint strap to record the dynamic tensile loads resulting from the crash event.

Two DTS G5 (Diversified Technical Systems, Inc., Seal Beach, California) data acquisition systems and a MIRO high-speed camera (Vision Research, Wayne, New Jersey) were mounted on the fuselage wall opposite the litter system. These units were located on a padded shelf fabricated by NASA personnel and mounted to the airframe specifically for this test.

Sensor data were collected on the G5 data acquisition systems at a sampling frequency of 10,000 Hz. MIRO high-speed video images were collected at a sampling frequency of 500 Hz. High-speed video and sensor data were synchronized using inter-range instrumentation group time codes. Sensor data collected during the test were analyzed using custom MATLAB (Math Works, Natick, Massachusetts) scripts. All electronic sensor data were filtered using a fourth order phaseless Butterworth filter designed in accordance with SAE J211 channel frequency class (CFC) recommendations for surface vehicles and ATDs.7

DC offsets were removed from all data before data analysis. Data were time shifted to set \( T = 0 \) with the instant of initial airframe contact with the ground. Ground contact was determined as the point of time where the aft floor-mounted vertical accelerometer exceeded a threshold of 1 G. Structural component data were analyzed from 0.05 seconds before impact to 0.15 seconds after impact. ATD data were analyzed from 0.05 seconds before impact to 0.20 seconds after impact. Injury criteria were assessed for the instrumented ATD based on body region and sensor regions.8–10

RESULT

Airframe velocity of 7.62 m/s vertically and 10.06 m/s longitudinally was reported by NASA during the crash.6 At impact, airframe pitch was 2.5° nose up, with 0.5° roll, and less than 1.0° yaw.6 Impact accelerations measured on the airframe floor, directly forward and aft of the patient litter system are reported in Table I.

Analysis of high-speed video revealed the timeline of litter system structural failure during the impact event (Fig. 3). Bending of all litter poles occurred immediately on fuselage impact, with the lower litter stirrups making contact with the airframe floor at \( T + 0.043 \) seconds. Litter support structure collapse occurred when the top aft outboard stanchion
bracket was observed to fail at $T + 0.077$ seconds, immediately followed by the failure of the top aft inboard bracket observed at $T + 0.079$ seconds. The middle aft outboard stanchion bracket failed shortly after with complete collapse of the litter system before $T + 0.200$ seconds. Post-test inspection of the litter support structure (Fig. 4) confirmed failure of both top aft inboard and outboard stanchion brackets. Additionally, post-test inspection identified the failure of the following stanchion brackets: top forward outboard, middle forward outboard, and middle aft outboard. All three litters exhibited fractures of the litter poles (Fig. 4).

Significant displacement of all three ATDs was observed in the high-speed video (Fig. 3). At the time of impact, the upper ATD was observed to have already displaced aft from its original position during the freefall acceleration ($T + 0.001$ seconds in Fig. 3). This aft displacement was the result of...
of the aircraft’s release from the two pull-back cables. The upper ATD displaced forward and laterally at $T + 0.175$ seconds and by $T + 0.683$ seconds, the upper ATD was observed to pivot laterally across the fuselage due to the litter and ATD remaining attached to the forward outboard stanchion bracket. At the time of aircraft impact, both the middle and lower ATDs remained in their original positions ($T + 0.001$ seconds in Fig. 3). The middle ATD started to displace forward at $T + 0.175$ seconds and by $T + 0.683$ seconds was observed contacting the airframe floor at a point forward of the litter system (Fig. 3). The bottom ATD displaced vertically during the crash, making contact with the airframe floor at $T + 0.053$ seconds. The middle ATD made contact with the bottom ATD at $T + 0.083$ seconds followed by the top ATD making contact with the middle ATD at $T + 0.121$ seconds.

Accelerations for the litter support structure measured during impact are reported in Table I. The peak forces measured in the occupant restraint belts ranged from 100.75 N to 675.02 N for five of the six load cells. Measured ATD accelerations during impact are reported in Table II. Measured ATD forces and moments during impact are reported in Table III. Calculated peak resultant neck shear was 573.43 N and peak resultant neck moment (excluding torsion) was 75.23 Nm. Peak resultant lumbar spine shear force was 2,619.24 N and calculated left femur peak resultant shear force was 1,243.87 N. Calculated peak resultant bending moments (excluding torsion) for the left and right femur were 453.23 Nm and 203.10 Nm, respectively. ATD injury criteria assessments are reported in Table IV.

During the impact crash event, a number of sensor signals were lost preventing data analysis. On the litter system the $x$-axis sensor at the forward outboard stanchion bracket as well as the $y$-axis and $z$-axis sensors at the aft inboard stanchion bracket were lost. The aft litter belt sensor securing the upper litter ATD was also lost. For the instrumented ATD the sensors measuring neck force in the $z$ axis, right femur force in the $y$ axis, and ATD lumbar spine moment in the $x$ axis were lost.

**DISCUSSION**

Post-crash analysis of collected high-speed video, litter system accelerations, and visual examination of the litter system components indicated the three-tiered strap and pole litter system experienced substantial overmatch, as evidenced by severe deformation and structural failure during the crash sequence.

**TABLE II. ATD peak accelerations (G)**

<table>
<thead>
<tr>
<th></th>
<th>Ax (G)</th>
<th>Time (s) **</th>
<th>Ay (G)</th>
<th>Time (s) **</th>
<th>Az (G)</th>
<th>Time (s) **</th>
</tr>
</thead>
<tbody>
<tr>
<td>Head Acceleration*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maximum</td>
<td>9.49</td>
<td>0.084</td>
<td>33.96</td>
<td>0.115</td>
<td>27.81</td>
<td>0.093</td>
</tr>
<tr>
<td>Minimum</td>
<td>−30.54</td>
<td>0.113</td>
<td>−6.07</td>
<td>0.090</td>
<td>−29.14</td>
<td>0.084</td>
</tr>
<tr>
<td>Chest Acceleration*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maximum</td>
<td>18.53</td>
<td>0.093</td>
<td>41.98</td>
<td>0.090</td>
<td>22.04</td>
<td>0.090</td>
</tr>
<tr>
<td>Minimum</td>
<td>−7.92</td>
<td>0.080</td>
<td>−6.26</td>
<td>0.143</td>
<td>−26.62</td>
<td>0.081</td>
</tr>
<tr>
<td>Pelvis Acceleration*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maximum</td>
<td>38.02</td>
<td>0.081</td>
<td>20.99</td>
<td>0.076</td>
<td>49.36</td>
<td>0.075</td>
</tr>
<tr>
<td>Minimum</td>
<td>−198.90</td>
<td>0.075</td>
<td>−11.46</td>
<td>0.080</td>
<td>−34.05</td>
<td>0.078</td>
</tr>
</tbody>
</table>

*Data filtered at CFC 1,000. **Time after impact.
As a result of structural collapse of the litter system, the ATDs experienced secondary response impacts, which were observed in the acceleration data collected. Both top aft stanchion brackets failed during impact, which resulted in a compromise of occupiable space for the lower and middle patients. Litter poles on all three litters exhibited bending failures, indicating inadequate structural strength for survivable crash loads. This suggests that the stanchion bracket and litter pole structural strength requirements should be revised to reflect the expected loads of survivable crashes. It should be noted that two different litter designs were used to represent potential operational variation. The difference in litters as well as their arrangement (i.e., Talon II on top) may have played a role in the failure of the litter system and warrants further study.

Review of the high-speed video revealed significant displacement and flail of the ATDs. All patient restraints were tightened before testing; however, at release, the aircraft experienced a forward acceleration (less than 1 G) and the occupants experienced weightlessness as the aircraft falls. The upper ATD restraint system did not provide adequate retention in this condition and allowed the upper ATD to displace unnaturally before the ground impact. The middle and bottom ATDs did not appear to displace before impact. Post-impact, the upper ATD experienced the greatest amount of displacement as it pivoted inboard about its feet after initial vertical displacement. The middle ATD initially displaced vertically during the impact and then was partially ejected from the litter forward into the fuselage, coming to rest on its side. This may have been caused by the interaction with the upper ATD during the litter system collapse. The bottom ATD remained relatively stationary on the litter compared to the other ATDs, only displacing vertically.

### TABLE III. ATD Peak Forces (N) and Moments (Nm)

<table>
<thead>
<tr>
<th></th>
<th>( F_x (N) )</th>
<th>Time (s)**️</th>
<th>( F_y (N) )</th>
<th>Time (s)**️</th>
<th>( F_z (N) )</th>
<th>Time (s)**️</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Neck Force</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maximum</td>
<td>565.99</td>
<td>0.084</td>
<td>242.34</td>
<td>0.090</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Minimum</td>
<td>−258.33</td>
<td>0.106</td>
<td>−133.35</td>
<td>0.057</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>Lumbar Spine Force</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maximum</td>
<td>2,570.05</td>
<td>0.081</td>
<td>249.26</td>
<td>0.151</td>
<td>3,457.45</td>
<td>0.092</td>
</tr>
<tr>
<td>Minimum</td>
<td>−2,600.40</td>
<td>0.088</td>
<td>−1,084.86</td>
<td>0.092</td>
<td>−2,292.53</td>
<td>0.081</td>
</tr>
<tr>
<td><strong>Femur Force</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maximum (Left)</td>
<td>996.74</td>
<td>0.083</td>
<td>212.96</td>
<td>0.095</td>
<td>3,666.27</td>
<td>0.076</td>
</tr>
<tr>
<td>Minimum (Left)</td>
<td>−1,236.78</td>
<td>0.077</td>
<td>−354.92</td>
<td>0.099</td>
<td>−355.20</td>
<td>0.200</td>
</tr>
<tr>
<td>Maximum (Right)</td>
<td>1,188.87</td>
<td>0.083</td>
<td>N/A</td>
<td>N/A</td>
<td>4,563.51</td>
<td>0.077</td>
</tr>
<tr>
<td>Minimum (Right)</td>
<td>−968.35</td>
<td>0.076</td>
<td>N/A</td>
<td>N/A</td>
<td>−724.41</td>
<td>0.156</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>( M_x (Nm) )</th>
<th>Time (s)**️</th>
<th>( M_y (Nm) )</th>
<th>Time (s)**️</th>
<th>( M_z (Nm) )</th>
<th>Time (s)**️</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Neck Moment</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maximum</td>
<td>8.89</td>
<td>0.084</td>
<td>69.00</td>
<td>0.106</td>
<td>5.85</td>
<td>0.091</td>
</tr>
<tr>
<td>Minimum</td>
<td>−30.73</td>
<td>0.105</td>
<td>−42.26</td>
<td>0.085</td>
<td>−12.71</td>
<td>0.161</td>
</tr>
<tr>
<td><strong>Lumbar Spine Moment</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maximum</td>
<td>N/A</td>
<td>N/A</td>
<td>89.80</td>
<td>0.072</td>
<td>76.47</td>
<td>0.075</td>
</tr>
<tr>
<td>Minimum</td>
<td>N/A</td>
<td>N/A</td>
<td>−124.42</td>
<td>0.081</td>
<td>−46.03</td>
<td>0.086</td>
</tr>
<tr>
<td><strong>Femur Moment</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maximum (Left)</td>
<td>123.11</td>
<td>0.098</td>
<td>42.93</td>
<td>0.164</td>
<td>44.05</td>
<td>0.153</td>
</tr>
<tr>
<td>Minimum (Left)</td>
<td>−168.83</td>
<td>0.077</td>
<td>−422.63</td>
<td>0.077</td>
<td>−4.70</td>
<td>0.078</td>
</tr>
<tr>
<td>Maximum (Right)</td>
<td>99.91</td>
<td>0.123</td>
<td>18.34</td>
<td>0.075</td>
<td>0.31</td>
<td>0.122</td>
</tr>
<tr>
<td>Minimum (Right)</td>
<td>−203.07</td>
<td>0.080</td>
<td>−15.52</td>
<td>0.082</td>
<td>−11.27</td>
<td>0.179</td>
</tr>
</tbody>
</table>

*Data filtered at CFC 1,000. **Data filtered at CFC 600. ***Time after impact.

### TABLE IV. Assessed Injury Criteria for Hybrid III 50th Percentile Male ATD.

<table>
<thead>
<tr>
<th>Recommended Value</th>
<th>TRACT 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Head Injury Criteria (HIC 15 ms)</td>
<td>700(^a), 51.68</td>
</tr>
<tr>
<td>Head Injury Criteria (HIC 36 ms)</td>
<td>1,000(^a), 71.73</td>
</tr>
<tr>
<td>Resultant Head Acceleration (G)</td>
<td>180(^b), 39.77</td>
</tr>
<tr>
<td>Neck Injury Criteria ( N_y )</td>
<td>1(^a) N/A(^b)</td>
</tr>
<tr>
<td>Neck Tension ( N )</td>
<td>4,170(^b), N/A(^a)</td>
</tr>
<tr>
<td>Neck Compression ( N )</td>
<td>4,000(^b) N/A(^a)</td>
</tr>
<tr>
<td>Neck Shear ( N )</td>
<td>3,100(^b), 573.43</td>
</tr>
<tr>
<td>Neck Flexion (Nm)</td>
<td>190(^b), 42.26</td>
</tr>
<tr>
<td>Neck Extension (Nm)</td>
<td>57(^a), 69.00</td>
</tr>
<tr>
<td>Chest Resultant Acceleration</td>
<td>60(^a), 573.43</td>
</tr>
<tr>
<td>Axial Lumbar Spine Load (N)</td>
<td>6,672(^a), 3,457.45</td>
</tr>
<tr>
<td>Femur Compression (N)</td>
<td>10,008(^a), ( L = 3,666.27 R = 4,563.31 )</td>
</tr>
</tbody>
</table>

\(^a\) Unable to calculate. **Lost data channel. *49 CFR §571.208. \(^b\) Mertz, 2002. \(^c\) Chandler, 1985.
Patient Litter System Response in a Full-Scale CH-46 Crash Test

during the crash. This may be due to the middle and upper ATDs landing on top of the bottom ATD, preventing it from further displacement.

The results illustrate that effective litter occupant restraint is critical in minimizing patient movement during a crash sequence. The difference in orientation of the ATDs (top feet forward, middle/bottom head forward) might have had an effect on litter system response; however, this cannot be determined from the current test and needs additional study.

The instrumented ATD (middle litter position) experienced both inertial and contact accelerations during the test. Low peak head accelerations and calculated HIC values are suggestive of a low risk of head injury. Recorded neck extension moment (69.0 Nm at T + 0.106 seconds) exceeded the federal motor vehicle occupant protection threshold of 57 Nm, indicating a failure to pass the criteria. Loss of the neck axial load channel prevented a complete risk assessment of neck injury. Measured chest acceleration did not exceed injury threshold of 60 G. Lumbar spine forces greater than 2,000 N were recorded in all directions except laterally; however, these forces did not exceed lumbar spine axial load limits (6,672 N) used by the Federal Aviation Administration.

High pelvis accelerations were recorded in the ATD’s posterior to anterior direction (198.90 G at T + 0.075 seconds); however, there are no known pelvis injury criteria. Timing of the peak pelvis acceleration suggests that interaction with the upper and lower ATDs contributed to the high measurements. This finding is consistent with Schamadan et al who measured vertical pelvis accelerations of 125 G in the upper ATD and 150 G in the lower ATD during upper and lower ATD contact. Peak pelvis accelerations for the current test are noticeably higher than the reported vertical pelvis accelerations by Schamadan et al and Weinberg. Weinberg reported peak vertical pelvis acceleration of 135 G for the upper ATD.

Force and bending moment measurements were consistent between the left and right femurs except in the case of bending around the y axis, which was higher in the left femur. The higher bending moment measured in the left femur (My ~422.63 Nm at T + 0.077 seconds) is of interest, though currently there are no known injury criteria for femur bending with which to compare these results. The timing of the high femur loading (right femur Fx 4563.31 N at 0.077 s; left femur Fx 3666.27 N at 0.076 seconds) suggests the load resulted from interaction with the lower and upper ATDs. It is likely that had the litter support structure not failed and the occupants been properly restrained to their respective litters, the high femur loading condition could have been avoided.

Even though the subject crash test was moderately severe, injury thresholds were exceeded for only one anatomical region—the neck. While even a single injury can be severe, the accuracy of these predictions is unknown since Hybrid III injury criteria may not be applicable to a supine litter patient configuration. The ATD injury criteria assessed in the current study were developed for seated occupants in car crash environments, and are neither representative nor validated for loading conditions in the supine position. Additionally, in real world aeromedical transport crash scenarios, the pre-existing medical condition (illness/injury) of the patients could affect the risk of injury. This may require more conservative ATD injury criteria depending on the nature and severity of the patient’s condition. Current doctrine states that the most injured patient is loaded on the bottom; however, from this work no further speculation can be made about the potential for additional injury that could occur to patients based on litter location due to limitations in the current study.

There were several limitations to the study that warrant discussion. The test articles for this study were a CH-46 airframe and a strap and pole litter system. The CH-46 is no longer in Military service and has a different litter support structure configuration than the Army HH-60 and UH-60 medevac platforms. Additionally the strap and pole litter system tested is not representative of the current configuration in medevac platforms. While the test articles differ from what is currently used for medevac transport, this unique test opportunity provides insight to other systems currently in use. The CH-46, while smaller and lighter, has similar configurations and mission to the Army CH-47 which can be configured with a multi-tiered strap and pole litter system for patient transportation if needed. Similar strap and pole litter systems are currently available for use in other platforms such as the Navy MV-22 and Air Force C-130 platforms. While not an evaluation of the system currently used by Army medevac, the results of this study illustrate the performance inadequacy of a statically qualified litter and litter support system in a dynamic environment.

The next limitation is the removal of high-mass components from the overhead structure of the airframe (i.e., engines, transmission, and rotors) without installing comparable masses. This configuration likely had a significant influence on the performance of the airframe sidewalls and the roof structure. Additionally, removal of the landing gear system and rear door altered the crash loading. Although these steps were taken to fulfill NASA test objectives, the impact loads observed and transmitted to the litter system in this test would vary from that expected in an operation crash event. For example, addition of the overhead engine and transmission mass could have caused more substantial roof deformation into the occupant spaces and potentially caused deflection or buckling of the sidewalls. Either of these two conditions would have altered the loading transmitted into the litter support systems and the observed responses. Medical equipment sets and patient care devices were also not included in this test. Medical equipment may become dislodged and could cause additional injuries in a medevac helicopter crash. Due to the number of available data channels, only one instrumented litter-borne ATD was utilized in this experiment limiting the amount of information that can be learned from the upper and bottom ATDs.
CONCLUSION

This system-level assessment of litter system performance in a dynamic test suggests that the structural integrity of statically tested litters and litter support structures does not equate to occupant protection during dynamic loading events. This finding is in agreement with previous testing of litter systems. Improved patient restraint systems are required to ensure proper retention of patients in dynamic environments. Additional research is needed to establish operationally relevant dynamic performance standards for litter systems and occupant protection criteria. These improvements are needed in order to provide wounded patients with an equivalent level of protection provided to the seated aircraft occupants.

ACKNOWLEDGMENTS

The authors would like to acknowledge the Medical Material Development Activity and the Military Operational Medicine Research Program at Fort Detrick, Maryland for funding support. The authors like to acknowledge the Aberdeen Test Center at Aberdeen Proving Ground Maryland and Naval Air Systems Command at Patuxent River Maryland for providing equipment for this test. Additionally, the authors would like to thank Dr. John Crowley, Mr. Chris Sullivan, Mr. Bobby Bowers, and Mr. Patrick Fonda for their expertise and support. This work was financially supported for the fiscal Year 2013 by Army Research Development Testing and Evaluation funds from U.S. Army Medical Materiel Development Activity (USAMMDA), U.S. Army Medical Research and Materiel Command (USAMRMC); and for the fiscal Year 2014/2015 by Army Research Development Testing and Evaluation funds from U.S. Army Military Operational Medicine Research Program (USAMOMRP), U.S. Army Medical Research and Materiel Command (USAMRMC).

REFERENCES

Modeling of Gender Differences in Thermoregulation

Anthony E. Iyoho, PhD; Laurel J. Ng, PhD; Lisa MacFadden, PhD

ABSTRACT In January 2013, the Department of Defense lifted a ban that had prevented women from holding combat positions in the military. However, innate differences in physical traits and physiology between men and women likely will result in differences in physical performance. Sex differences in thermoregulation is a key area that needs to be examined due to the potential impact on physical performance. Therefore, we expanded our previously developed thermoregulation model (TRM) to include the effects of gender. Women have been found to have a lower sweat output in heat stress and lesser shivering in cold stress than men; therefore, the equations for sweat mass loss rate and shivering heat generation were modified for women accordingly. The updated TRM showed good agreement with female data collected from exercise in cool to hot conditions, cold air exposure, and cold water immersion. Gender differences in sweat evaporation appear minimal except for sufficiently high exercise-heat combinations. Gender differences in core temperature and heat generation during cold stress are significant. The expanded TRM can be used to assess gender-specific thermal response with future application to predicting performance differences and optimizing warfighter effectiveness for a wide range of military relevant tasks.

INTRODUCTION
In January 2013, the Department of Defense lifted a ban that had prevented women from holding combat positions in the military, raising concerns as to whether women are physically capable of performing the tasks required of many combat positions.1 Innate differences in physical traits and physiology between men and women likely will result in differences in physical performance, which must be quantified to fully understand the physical capacity of each gender. It is necessary to understand the collective contribution of these gender differences using quantitative model-based methodologies such that better planning and occupational guidelines for physical demand can be developed. Particularly, gender differences in thermoregulation is a key area that needs to be examined due to the potential impact on physical performance in military and civilian settings.

Gender differences in thermoregulation likely lead to discrepancies in physical performance. It is well known that increased body temperature degrades performance for whole body exercise.2–4 Hyperthermia also leads to an attenuation in voluntary activation during isometric sustained contractions.5,6 The investigation of sex differences in thermoregulation during acute heat stress started in the 1940s but gained steam in the 1960s.7–12 Gender differences also have been observed during cold stress with women experiencing lower core temperatures ($T_{core}$) and lesser shivering heat generation than men for the same degree of cooling.13

A big question remains: do women and men thermoregulate differently due to physiologic gender differences or simply due to differences in physical characteristics like body mass, height, and body fat lead to the discrepancies? Until recently, the general consensus was that gender differences in thermoregulation are explained by differences in physical characteristics and fitness.14 In general, women have less lean mass and muscle strength, lower body weight, higher body fat than men, and are typically shorter. Women also have different levels of sex hormones (fluctuating throughout their menstrual cycle), which can affect thermoregulation.15 Recent studies, however, have challenged the idea that gender differences can be explained solely by differences in physical characteristics.16,17 For example, one study isolated the effect of gender and determined that women likely have a limited sweating capacity in comparison to men.7 Additionally, Glickman–Weiss18 concluded that physical characteristics could not alone explain gender differences observed during cold response.

Various human thermoregulation models (TRMs) have been developed to predict body temperatures for a range of environmental conditions to assess responses ranging from thermal comfort to heat strain casualties. The earliest models did not incorporate physiology but instead relied on linear regression fits of environmental test conditions to measured variables.19 In the 1960s, more models, which incorporated the complexities of human thermoregulation, were developed as computing power increased.20,21 Later, models that took a more physiologic approach with an anatomic representation for the body and utilizing Pennes’ energy balance22 in a tissue volume could provide a real-time assessment of the effects of environment, clothing, and variable activity levels.23–29 Into the 21st century, TRMs are still being developed containing useful modifications to their predecessors.14,30–34 However, we found only 1 TRM in literature that attempted to account for gender differences; but,
this model-treated sex differences as purely physical in nature and was applied to heat stressed conditions only. Therefore, there is a need for a TRM to incorporate physiologic gender differences to more accurately predict thermal response for the entire range of operational scenarios (i.e., cold to hot environment, activity levels, and clothing types).

We previously developed a TRM to predict body temperature, evaporative heat loss, and shivering heat response for a wide range of environmental conditions, clothing types, and exercise levels. However, our TRM was validated against male data only. Therefore, the overall objective of this work was to expand the TRM to model the thermoregulatory response of women to make predictions which in the future can assist in understanding performance limitations for women performing physical tasks in cold to hot environments.

METHODS

TRM Overview

A full description of the TRM is provided in the supplementary materials of Iyoho et al. The TRM consists of passive and active components that govern heat exchange within the human body and the environment. The human anatomy is approximated with 10 body segments, where each segment is composed of a core, muscle, fat, and skin layer. All body segments are treated as cylinders, except for the head which is given a spherical representation. The model considers heat conduction between tissues (radial only), dry heat exchange with the environment (convection and radiation), sweating heat loss, shivering heat generation, respiratory heat exchange, and vasomotor blood flow changes. Blood circulates in a closed loop from the central artery to the 10 regions, undergoing countercurrent heat exchange between the arterioles and venules, and heat exchange in the tissues before collecting in the central vein where it is pumped by the heart to the lungs, and back to the central artery. The TRM equations are simultaneously solved using Matlab with Simulink toolbox (The MathWorks, Inc., Natick, Massachusetts). A schematic of the TRM is shown in Figure 1.

For the physical system, the model allows for the input of individual characteristics including body mass, height, percent body fat, and material properties like clothing resistance and clothing permeability. Environmental inputs like air temperature, water immersion temperature, radiant temperature, relative humidity, and wind speed are specifiable. Additionally, parameters related to activity such as work level and net mechanical efficiency can also be specified.

Gender-Specific Modifications

There is evidence that men have greater sweat output than women due to a higher output per gland, which is likely only apparent for severe exercise-heat stress. For women, sweat output begins to plateau with increasing composite body temperature (which is a weighting of $T_{\text{core}}$ and average skin temperature $[T_{\text{sk}}]$) even when the environment does not...
limit evaporation. In Gagnon and Kenny, a plot of evaporation heat loss rate vs. body temperature is provided for women; where the rate of mass loss due to sweating is acquired by dividing this evaporation heat rate by the latent heat of vaporization. A least-squares fit to the total rate of mass loss due to sweating (\( m_{sw} \) [g/h/m²]) for women to this data were done, assuming an exponential shape with respect to body temperature:

\[
A_{w}^{momen} = \frac{376.5}{A} \left[ 1 - \frac{1}{\exp\left\{ 1.983 \left( T_{art} - T_{art,0} - \Delta T_{art,thresh} \right)^{+ve or 0} \right\} } + \frac{1}{A} \left[ 28.35 \left( T_{sk} - T_{sk,0} \right)^{+ve or 0} \right] \right]
\]  

(1)

where \( \Delta T_{art,thresh} \) represent the body surface area, arterial (average skin), and threshold change in arterial temperatures (\( T_{art} \)), respectively. All temperatures are in °C. In calculating the sweat response, we used arterial and \( T_{sk} \) error signals (i.e., difference between the value and initial value) which are involved in the feedback loop that the body uses to regulate sweating.\textsuperscript{27} The temperature error signals were only allowed to be positive (+ve or 0) to prevent \( m_{sw} \) from dropping below zero, which is not physical. It was assumed that \( T_{art} \) represents the composite body temperature as they are both a weighting of core and skin temperatures. We fitted the equation assuming the coefficient of the sweat response due to \( T_{sk} \) was 85% of the male coefficient as it was found that women matched for body mass had about 85% less sweat production than men.\textsuperscript{16} The value of \( \Delta T_{art,thresh} \) was 0.05°C and the distribution of sweat loss to each compartment was taken from Stolwijk.\textsuperscript{27}

Regarding gender differences during cold stress, women have been observed to cool more rapidly than men in cold water and have a lower shivering response for the same degree of cooling.\textsuperscript{13} Studies with responses to cold air show a bit more controversy with women exhibiting slightly lower shivering response.\textsuperscript{13,36} In all, this suggests that men have greater shivering heat production capacity compared to women.\textsuperscript{13}

The shivering control system receives three error signals: skin heat flux, \( T_{core} \), and \( T_{sk} \) and is derived from the work of Gordon et al.\textsuperscript{27} and Stolwijk.\textsuperscript{27} There is evidence that skin heat flux plays a role in shaping the shivering response, and previous modeling work has included this variable along with \( T_{core} \) and skin temperature error signals.\textsuperscript{37} For men, the shivering response, \( q_{shiv} \) [W], is calculated as

\[
q_{shiv, men} = \frac{40.0}{h_{tot}} \left( q_{sk} - q_{sk,thresh} \right)^{+ve or 0} + 80 \left( T_{core,0} - T_{core} - \Delta T_{sk,thresh} \right)^{+ve or 0} + 13 \left( T_{sk,0} - T_{sk} - \Delta T_{sk,thresh} \right)^{+ve or 0}
\]

(2)

where \( h_{tot} \) [W/°C] is the combined convection and radiation coefficient summed for all 10 segments, \( q_{sk} \) is the total heat rate loss at the skin surface and \( q_{sk,thresh} \) is the threshold skin heat rate loss. The symbols \( T_{core,0} \) (\( T_{sk,0} \)) and \( \Delta T_{core,thresh} \) (\( \Delta T_{sk,thresh} \)) represent the core (average skin), initial core (average skin), and threshold change in core (average skin) temperatures. The values of \( \Delta T_{core,thresh} \), \( q_{sk,thresh} \), and \( \Delta T_{sk,thresh} \) are 0.2°C, 106 W/m², and 3°C, respectively. The error signals were only allowed to be positive (+ve or 0) to prevent \( q_{shiv} \) from dropping below zero, which is not physical.

The shivering equation for women reflects a reduced shivering output in comparison to men. For women, the shivering response, \( q_{shiv} \) [W], is calculated as

\[
q_{shiv, women} = \frac{20.0}{h_{tot}} \left( q_{sk} - q_{sk,thresh} \right)^{+ve or 0} + 40 \left( T_{core,0} - T_{core} - \Delta T_{core,thresh} \right)^{+ve or 0} + 13 \left( T_{sk,0} - T_{sk} - \Delta T_{sk,thresh} \right)^{+ve or 0}
\]

(3)

For women, the values of \( \Delta T_{core,thresh} \), \( q_{sk,thresh} \), and \( \Delta T_{sk,thresh} \) are 0.2°C (same as men), 106 W/m² (same as men), and 6°C for air (greater than men’s), respectively.

The shivering equations for men and women were fitted to best match \( T_{core} \) and metabolic response data (studies referenced in the Results section) during cold stress (air and water) from literature. Furthermore, we constrained the model assuming that the shivering equation for men and women should be the same for both air and water exposures. The shivering response cannot exceed a theoretical shivering maximum, which is calculated according to Eyolfson et al.\textsuperscript{38} For shivering control, the \( T_{core} \) is set equal to the torso \( T_{core} \), which resulted in the better agreement with data than using \( T_{art} \). The heat from shivering is distributed throughout the compartments with most of the heat generated at the torso core according to Stolwijk.\textsuperscript{27}

Increases in \( T_{core} \) of 0.3 to 0.5°C have been consistently shown during the luteal phase in comparison to the follicular phase; and this is due to an increase in thermoregulatory set point in \( T_{core} \).\textsuperscript{15} Previous studies have shown that sweat onset thresholds (i.e., the body temperature at which sweating begins) are greater in the luteal phase in comparison to the follicular phase.\textsuperscript{39–41} Changes in onset thresholds are indicative of an alteration in thermoregulatory setpoints.\textsuperscript{17} The TRM already includes core set point temperatures, and therefore can incorporate the effects of menstrual cycle. Previous studies have shown that sweat sensitivity (i.e., the degree that sweat increases or decreases for changes in body temperature) is similar between menstrual phases.\textsuperscript{39–41} Therefore, the coefficients and structure of the rate of sweat mass loss shown in Equation (1) were unaltered for the different menstrual cycle phases.

There is considerable evidence from literature that blood flow onset threshold (i.e., the body temperature at which blood flow increases above some basal value) and sensitivity (i.e., the degree that blood flow increases or decreases for
TRM Validation for Women

TRM predictions have already been shown to be accurate for male subjects over a wide range of air temperatures, activity levels, and clothing. For the simulations, the intrinsic clothing heat resistance for soaked clothing is assumed to be 7% of the dry intrinsic clothing resistance. For water immersion, a mean convection coefficient of 220 W/m²/K was chosen, which is a reasonable value assuming low effective water velocity. It was assumed that no evaporation occurs for immersed body segments. For cold stress, the drop in $T_{core}$ has been shown to depend significantly on body fat composition. Therefore, the actual percent body fat reported in each cold stress study were used as inputs to the model. The environmental conditions and average physical parameters from each experiment were also used as input to the simulations.

To validate the model for female subjects, a literature survey was first conducted to attain thermoregulatory data for exercise, resting cold air, and resting cold water conditions. In particular, 14 female exercise experiments including air temperatures of 20 to 48°C and exercise levels of 3 to 6.4 METs were used to challenge the TRM. This exercise data were not used to parameterize the TRM and thus provided direct test data for the heat stress response. Female cold air exposures ranging between 5 and 20°C and cold water exposures ranging from 18 to 28°C were compared against the TRM. Due to lack of available data, this same cold stress data were used to parameterize the shivering heat calculation shown in Equation (3).

The root mean square deviation (RMSD) was calculated between the experimental data points and model predictions for $T_{core}$, $T_{sk}$, and metabolic thermogenesis to quantify the agreement between model and data:

$$RMSD[^G]\left[C\right]=\left[\frac{\sum_{i=1}^{n}(X_i-Y_i)^2}{n}\right]^{1/2}$$

where $X$ is the difference between the measured and predicted variable for $n$ data points. It has been suggested that RMSD values below 0.5 and 2.0°C for $T_{core}$ and $T_{sk}$, respectively, are acceptable error thresholds; this criteria was used to assess TRM agreement with experimental data. These RMSD bounds are also close to the maximum average standard deviations found by Haslam and Parsons for a large data set comprised of 80 different combinations of environmental conditions. For metabolic rate (MR) during cold exposure, a standard deviation of 33 W/m² has been observed, so an RMSD below this value was assumed to be acceptable.

Additionally, Bland–Altman plots were also constructed to determine the limits of agreement (LoA) and bias between predicted and observed $T_{core}$ and MR. LoA, calculated as $\pm 1.96 \times SD$ of the differences between predicted and observed, determines the range of error that 95% of predictions should fall within. Furthermore, using the Bland–Altman plots, we determined whether gender-specific sweating and shivering responses were necessary. In particular, we simulated the model using female vs. male sweating and shivering equations for the set of female data and compared these results using Bland–Altman analysis. This determined the prediction accuracy if it is assumed that women have the same sweating or shivering response as men.

RESULTS

Exercise Test Conditions

Thermoregulatory responses predicted by the model for women exercising in air ranging from cool to hot were compared against experimental data. Model predictions showed good agreement with experimental data as evidenced by the RMSD values, which were all within the acceptable bound (Table I). Bland–Altman plots for $T_{core}$ are shown in Figure 2a and Figure 2b using the female and male sweat equations, respectively. The LoA and bias of the $T_{core}$ prediction using the female sweat equation (Equation 1) was 0.36 and 0.017°C, respectively. The LoA and bias using the male sweat equation was 0.39 and 0.024°C; which was similar to the results using the female equation. The LoAs were below the 0.5°C threshold of acceptable RMSDs and were more than adequate for a predictive $T_{core}$ model, and the biases were small. The $T_{core}$ probes used in each study were accurate to ±0.1°C. The predicted evaporative heat loss

<table>
<thead>
<tr>
<th>Reference</th>
<th>$T_{air}$</th>
<th>rh</th>
<th>act</th>
<th>$T_{core}$</th>
<th>$T_{sk}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>51</td>
<td>40</td>
<td>30</td>
<td>4.2</td>
<td>0.19</td>
<td>1.06</td>
</tr>
<tr>
<td>50</td>
<td>30</td>
<td>30</td>
<td>5.7</td>
<td>0.16</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>35</td>
<td>12</td>
<td>5.8</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>35</td>
<td>12</td>
<td>5.1</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td>47</td>
<td>36</td>
<td>64</td>
<td>3.6</td>
<td>0.14</td>
<td>0.68</td>
</tr>
<tr>
<td>55</td>
<td>48</td>
<td>13</td>
<td>4.6</td>
<td>0.26</td>
<td>1.98</td>
</tr>
<tr>
<td>48</td>
<td>48</td>
<td>14</td>
<td>3.1</td>
<td>0.23</td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>37</td>
<td>60</td>
<td>3</td>
<td>0.16</td>
<td>0.59</td>
</tr>
<tr>
<td>49</td>
<td>48</td>
<td>14</td>
<td>3</td>
<td>0.15</td>
<td>1.07</td>
</tr>
<tr>
<td>52</td>
<td>20</td>
<td>45</td>
<td>5.3</td>
<td>0.07</td>
<td>0.98</td>
</tr>
<tr>
<td>52</td>
<td>32</td>
<td>60</td>
<td>5.3</td>
<td>0.33</td>
<td>0.97</td>
</tr>
<tr>
<td>54</td>
<td>22</td>
<td>60</td>
<td>5.8</td>
<td>0.1</td>
<td>1.4</td>
</tr>
<tr>
<td>55</td>
<td>35</td>
<td>70</td>
<td>6.4</td>
<td>0.2</td>
<td></td>
</tr>
<tr>
<td>53</td>
<td>40</td>
<td>35</td>
<td>6.4</td>
<td>0.17</td>
<td></td>
</tr>
</tbody>
</table>

| TABLE I. Exercise Test Conditions and RMSD Values. The Symbols $T_{air}$, rh, and act Denote the Air Temperature, Relative Humidity, and Activity Level, Respectively. RMSD Values are Calculated for Core Temperature ($T_{core}$) and Average Skin Temperature ($T_{sk}$). |
using the female and male equations were similar and compared well with data as shown in Figure 3.

**Cold Stress Test Conditions**

Cold air (5–20°C) and water (18°C–28°C) test conditions were simulated and compared with female experimental results. Input conditions used for the cold air and cold water temperature simulations are shown in Tables II and III, respectively. RMSD values for $T_{core}$ and MR were all acceptable as shown in Tables II and III for cold air and cold water, respectively. Bland–Altman plots of $T_{core}$ for the female dataset using the female and male shivering equations are shown in Figures 4a and b, respectively. Results using the female equation show an LoA and bias of 0.36 and $-0.0028^\circ C$; where the bias was extremely small and the LoA was below the 0.5°C threshold of acceptable RMSDs and is more than adequate for a predictive $T_{core}$ model. Results using the male equation fared much worse featuring an LoA and bias of 0.47 and 0.22°C. The $T_{core}$ probes used in each study were accurate to ±0.1°C. Bland–Altman plots of MR for the female dataset using the female and male shivering equations are shown in Figures 5a and b, respectively. Using the female shivering equation, the LoA and bias were 30 and $-0.063$ W/m², respectively; where the bias was negligible and the LoA was below the 33 W/m² acceptable criteria for RMSD values. However, the metabolic results using the male equation showed an unacceptable LoA (51 W/m²) and a large bias (17 W/m²).

**DISCUSSION**

With the recent allowance of women in combat positions, it is necessary to understand the capabilities of women performing physical tasks in different environments. Gender differences in thermoregulation is an important area that needs to be examined due to its impact on physical performance. Therefore, in this work, we expanded an existing model of thermoregulation to be applicable to women. The TRM, previously developed for male subjects, was augmented to model the thermoregulatory response of women. On average, women have lower body mass, shorter stature, less muscle mass, and more body fat than men. The TRM already takes into account these types of physical characteristics. However, there is strong evidence that the active thermoregulatory control components (i.e., sweating and shivering) vary with gender. Our results indicate that gender differences in sweat evaporation during exercise-heat stress are small; at least for the cases studied in this article. This is in-line with the prevailing view that gender differences in thermoregulation during heat stress can be explained solely by physical characteristics. In regard to sweating, men have been shown to have similar onset thresholds (compared to women in the follicular phase) but significantly greater sweat sensitivity than women. Women have a larger population density of heat-activated sweat glands per area than men, but this is offset by a lower sweat output per gland. Gender differences in the physical properties of glands likely leads to the differences in output. In particular, the higher sweat output of men in comparison to women is partly due to men’s larger sweat gland size. However, even given this evidence of men’s higher sweat output, it is likely that differences in evaporative capacity between men and women are really only seen at sufficiently high workloads in a dry environment.

One possible limitation of our study is that our test cases (Table I) did not feature enough data with a sufficiently high
combination of heat and workload to see gender differences. Our test data do feature a few cases where the environment is extremely hot (>40°C) and dry (<30% rh); which is an environment that should most expose gender differences in heat stress response. Research has shown that women can approach evaporative heat capacities, even in a dry environment where sweat evaporation is not limited, whereas men do not. This characteristic of the female sweat response is reflected by Equation (1), which shows an eventual plateau in sweat output with increased body temperature. However, when we compared $T_{core}$ predictions using the female sweat equation vs. the male equation, there were negligible differences (plot not shown) as evidenced by similarities in LoA and bias (Fig. 2). In particular, for the “Wells1974” case (48°C and 13% rh) in Figure 3, evaporative heat loss using the male equation is visibly larger than the female equation; however, by the end of the simulation, $T_{core}$ resulting from the female equation is only about 0.25°C higher than that of the male equation (plot not shown). Although, it is possible that differences in $T_{core}$ would have become significantly large if the exercise duration lasted longer than the prescribed 40 minutes. For instance, Gagnon and Kenney recently showed that women matched for body mass with men had higher rectal temperatures (~0.46°C higher on average) reaching statistical significance after 90 minutes of cycling in a hot-dry environment. In all, it appears that men do have a greater capacity for sweat heat losses; however, it is only seen for a sufficiently high combination of heat and workload.

Gender differences in cold stress response are much less studied than heat stress; however, it appears that women have a lower heat generation for the same degree of cooling than men which is more pronounced in cold water in comparison to cold air. The lower shivering output in cold air and cold water seen in women is reflected by the parameters of Equations (2) and (3). The parameters of these equations show a shivering response with reduced $T_{core}$ sensitivity (i.e., reduced $T_{core}$ coefficient) and reduced heat

**TABLE II.** Cold Air Exposure Test Conditions and RMSD Values. The Symbol $T_{air}$ Denotes the Air Temperature. RMSD Values are Calculated for Core Temperature ($T_{core}$), Average Skin Temperature ($T_{sk}$), and Metabolic Rate (MR)

<table>
<thead>
<tr>
<th>Reference</th>
<th>$T_{air}$ (°C)</th>
<th>Fat (%)</th>
<th>$T_{core}$ (°C)</th>
<th>$T_{sk}$ (°C)</th>
<th>MR (W/m²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>56</td>
<td>5</td>
<td>23.2</td>
<td>0.15</td>
<td>1.36</td>
<td>27.3</td>
</tr>
<tr>
<td>56</td>
<td>10</td>
<td>22.8</td>
<td>0.26</td>
<td>0.45</td>
<td>7.8</td>
</tr>
<tr>
<td>56</td>
<td>15</td>
<td>22.8</td>
<td>0.21</td>
<td>0.28</td>
<td>6</td>
</tr>
<tr>
<td>56</td>
<td>20</td>
<td>22.8</td>
<td>0.04</td>
<td>0.51</td>
<td>1.4</td>
</tr>
</tbody>
</table>

**TABLE III.** Cold Water Exposure Test Conditions and RMSD Values. The Symbol $T_{water}$ Denotes the Water Temperature. RMSD Values Are Calculated for Core Temperature ($T_{core}$) and Metabolic Rate (MR)

<table>
<thead>
<tr>
<th>Reference</th>
<th>$T_{water}$ (°C)</th>
<th>Fat (%)</th>
<th>$T_{core}$ (°C)</th>
<th>MR (W/m²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>58</td>
<td>20</td>
<td>25.2</td>
<td>0.2</td>
<td>13.4</td>
</tr>
<tr>
<td>58</td>
<td>24</td>
<td>25.2</td>
<td>0.2</td>
<td>10.4</td>
</tr>
<tr>
<td>58</td>
<td>28</td>
<td>25.2</td>
<td>0.11</td>
<td>16.3</td>
</tr>
<tr>
<td>58</td>
<td>20</td>
<td>18.5</td>
<td>0.23</td>
<td>12</td>
</tr>
<tr>
<td>58</td>
<td>24</td>
<td>18.5</td>
<td>0.07</td>
<td>13.9</td>
</tr>
<tr>
<td>58</td>
<td>28</td>
<td>18.5</td>
<td>0.08</td>
<td>9.6</td>
</tr>
<tr>
<td>59</td>
<td>20</td>
<td>21.9</td>
<td>0.2</td>
<td>12.7</td>
</tr>
<tr>
<td>59</td>
<td>28</td>
<td>21.9</td>
<td>0.23</td>
<td>8.7</td>
</tr>
<tr>
<td>60</td>
<td>18</td>
<td>22.4</td>
<td>0.11</td>
<td>16.4</td>
</tr>
<tr>
<td>57</td>
<td>20</td>
<td>22.2</td>
<td>0.26</td>
<td>20.3</td>
</tr>
<tr>
<td>57</td>
<td>20</td>
<td>34</td>
<td>0.13</td>
<td>23.7</td>
</tr>
</tbody>
</table>

**FIGURE 3.** Evaporative heat loss (Evap) response using female sweat equation (solid lines) and male sweat equation (dashed lines) vs. experimental data (circles) for exercise test conditions.
flux sensitivity. Bland-Altman plots of $T_{\text{core}}$ and MR for the cold stressed conditions using female and male shivering equations (Figure 4a vs Figure 4b and Figure 5a vs. Figure 5b) show a stark reduction in prediction accuracy if it is assumed that women have the same shivering response as men. Therefore, gender-specific shivering equations should be used when predicting cold stress response for men and women.

To our knowledge, our TRM is the first to incorporate physiologic sex differences in heat and cold stress responses; albeit the gender differences in heat stress response appear only to be a factor for sufficiently high combinations of ambient temperature and workload. A previous TRM individualized its parameters to account for varying characteristics but accounted for gender through physical characteristics only, which appears to be a fair assumption when the exercise-heat stress is not sufficiently high. However, our results and literature would suggest that there are significant differences between men and women in regards to the cold stressed response. Furthermore, Glickman-Weiss concluded that anthropomorphic characteristics could not alone explain gender differences observed during cold response. Therefore,
the TRM is in a position to predict physiologic sex differences in thermoregulatory response especially in cold conditions.

CONCLUSION

The goal of this study was to expand a previously developed TRM to include the effects of gender with future applications for prediction of performance differences including thermal response in extreme environmental conditions and optimization of warfighter effectiveness for military relevant tasks. Our analysis showed that using male control equations to predict female response produced inaccurate predictions of women’s thermoregulatory response particularly in cold conditions. In regards to exercise in the heat, many military activities will result in increased body temperatures and sweating. Evidence from literature would suggest that women performing work at the same absolute level as men will experience greater rises in $T_{core}$ due to their lower body mass, higher fat content, and lower sweat output. Women matched for physical characteristics and working at the same absolute level as men will likely have greater rises in $T_{core}$ when the exercise-heat stress is sufficiently high; this is attributed to women in general exhibiting a somewhat lower evaporative output especially in a hot-dry environment. However, women may have a lower risk of dehydration due to this lower sweat rate. During cold exposure, women will likely experience greater hypothermia than men, even though they have greater fat content, due to a lower shivering capacity. However, there is still uncertainty in where these sex differences would limit or benefit women’s ability to accomplish certain tasks, as the specific demands and environment must be taken into account to better assess the physical and mental outcomes of men and women.

ACKNOWLEDGMENTS

This work was funded through the U.S. Army Medical Research and Materiel Command under contract W81XWH-11-D-0011.

REFERENCES

Modeling of Gender Differences in Thermoregulation


MILITARY MEDICINE, Vol. 182, March/April Supplement 2017
Proposing Using Waist-to-Height Ratio as the Initial Metric for Body Fat Assessment Standards in the U.S. Army

COL Stephen A. Berstein, MC USA (Ret.); Michael Lo, MSPH†; W. Sumner Davis, PhD‡

ABSTRACT  Objectives: Soldiers failing to meet Army Regulation 600-9 (AR 600-9) for weight-to-height standards are required to undergo body fat taping assessment. This article describes a clinical performance improvement project in which battalion medical staff identified a simpler way to improve on meeting AR 600-9 standards with proposing to use the waist-to-height ratio (WtHR) in place of the current methodology. Methods: During a yearlong combat deployment to Iraq from 2009 to 2010, 42 Soldiers (34 males, 8 females) were evaluated and monitored by battalion medical staff for weight and body fat loss. Mean body mass index and waist circumference were compared between baseline or initial assessment and final assessment. The percentage of Soldiers meeting body fat standards was compared among those who had attained a WtHR ≤ 55% versus those who had not. Results: By the final assessment, mean body mass index had decreased 2.21 kg/m² or 6.6% (p = 0.002) and mean waist circumference had decreased 6.0 cm or 5.8% (p = 0.008). All Soldiers who had attained a WtHR ≤ 55% met AR 600-9 body fat taping assessment standards and presented a professional military appearance. Conclusion: Attaining a WtHR ≤ 55% is an effective measure of body composition that was able to predict a Soldiers’ achievement of body fat standards metric for meeting AR 600-9 standards and achieving a professional military appearance.

INTRODUCTION
As an all-volunteer force, the U.S. military is a reflection of society, in which the combined prevalence of overweight and obesity, currently estimated at 69% among adults age 20 years and older, is a national public health concern. An estimated 23 to 27% do not qualify for enlistment due to excessive weight. Current societal issues are affecting recruitment, as evidenced by the 7,000 Army enlisted personnel and officers responding to the 2011 Department of Defense (DoD) Health Related Behaviors Survey of Active Duty Military Personnel. Over one-half (51.9%) qualified as overweight, defined as a body mass index (BMI) of 25 kg/m² or greater; furthermore, an additional 15.8% of Army respondents were found to have been obese, defined as a BMI of 30 kg/m² or greater. To qualify for service in the Army, 40.5% of male respondents and 22.3% of female respondents were required to lose 20 pounds or more.

Obese and overweight Soldiers, along with eligible beneficiaries, impact military readiness, overall health, and total health-related costs. The health impacts of obesity are widely known; what is not appreciated is the annual cost of $60 million for replacing the discharged first-term enlistees due to being overweight. While this pales in comparison to the cost of overweight-related morbidities to the DoD Military Health System, estimated to be over $1 billion annually, the loss due to missed and lower productivity are incalculable. This is more than the costs of treating alcohol and tobacco-related issues combined. Excessive body weight and obesity are more than just issues with body fat and composition; they also transcend into the physical rigors and injuries associated with military service.

Military bearing and appearance, including health, are tied to professional advancement, and hence, the reason for the height and weight standards. Per Army Regulation 600-9 (AR 600-9; Army Body Composition Program [ABCP] and formerly titled the Army Weight Control Program), unit commanders and supervisors are tasked with enforcing military bearing and fitness. The primary objective is “to ensure all Soldiers achieve and maintain optimal well-being and performance under all conditions.” This includes operational readiness, physical fitness, good health, and professional military appearance. Compliance is assessed semi-annually, in conjunction with the Army Physical Fitness Test (APFT).

To determine if the service member (SM) meets the Army standard for body composition, the assessment process begins with using a screening allowable weight for height table, organized by sex and age. If the SM meets the standard, the SM passes; if the SM fails to meet the standard, then a body fat assessment is required using taped anthropometric measurements described below with a subsequent formula calculation (see Fig. 1 [DA Form 5500] and Fig. 2 [DA Form 5501]). Men have waist and neck circumference measurements taken; women have the waist, neck, and hips measured. Allowable calculated body fat percentage standards vary with age and gender. If the SM’s...
calculated percentage body fat fails to meet the standard, the SM is referred for medical assessment to determine if there is a medical cause for being out of standards. If none exists, the SM is enrolled in the unit ABCP, and a Report to Suspend Favorable Personnel Actions (DA Form 268), is issued. There are monthly ABCP assessments to document progress and health and fitness knowledge. Failure to progress within 3 months or re-entry into ABCP within 12 months after release from ABCP is the basis for administrative separation.6

With deployment and recurrent deployment cycles, the year of training and preparation often takes precedence and priority; as a result, unit leaders place less emphasis on such secondary issues as weight control. Longer work hours and multiple field training exercises also disrupt normal diet and exercise regimens for Soldiers. Manning and staffing for

FIGURE 1. DA form 5500, body fat content worksheet (male).
deployment into a combat zone takes precedence over the issues of being out of weight standards before the end of enlistment (known as trying to eat one’s way out of deployment or out of one’s enlistment contract). Combine all of these factors with a group of formerly overweight new and junior Soldiers, who had to lose weight to meet entry and retention standards, and it is no surprise to see these Soldiers gain weight and failing standards again.

This was the background for one deploying Army battalion in 2009. The Battalion Commander recognized this cyclical problem and decided to approach the issue differently by utilizing his battalion medical section differently while deployed. His order was not just to support the unit leaders and evaluate the overweight Soldiers, but rather also to teach, assist, track, and lead them in their efforts to reduce weight and improve their overall health and physical fitness.

The Battalion Commander’s intent was to take an administrative burden off the Company Commanders and place as much as possible on the medical section so the companies might focus on the combat mission; this unique approach worked for this deployed battalion.

What came out of this work was the key to success was the Soldiers attaining a healthy waist circumference that allowed them to pass the Army standards. This observation—reducing the waist circumference being key to passing the calculated body fat percentage standard—is what led to considering and proposing the concept that instead of doing the current process of assessing body fat percentage as the criteria for enrollment into ABCP, that perhaps the Army might look at just the waist circumference and the waist-to-height ratio (WtHR). WtHR is simply dividing the waist circumference at the navel by the height—a unit-less
measure, much simpler to collect, and much quicker to calculate than the current process.

**METHODS**

The data used to support this proposition came from a prospective, observational performance improvement project. This used a standard medical approach for weight loss with the battalion’s overweight junior Soldiers across demographics while deployed to Iraq in 2009 to 2010. The chain of command, the brigade surgeon, the division preventive medicine officer, and the division surgeon approved this plan. The main focus of this project was in line with the Battalion Commander’s intent—insulate and assist line commanders in tracking and counseling Soldiers by performing the bulk of the work in the battalion aid station. With this being execution of normal military duties required per Army regulations and a performance improvement project, the need for institutional review board review and approval for this project was deemed exempt.

The project followed 42 overweight Soldiers identified for enrollment in the ABCP just before deployment. Twenty-six of these 42, also failed their last APFT. Once in the combat theater, each was seen in clinic receiving a comprehensive medical history and physical along with dietary and exercise counseling, care for medical issues impeding progress and continuity of care throughout the deployment—all in accordance with AR 600-9. Services were limited for outside referrals or consultations. Clinical laboratory abnormalities were identified in this group, which highlighted adverse metabolic and cardiovascular effects that had not been previously described in the literature regarding younger (age < 35) overweight adults to include hyperglycemia, metabolic syndrome, and premetabolic syndrome. One Soldier was administratively separated for nonweight/fi

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Male (n = 34)</th>
<th>Female (n = 8)</th>
<th>Total (n = 42)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, Years, Mean</td>
<td>26.0 ± 5.1</td>
<td>25.3 ± 3.3</td>
<td>25.9 ± 4.8</td>
</tr>
<tr>
<td>20–25</td>
<td>19 (55.9%)</td>
<td>5 (62.5%)</td>
<td>24 (57.1%)</td>
</tr>
<tr>
<td>26–30</td>
<td>7 (20.6%)</td>
<td>2 (25.0%)</td>
<td>9 (21.4%)</td>
</tr>
<tr>
<td>31–35</td>
<td>6 (17.6%)</td>
<td>1 (12.5%)</td>
<td>7 (16.7%)</td>
</tr>
<tr>
<td>36–40</td>
<td>2 (5.9%)</td>
<td>0 (0.0%)</td>
<td>2 (4.8%)</td>
</tr>
</tbody>
</table>

Pay Grade

<table>
<thead>
<tr>
<th>Pay Grade</th>
<th>Male (n = 34)</th>
<th>Female (n = 8)</th>
<th>Total (n = 42)</th>
</tr>
</thead>
<tbody>
<tr>
<td>E2</td>
<td>3 (8.8%)</td>
<td>0 (0.0%)</td>
<td>3 (7.1%)</td>
</tr>
<tr>
<td>E3</td>
<td>8 (23.5%)</td>
<td>2 (25.0%)</td>
<td>10 (23.8%)</td>
</tr>
<tr>
<td>E4</td>
<td>16 (47.1%)</td>
<td>5 (62.5%)</td>
<td>21 (50.0%)</td>
</tr>
<tr>
<td>E5</td>
<td>7 (20.6%)</td>
<td>1 (12.5%)</td>
<td>8 (19.1%)</td>
</tr>
</tbody>
</table>

BMI, kg/m², Mean

<table>
<thead>
<tr>
<th>BMI (≤25)</th>
<th>Male (n = 34)</th>
<th>Female (n = 8)</th>
<th>Total (n = 42)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0 (0.0%)</td>
<td>0 (0.0%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Overweight (25.1–30.0)</td>
<td>4 (11.8%)</td>
<td>4 (50.0%)</td>
<td>8 (19.0%)</td>
</tr>
<tr>
<td>Mildly Obese (30.1–35.0)</td>
<td>17 (50.0%)</td>
<td>4 (50.0%)</td>
<td>21 (50.0%)</td>
</tr>
<tr>
<td>Moderately Obese (35.1–40.0)</td>
<td>11 (32.3%)</td>
<td>0 (0.0%)</td>
<td>11 (26.2%)</td>
</tr>
<tr>
<td>Morbidly Obese (&gt;40.0)</td>
<td>2 (5.9%)</td>
<td>0 (0.0%)</td>
<td>2 (4.8%)</td>
</tr>
</tbody>
</table>

Army Physical Fitness Test

<table>
<thead>
<tr>
<th>Test</th>
<th>Male (n = 34)</th>
<th>Female (n = 8)</th>
<th>Total (n = 42)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Passed</td>
<td>14 (41.2%)</td>
<td>2 (25.0%)</td>
<td>16 (38.1%)</td>
</tr>
<tr>
<td>Failed</td>
<td>20 (58.8%)</td>
<td>6 (75.0%)</td>
<td>26 (61.9%)</td>
</tr>
</tbody>
</table>

BMI, body mass index; SD, standard deviation.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Mean</th>
<th>SD</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>BMI, kg/m²</td>
<td>33.7</td>
<td>2.9</td>
<td>27.9, 40.4</td>
</tr>
<tr>
<td>Waist Circumference, cm</td>
<td>106.0</td>
<td>7.3</td>
<td>93.5, 126</td>
</tr>
<tr>
<td>Glucose, mg/dL</td>
<td>105.4</td>
<td>7.6</td>
<td>92, 119</td>
</tr>
<tr>
<td>Total Cholesterol, mg/dL</td>
<td>157.5</td>
<td>22.6</td>
<td>97, 203</td>
</tr>
<tr>
<td>High-Density Lipoprotein, mg/dL</td>
<td>37.0</td>
<td>6.7</td>
<td>24, 53</td>
</tr>
<tr>
<td>Triglycerides, mg/dL</td>
<td>104.9</td>
<td>46.9</td>
<td>45, 234</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Mean</th>
<th>SD</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>BMI, kg/m²</td>
<td>30.5</td>
<td>3.1</td>
<td>27.4, 34.5</td>
</tr>
<tr>
<td>Waist Circumference, cm</td>
<td>96.3</td>
<td>5.5</td>
<td>88, 106</td>
</tr>
<tr>
<td>Glucose, mg/dL</td>
<td>102.2</td>
<td>11.1</td>
<td>90, 121</td>
</tr>
<tr>
<td>Total Cholesterol, mg/dL</td>
<td>188.6</td>
<td>52.3</td>
<td>131, 299</td>
</tr>
<tr>
<td>High-Density Lipoprotein, mg/dL</td>
<td>52.5</td>
<td>8.8</td>
<td>44, 71</td>
</tr>
<tr>
<td>Triglycerides, mg/dL</td>
<td>115.0</td>
<td>33.7</td>
<td>79, 173</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Mean</th>
<th>SD</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>BMI, kg/m²</td>
<td>33.1</td>
<td>3.2</td>
<td>27.4, 40.4</td>
</tr>
<tr>
<td>Waist Circumference, cm</td>
<td>104.1</td>
<td>7.9</td>
<td>88, 126</td>
</tr>
<tr>
<td>Glucose, mg/dL</td>
<td>104.8</td>
<td>8.3</td>
<td>90, 121</td>
</tr>
<tr>
<td>Total Cholesterol, mg/dL</td>
<td>163.6</td>
<td>32.3</td>
<td>97, 299</td>
</tr>
<tr>
<td>High-Density Lipoprotein, mg/dL</td>
<td>40.0</td>
<td>9.4</td>
<td>24, 71</td>
</tr>
<tr>
<td>Triglycerides, mg/dL</td>
<td>106.9</td>
<td>44.4</td>
<td>45, 234</td>
</tr>
</tbody>
</table>

BMI, body mass index; SD, standard deviation.
RESULTS

Overall, 27 of the 41 (65.8%) Soldiers succeeded in lifestyle changes to meet body fat standards by the end of the 1-year deployment period. Further, 20 of the 26 (77%) of those Soldiers who had failed previous APFT succeeded in passing the test. The female group demonstrated an average loss of 17.00 lb (8.8%), with a 2.77 reduction in BMI and 9.50-cm (9.9%) reduction in waist circumference. The male group showed an average loss of 14.21 lb (6.6%) with a 2.07 reduction in BMI and 5.18-cm (5.0%) reduction in waist circumference. The male group of 41 (58%) achieved this benchmark with all passing the Army body fat taping standard. Only 6 of 17 (35%) that did not attain a waist circumference of ≤ 55% of height passed the Army body fat taping standard (p < 0.00072) (Table III).

DISCUSSION

The weight control program is the responsibility of the Commander, rather than a medical program. Commanders may order a body fat assessment at any time for Soldiers not presenting a military appearance; the norm is that all Soldiers are screened every 6 months in conjunction with the APFT—hence, weight and physical fitness testing are tied together. The overweight Soldier works with the unit master fitness trainer for exercise guidance and monitoring. The overweight Soldier is referred to the medical unit for health evaluation and to investigate for the rare medical causes for obesity. If a medical cause of the obesity is found, the condition is treated and controlled to the point of being a nonfactor or is grounds for medical separation evaluation. Nutrition services, where available, are consulted and provide dietary guidance—if not available, the health care provider does this. Thus, the program designed and executed in this deployment differed from the norm, as the medical unit was primarily responsible for direction and monitoring. Beyond providing medical care and support, when the Soldier’s waist circumference shrank to or below 55% of his/her height, she/he was able to meet the formal Army body fat assessment—this gave him/her an easy metric to follow.

Several methods are used for measuring body composition/body fat in the clinic, but the problem lies with method simplicity. Many electronic health records calculate BMI, but it does not account for muscle mass or body build. Alternate modalities for body composition analysis include bioelectrical impedance, calipers to measure skinfold thickness, dual-energy X-ray absorptiometry scan, and hydrostatic weighing; while these examinations are accurate, they are also time-intensive, costly, and used primarily in medical research or nutritional care offices. Physical measurements for waist-to-hip ratio, being the waist circumference divided by the hip circumference, or WtHR, being the waist

<table>
<thead>
<tr>
<th>Weight Category</th>
<th>Male WtHR</th>
<th>Male Examples</th>
<th>Female WtHR</th>
<th>Female Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abnormally Slim</td>
<td>&lt;0.35</td>
<td></td>
<td>&lt;0.35</td>
<td>Barbie Doll (0.25)</td>
</tr>
<tr>
<td>Extremely Slim</td>
<td>0.35–0.43</td>
<td>Ken Doll (0.36), College Swimmers (0.428)</td>
<td>0.35–0.42</td>
<td>College Swimmers (0.424)</td>
</tr>
<tr>
<td>Slender, Healthy</td>
<td>0.43–0.46</td>
<td>Willoughby Ideal (0.458)</td>
<td>0.42–0.46</td>
<td></td>
</tr>
<tr>
<td>Normal, Healthy</td>
<td>0.46–0.53</td>
<td>General Healthy Cutoff (0.50), NHANES BMI = 25* (0.51)</td>
<td>0.46–0.49</td>
<td>WHO Increased Risk (0.492)</td>
</tr>
<tr>
<td>Overweight</td>
<td>0.53–0.58</td>
<td>WHO Increased Risk (0.536), NHANES BMI = 30** (0.57)</td>
<td>0.49–0.54</td>
<td>WHO Substantially Increased Risk (0.582)</td>
</tr>
<tr>
<td>Obese</td>
<td>0.58–0.63</td>
<td>WHO Substantially Increased Risk (0.582)</td>
<td>0.54–0.58</td>
<td>WHO Substantially Increased Risk (0.541)</td>
</tr>
<tr>
<td>Highly Obese</td>
<td>&gt;0.63</td>
<td></td>
<td>&gt;0.58</td>
<td></td>
</tr>
</tbody>
</table>

circumference divided by height, are rarely performed in the clinic. Both are more straightforward methods, and the simplest is measuring the waist to the height ratio—the WtHR. Measurement of the waist circumference at the umbilicus is easy, reproducible, and quick, and allows for giving the patient a real tangible target to monitor and follow at home. This facilitates discussion and counseling with the patient on health risk concerns as the waist circumference more accurately reflects central adiposity, which is directly linked to cardiometabolic issues and risk.9,10

Using the WtHR as a counseling tool, allows a quick, simple, easy-to-understand metric and goal, which patients can work toward to improve their health. The WtHR allowed an improved comprehension of the true meaning of the BMI for clinical use, such as in the case of an individual with an athletic, muscular build who has a high BMI, but also has a normal WtHR. WtHR further defines metabolic syndrome criteria for the waist circumference, adding context to the level of central adiposity.9,10 The WtHR is neutral for gender, ethnicity, and age, something vital for use with the military. A clinical guideline table highlights the subtle differences for practical clinic use (Table IV).11 Furthermore; WtHR actualizes proper military appearance and bearing at all times.6

Table IV also provides a review of various WtHR norms. Fifty percent (or 0.50) is the general healthy normative value.12,13 Waist circumference criteria for metabolic syndrome is set at the arbitrary level of 40 inches or higher for males and 35 inches or higher for females, with no consideration of the person’s height. Fifty-five percent (or 0.55) was chosen for analysis with simple mathematics applied—waist equal to half of the height (the general norm) plus 10%. This value represents and covers those at a level of increased risk in both men and women. Further study and analysis ought to be done to further refine this value, but the 0.55 benchmark serves as an easy starting point in this project for two reasons. First is that in the analysis of the sample project, all who attained a level of 0.55 or less, passed the AR 600-9 body fat calculation standards. Second is that this value allows for simple mathematics whether in clinic from a medical perspective or from a unit commander perspective—if the SM is 70 inches tall, then the waist needs to be less than or equal to 38.5 inches (35 + 3.5). Likewise, a 66-inch tall SM needs to be less than or equal to 36.3 inches (33 + 3.3).

In conclusion, all 42 SMs in this project began with a waist circumference that was > 55% of their height. Measuring the waist was simple and reproducible, giving Soldiers a metric that they could use to establish goals. All who attained a WtHR ≤ 55% of height passed Army body fat standards. Furthermore, these Soldiers had a proper military appearance, felt better, and had improved lab indices. Although noting the limitation of a small sample size and noting this group may not be representative of the general Army population, this project supports the need for further research into using WtHR as an initial metric for Soldiers. Based on observation, the WtHR is a valid test for this sample. This metric may be a more useful and better clinical metric for use within the Army and DoD in trying to curb the issues with weight and body fat as well as the rising costs of treating obesity-related conditions. All Soldiers in this project met the body fat standard, which supports the use of WtHR as an alternative to the existing process. The WtHR for Soldiers made personal compliance more feasible and properly enforceable for this sample.

REFERENCES

An Augmented Reality-Based Approach for Surgical Telementoring in Austere Environments

Dan Andersen, MS†; Voicu Popescu, PhD†; Maria Eugenia Cabrera, MS‡; Aditya Shanghavi, MS‡; CDR Brian Mullis, MC USN (Ret.)‡; Sherri Marley, BSN, RN‡; Gerardo Gomez, MD‡; Juan P. Wachs, PhD†

ABSTRACT Telementoring can improve treatment of combat trauma injuries by connecting remote experienced surgeons with local less-experienced surgeons in an austere environment. Current surgical telementoring systems force the local surgeon to regularly shift focus away from the operating field to receive expert guidance, which can lead to surgery delays or even errors. The System for Telementoring with Augmented Reality (STAR) integrates expert-created annotations directly into the local surgeon’s field of view. The local surgeon views the operating field by looking at a tablet display suspended between the patient and the surgeon that captures video of the surgical field. The remote surgeon remotely adds graphical annotations to the video. The annotations are sent back and displayed to the local surgeon while being automatically anchored to the operating field elements they describe. A technical evaluation demonstrates that STAR robustly anchors annotations despite tablet repositioning and occlusions. In a user study, participants used either STAR or a conventional telementoring system to precisely mark locations on a surgical simulator under a remote surgeon’s guidance. Participants who used STAR completed the task with fewer focus shifts and with greater accuracy. The STAR reduces the local surgeon’s need to shift attention during surgery, allowing him or her to continuously work while looking “through” the tablet screen.

INTRODUCTION
Treating combat trauma injuries effectively and rapidly is currently a key challenge in the military. Such injuries are often complex and affect multiple organs (polytrauma). For example, trauma resulting from blast or fragmentation injuries in the thorax is the main cause of mortality in 25% of military victims, and a contributing factor in an additional 25% of trauma-related deaths. Such trauma is difficult to treat, requiring appropriate care from specialized surgical experts. Most combat deaths occur before the injured soldier can reach a medical facility, but approximately one in four battlefield fatalities have been deemed potentially survivable; reducing the time interval between the point of injury and a surgical operation can significantly improve outcomes for such patients.

Kotwal et al. found that an institutional mandate to reduce the time between injury and receiving of care improved patient outcomes, which emphasizes the importance of targeting the “golden hour” in treatment of combat injuries. However, evacuation of soldiers with such injuries to a hospital may not always be logistically possible. Even in such cases where transport is possible, the required time of transport (more than 30 minutes) can reduce the likelihood of patient survival. As a result, combat trauma injuries may be restricted to an austere environment, such as a Role 2 Forward Surgical Team, where the necessary expert surgeons are not physically present. Improving the outcomes for these patients requires solving the problem of providing diverse, specialized, and expert care in such austere environments. Providing adequate care during the “golden hour” in future military scenarios (such as prolonged field care and long-distance evacuation) will require an emphasis on surgical measures that can be carried out regardless of patient location.

One promising approach to reducing the number of combat deaths due to complex trauma injuries is to bring the specialized expertise of a surgeon into an austere environment where that surgeon is not physically present, using information technology. For example, surgical telementoring is a method of providing a remote expert surgeon’s expertise to a local, less-experienced surgeon in real time, during a surgical operation itself. In such a method, the remote surgeon receives information about the operating field, and then provides real-time instruction or guidance to the local surgeon on how best to proceed with the surgery. It should be noted that, in the austere environments targeted by such a telementoring system, there is an assumption and requirement of basic medical skills on the part of the local surgeon, which can be augmented by remote expert assistance.

The goal of surgical telementoring is to improve patient outcomes by bringing the expertise and mentoring abilities of a remote surgeon into the operating room. Providing the surgeons with a sense of co-presence, where both the remote

*Department of Computer Science, Purdue University, 305 North University Street, West Lafayette, IN 47907.
†School of Industrial Engineering, Purdue University, 315 North Grant Street, West Lafayette, IN 47907.
‡Indiana University School of Medicine, 340 West 10th Street, Suite 6200, Indianapolis, IN 46202.

Presented at the Military Health System Research Symposium, Fort Lauderdale, FL, August 17–20, 2015.

This work was supported by the Office of the Assistant Secretary of Defense for Health Affairs under Award No. W81XWH-14-1-0042. Opinions, interpretations, conclusions, and recommendations are those of the author and are not necessarily endorsed by the Department of Defense.

doi: 10.7205/MILMED-D-16-00051
and local surgeons are able to interact as if the other were physically co-located, has the potential to enhance the effectiveness of telementoring. As the level of communication and visualization between the surgeons increases, the level of co-presence increases, which moves the telementored experience toward two objectives: (1) that the remote surgeon receives all relevant information as if he or she were present in the operating room and (2) that the local surgeon receives the remote surgeon’s guidance as if he or she were physically present.

While current telementoring systems have demonstrated the usefulness and potential of telementoring, the interaction between surgeons is constrained by the limitations of communication and visualization in these systems. For example, current approaches allow an expert surgeon to draw visual annotations overlaid on an image of the operating field in a process called telestration, but the local surgeon can only view these annotations by looking at a nearby computer monitor. As a result, the local surgeon is forced to repeatedly shift focus between the monitor and the operating field during the course of the procedure. Such focus shifts can be distracting and reduce the sense of co-presence. Further, it requires the local surgeon to mentally transform the annotated view to his or her direct view of the operating field. This could lead to delays or errors in medical treatment.

In this article, the STAR system is described and it is shown how it can improve the sense of co-presence by using augmented reality (AR). Such a system removes the need for distracting impediments such as focus shifts from the surgical telementoring process. The technique of AR overlays relevant virtual information onto imagery of a real-world scene. The STAR overlays the remote surgeon’s textual and graphical annotations onto imagery of the operating field, as with current telementoring systems, but it also overcomes the focus shift limitation by placing this augmented imagery directly into the field of view of the local surgeon. The augmented imagery appears to the local surgeon using a tablet suspended between the operating field and the local surgeon’s head, which allows for continuous viewing of both the operating field and the remote surgeon’s instructions in a form factor that avoids bulky eyewear or headwear. By improving the sense of co-presence in this way, telementoring becomes more of a feasible solution for the problem of combat trauma injuries, allowing experts to provide more assistance closer to the point of injury within the “golden hour” where intervention can improve patient outcomes.

This article first provides a general overview of STAR and how each of its components interacts to provide surgical telementoring without the need for focus shifts. Second, it describes formative evaluation of a prototype version of our system. Third, it describes two experiments that were conducted to evaluate the system from a technical perspective and from a user validation perspective.

MATERIALS AND METHODS

Overview of STAR

STAR is divided into two general components: the remote subsystem and the local subsystem. The remote subsystem’s function is to display real-time imagery of the local surgeon’s operating field to the remote expert surgeon, and to provide the remote surgeon with a touch-based user interface to annotate the imagery of the operating field. The local subsystem’s function is to capture imagery of the operating field at the local site, and to display the remote expert surgeon’s annotations directly in the local surgeon’s field of view.

In a typical workflow using STAR, each surgeon can communicate verbally with each other using an audio channel. The remote surgeon views live images of the operating field, and then creates and transmits a relevant annotation of the next step of the operation to the local surgeon. After the local surgeon has either followed the instruction or indicated an understanding of the next step, the remote surgeon can erase the annotation if desired and illustrate the next step of the operation.

Figure 1 shows the local subsystem. The local subsystem is an off-the-shelf tablet that is positioned between the operating field and the local surgeon’s head. It should be noted that the form factor of a tablet is lightweight, portable, and suitable for prolonged field care missions where space is a planning constraint and heavy or bulky equipment would be suboptimal. The tablet captures and displays a live video feed of the operating field using its on-board camera. As a result, the local surgeon is always able to view the operating field by looking at the tablet display, as if the display were a transparent window. The captured video feed of the operating field is wirelessly sent to the remote subsystem.

Figure 2 shows the remote subsystem. The remote subsystem contains a display that allows a remote expert surgeon to view the operating field in real time. Because the

FIGURE 1. The STAR local subsystem, as seen from the local surgeon’s perspective.
position of the local subsystem’s camera is similar to the local surgeon’s viewpoint, the remote surgeon is able to view the operating field from approximately the same perspective as the local surgeon. The remote subsystem also contains a touch-based user interface, which allows the remote surgeon to draw annotation on top of the surgical view. These annotations include instructional visuals (e.g., lines and shapes), and icons of surgical instruments or predefined textual labels. The annotations are actionable physical knowledge conveyed to the local surgeon. These placed annotations can be moved, rotated, or scaled by the remote surgeon. Once the remote surgeon has finished creating a set of annotations of the current operating field imagery, the annotations are wirelessly transmitted to the local subsystem.

Once the local subsystem receives the remote surgeon’s annotations, it overlays the annotations (e.g., a polygon representing an incision region) onto the live imagery of the operating field. In this way, the local surgeon is able to view the operating field with relevant expert-provided annotations without ever needing to shift focus.

In the implementation of STAR described in this article, wireless communication was achieved using a Wi-Fi Direct connection to create an ad-hoc local area network between the two subsystems. Video frames from the local subsystem to the remote subsystem were delivered as low-resolution (320 × 200) PNG images at approximately 20fps. Annotation data from the remote subsystem to the local subsystem was in the form of text strings ranging between 1 and 100 KB depending on the complexity of the drawn annotation. Data were sent over a TCP connection for increased reliability.

During surgery, regions of the operating field move with respect to each other as the tissue is deformed as a result of the intervention. If the expert-provided annotations were static and unchanging in their appearance over time, then the operating field may shift underneath an annotation. This could result in an incorrect instruction, increased cognitive load for the less-experienced local surgeon who must mentally remap the now-invalid instruction to its correct location, and increased cognitive load for a remote surgeon who must recognize when an annotation has become invalid. To solve this, the local subsystem uses a series of computer vision algorithms to perform minor repositioning of the expert-provided annotations, such that the annotations appear anchored to the operating field elements that they describe. This annotation anchoring process allows the annotations to remain valid and useful even if the local subsystem is repositioned or if elements of the operating field change their relative position. To accomplish this annotation anchoring, the system first extracts data about salient features of the operating field near the location where the annotation was first created. Areas of high contrast, such as edges and corners, are automatically selected as salient features. In subsequent video frames, the local subsystem detects matching salient features, and, if these features have moved relative to the video frame, the annotation’s position is moved with it.

To validate that the STAR approach to telementoring was viable, an early STAR prototype underwent formative evaluation by surgeons during an Advanced Trauma Operative Management course at the Indiana University School of Medicine. A trauma surgery faculty operated the STAR remote subsystem in order to provide remote telementored guidance to a resident surgeon who was performing a fasciotomy on a euthanized porcine model (Figures 3 and 4). The resident surgeon was able to use the drawn annotations, as viewed through the local subsystem, to complete the operation.

Subsequent development incorporated the formative feedback provided by the surgeons after the conclusion of the test operation. For example, the mechanical assembly holding the local subsystem over the operating field was redesigned to be more stable. In addition, the remote expert surgeon reported that more free-form drawing tools like lines and circles were favored as annotations over more complex placement of icons. This suggests that the value of...
telementored annotation was in quick arrow-like indications, rather than a more detailed but slower construction of a more photorealistic illustration.

During the development of the current STAR prototype, two experiments were conducted, to validate the quality of the STAR system’s implementation, and to compare user performance of tasks between STAR and a traditional telementoring approach. This section provides an overview of each experiment.

**Experiment 1: Accuracy of Annotation Anchoring**

When an annotation sent from the remote subsystem is received by the local subsystem, it should be drawn at a location on the tablet screen that matches the operating field region at which it was first created. To evaluate the annotation anchoring algorithms used, pregenerated annotations were overlaid onto captured video frames of a simulated operating field (represented by either an anatomical print or a surgical simulator). In the video frames, the position of the simulated operating field was translated, rotated, and zoomed relative to the local subsystem. In the case of the surgical simulator, the operating field was also occluded and deformed.

Given these video frames as the input to the annotation anchoring algorithm, the local subsystem updated the positions of the pregenerated annotations each frame. The resulting annotation positions were compared against a ground truth position in each frame, and the distance between the positions was the annotation error. Errors above a particular threshold (20 pixels) were considered an anchoring failure, and errors below the threshold were considered an anchoring success.

**Experiment 2: User Validation**

To compare STAR with a traditional telementoring approach which relies on telestration (i.e., where annotations are provided on a monitor outside the local surgeon’s field of view), a user study was conducted. Twenty premedical and medical students performed two simulated surgical tasks under telementored guidance. The first task was to place a set of adhesives, representing the location of incision ports, onto the neck region of a patient simulator. The second task was to perform a multi-step abdominal incision on a patient simulator using several surgical instruments to be placed at precise locations. One half of the test subjects used STAR, while the other half used a telementoring approach that displayed each step of the operation on a nearby screen. Figure 5 shows the abdominal incision task for the STAR condition, where a series of icons of surgical instruments were displayed to guide the local surgeon to place real-life surgical instruments in the correct locations on the patient simulator. The independent variable was the choice of telementoring system, and the dependent variables were task completion time, placement error, and the number of times the subject shifted focus away from the operating field. Placement error was defined as the physical distance between the location on the operating field where the remote surgeon instructed the local surgeon to place a surgical instrument, and the location on the operating field where the local surgeon actually placed the surgical instrument.

**RESULTS**

In this section, results are provided for each of the 2 experiments, described earlier, that were conducted to validate the current STAR prototype from a technical perspective and from a user task performance perspective.

**Experiment 1: Accuracy of Annotation Anchoring**

Annotation anchoring was robust to translation, minor occlusion, and zooming out, succeeding in approximately 90 to 100% of the frames. However, major occlusion, in which large portions of the operating field were occluded by the local surgeon’s hands, succeeded only 60 to 74% of the time.
Deformation of the surgical field also led to lower success rates (15–63%), an expected result as the current anchoring approach assumes a rigid, planar operating field. Zooming is intended to result in more anchoring failures than zooming out; this is because when the operating field is zoomed in, there is less of the original operating field that is visible. As a result, the annotation anchoring algorithms have fewer salient features in the image with which to track the operating field elements relative to their original location.

**Experiment 2: User Validation**

For the two tasks (adhesive placement and abdominal incision) performed by the participants in the study, the subjects using STAR had lower placement error (45 and 68%, \( p = 0.003 \)), and also shifted focus away from the operating field less often (86 and 44%, \( p = 0.0003 \)) than the subjects who used the telestrator system. Those who used STAR also were 19% slower on average for each task, but this result was not statistically significant (\( p = 0.165 \)).

**DISCUSSION**

The initial results from user studies are promising, indicating that an AR transparent display approach to surgical telementoring can improve the quality of telementoring by improving surgeon accuracy and decreasing focus shifts. However, some limitations of the current STAR prototype exist: the remote surgeon is constrained to a small tablet-based system, the network connection between the subsystems is suboptimal in its bandwidth efficiency, and the local tablet’s display only gives an approximate sense of transparency to the local surgeon.

One limitation is that the STAR remote subsystem is currently constrained to a small tablet form factor, which reduces the ability of the remote surgeon to view the operating field in the context of the entire patient’s body. In future work, the remote subsystem will be adapted to a larger patient-sized interaction table, which will enhance the sense of co-presence for the remote surgeon by placing the operating field in its proper context.

There are some limitations in the described implementation of STAR regarding the network connection between remote and local subsystems. The use of lossless video frame encoding increases the bandwidth needed for transmitting video frames from the local site to the remote site. Future work will integrate more robust video streaming solutions for guaranteed delivery of video and annotation data in the austere environments that STAR would be used in. A surgical telementoring system for use in forward operating bases would require a robust solution to connectivity so that latency does not impede the ability for the local surgeon to receive adequate guidance. In addition, data transmitted between the subsystems should be encrypted for security reasons in a fully deployed system. However, it should be noted that any surgical telementoring system would need to address these issues, and that STAR’s approach to telementoring is not unique in this regard.

The STAR local subsystem acts as a window through which the local surgeon views the operating field, but currently it only gives an approximate window-like appearance. The captured video frames of the operating field are directly shown on the tablet display, and does not change based on the local surgeon’s current viewpoint. As a result, objects in the operating field may appear with a position or size that is different from what the local surgeon would see if the tablet were not present. Work is ongoing into investigating simulating a transparent display by capturing the 3D geometry of the operating field, and tracking the local surgeon’s head position with cameras, to generate a real-time image on the tablet display that appears transparent to a particular viewpoint. Integrating such simulated transparent display technology into the STAR telementoring approach will result in a system that enhances the co-presence perceived by the local surgeon.

It is also valuable to consider the wider application of STAR’s approach to surgical telementoring. Beyond military medicine, STAR also has potential applications for public and rural health care, where trauma patients at hospitals that lack specialized expert surgeons may receive urgent, lifesaving care by a local general surgeon receiving telementored guidance. Outside the context of urgent trauma care, STAR has the potential to enhance teleproctoring by allowing a small number of expert surgeons to oversee and interact with a large number of remote students of surgery. By increasing access to specialized training, improved telementoring systems can act as a force multiplier for the limited resource of expert surgeons in the world.

Additional work will involve enhancing the remote surgeon’s experience by providing a large interaction platform, complete with hand gesture controls, to allow the remote surgeon to perform more intuitive actions in the course of guiding a less-experienced surgeon through a surgical procedure. Furthermore, the annotation anchoring capabilities of the local subsystem will be enhanced to better support the occlusion and deformation of the operating field that frequently occurs during an operation. Finally, as the field of AR continue to improve, STAR will serve as a robust sandbox platform for further testing and validation of new AR technologies, such as Google Glass and Microsoft HoloLens, which may prove to be useful tools for enhancing the ability of the remote and local surgeons to interact as if they were physically co-located.

**CONCLUSIONS**

STAR improves on traditional telementoring systems by reducing the need for less-experienced local surgeons to shift focus away from the operating field, and by anchoring virtual expert annotations to the operating field elements that they describe. Initial user studies indicate that STAR can improve a less-experienced provider’s accuracy when
performing surgical tasks. These benefits promise to improve the care provided to soldiers who receive traumatic combat injuries by increasing the level of co-presence between local and remote surgeons in a telementored operation in the austere environment of a forward operating base.

ACKNOWLEDGMENTS
This work was supported by the Office of the Assistant Secretary of Defense for Health Affairs under Award No. W81XWH-14-1-0042. Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the Department of Defense.

REFERENCES
Telesurgery With Miniature Robots to Leverage Surgical Expertise in Distributed Expeditionary Environments

Mark Reichenbach, BS*; Tom Frederick, MS*; Lou Cubrich, BS*; Walter Bircher, BS†; Nathan Bills, PhD, MBA‡; Marsha Morien, MSBA, FHFMAMISS; Shane Farritor, PhD†; Dmitry Oleynikov, MD, FACS‡

ABSTRACT This study aimed to evaluate the capability of performing telesurgery via radio transmission for military arenas where wired internet connections may not be practical. Most existing robotic surgery systems are too large to effectively deploy with first responders. The miniature surgical platform in this study consists of a multifunctional robot suite that can fit easily into a briefcase. Methods: The focus of this study is to explore the implications of radio control of the robot. The hypothesis is that an in vivo robot and its control boards can be controlled using off-the-shelf wireless components. An experiment was designed with off-the-shelf wireless components to test the capability of our newest generation of miniature surgical robot to become battery-operated and wireless. Results: Wireless transmission of control signals has provided proof of concept and has exposed areas of the software that can be built upon to improve responsiveness. Wireless transmission of the video feed can be adequately performed with basic off-the-shelf components.

INTRODUCTION
As the military disperses surgical facilities so that patients receive care closer to where they are stationed, it becomes important to maintain and improve outcomes by making the most sophisticated diagnostic and interventional care as widely available as soon as possible. Since it is not possible to deploy limited medical personnel and resources to all arenas, secure high-speed data links available can be leveraged to more efficiently deploy specialists. Several robotic surgery systems have been developed, and some have demonstrated telesurgery over the internet. However, current systems, such as the da Vinci Surgical System (Intuitive Surgical, Inc., Sunnyvale, California), are prohibitively large and cannot be easily deployed. Smaller systems would lend themselves better to deployment in Combat Support Hospitals (Level 3), or with Forward Surgical Teams (FSTs) (Level 2). When an emergency surgery is required, or when the patient cannot be immediately transferred to the next-level care facility, the availability of advanced surgical options can make all the difference. Improved evacuation procedures and an emphasis on rapid transit to higher levels of care are decreasing combat died of wounds and death on the battlefield.1-3

Research at the Center for Advanced Surgical Technology and the Advanced Surgical Technologies Lab at the University of Nebraska has the goal of enabling telesurgery through remotely controlled miniature surgical robots. This is a force multiplier for a consolidated medical footprint, making expert highly skilled surgeons available for procedures in remote-distributed environments that lack conventional surgical resources. The miniature robots are small and simple enough to be deployed and set up by an FST; an experienced surgeon could then log in remotely and perform emergency procedures.

Presented here is the introduction of a wireless communication system that would allow deployment of these robots to be controlled remotely by a skilled surgeon. Testing was focused on the performance of wireless response of the robot, future work will evaluate the effects on robot operation. This advance in surgical technology can reach remote locations where existing communication infrastructure is not available. Research is ongoing to increase the clinical capabilities of the robot, but developing a wireless control network would greatly increase range of potential deployments of the telesurgical technology.

BACKGROUND
Advances in surgical technology have improved the patient experience greatly. Laparoscopic surgery revolutionized how many procedures are performed. Research and development have begun producing robotic Minimally Invasive Surgery (MIS) systems that further improve patient safety;
including decreased operative blood loss, postoperative pain, faster return of normal function, and fewer analgesics.\textsuperscript{4,5} However, these improvements have not made their way out of the hospital-based surgery suites. With operations all over the world, the U.S. military deploys thousands of personnel to dozens of countries. Some areas are war torn and present a front line where soldiers or civilians may be injured or are in need of emergency care. Other cases may involve humanitarian aid for populaces of affected countries. Bringing advanced surgical technologies to these emergency scenarios could save thousands of lives all over the world.

We have designed a compact surgical robot suite comprised of a miniature surgical robot, high-quality vision system, user interface software, haptic controllers, and telesurgery capabilities. This system has been tested in vivo in multiple porcine surgeries, and fits into the abdomen through a two-inch incision at the navel, thus enabling remote, MIS. The design of this robot has been licensed to a private company, Virtual Incision Corporation, which has performed two first inhuman surgeries and is currently in discussions with the FDA to submit 510(k). This study focuses on controlling the compact surgical robot suite wirelessly. In the future, using existing high-speed defense communication backbones and protocols, skilled surgeons located anywhere will be able to remotely assist in complex procedures taking place hard to reach or remote locations where FST are deployed.

The commercially available da Vinci Surgical System has been the leading clinical robotic MIS system since it received FDA approval in 2000, and is currently the only surgical robot with U.S. FDA approval for use in laparoscopic surgery.\textsuperscript{6} Surgeons can control laparoscopic tools from a remote workstation. The system can filter tremors, and provides 3D vision,\textsuperscript{7} increasing the quality of the surgery. The da Vinci system has successfully demonstrated the ability to perform procedures on a porcine model over the internet. The distances were 1,300 and 2,400 miles. Round-trip delays of 450 to 900 milliseconds were demonstrated. The system used public internet for communication.\textsuperscript{8} Later testing across a 17 MB/s bandwidth VPN network over 1,770 miles showed 370 milliseconds of delay.\textsuperscript{9}

Other systems are under development. The Raven-II is a collaborative research effort that is built around three 3-degree of freedom (DOF) arms with interchangeable 4-DOF instruments. Although the da Vinci is a commercial product, the Raven-II is an open-source platform jointly built by seven universities.\textsuperscript{10} The system is capable of similar teleoperation to the da Vinci, and has been controlled by various off-the-shelf controllers.\textsuperscript{11} The Raven-II has also demonstrated telesurgery capabilities; from 100 m (wireless) to 4,700 miles (commercial internet). The system was tested from Seattle, Washington, to London, England, and exhibited 140 milliseconds of internet latency.\textsuperscript{12} Similar robotic systems are being developed\textsuperscript{13} using arms outside the body to position tool tips. The Raven and da Vinci have both demonstrated telesurgery capabilities; deploying these systems, with large arms and actuators around the patient, do not lend themselves to easy set up in emergencies or remote sites. In an emergency, there is a huge range of injuries that may require treatment. Not every scenario will be applicable for these robotic systems. Most of the current robotic surgery systems focus on abdominal surgery due to the large workspace available, which leaves a gap that negatively impacts the systems’ applicability. However, as the systems’ functionality increases, there is no doubt that the ability for a surgeon to remotely control two dexterous tools would save lives. Although current procedures for evacuation are improving, the rate of combat casualties has decreased.\textsuperscript{1} Inevitably, there will be situations where evacuation beyond Level 2 is not possible; a wireless system could improve the treatment of combat injuries in these scenarios. Nothing will be able to replicate a controlled surgical setting. However, it is possible to deploy the proposed system for surgery with the skills and equipment available to a FST that include administering anesthesia and inserting a port. Since the type of injuries can vary greatly, and the robots being developed are mostly specialized, there is not much overlap. Different surgical procedures pose unique difficulties including access to a specific site, large workspace, numerous veins, and vessels to control.\textsuperscript{14} An insertable robot has easier access to the entire abdomen; where the da Vinci and Raven have difficulty maneuvering due to the robotic arms outside of the patient. Insertable robots simply need to be rotated on the axis of insertion to reach the entire abdomen. The system is more manageable to deploy and set up requires minimal motion compared to the larger systems that are currently available.

Several different types of insertable robots have been developed; however, the only robot with FDA approval for use in laparoscopic surgery in the United States is the da Vinci Surgical System. The BioRobotics institute in Pisa, Italy, has developed a two-arm robot, SPRINT, which is inserted through a single port; control has been demonstrated, but a specialized trocar is necessary.\textsuperscript{15} SPRINT uses an off-board control system, with cables running to external controllers, which negatively impacts its ability to be transported easily, as in an emergency situation. Further, the system has demonstrated no telesurgery capabilities.\textsuperscript{16,17}

A snake-like robot, developed by Waseda University, is inserted through a single port; the system deploys tools from the main tubular body. The robot is positioned by a robotic arm; the end effectors are actuated using a cable drive system. The system has no telesurgery capabilities\textsuperscript{18} and the positioning arm and off-board actuators make deployment difficult.

The i-Snake robot, from Imperial College London, is inserted through a standard trocar and has a flexible head that deploys; in addition, two cable-driven arms are inserted to manipulate tissue. However, the system cannot properly position tools for tissue manipulation. The system has demonstrated no telesurgery capabilities.\textsuperscript{19,20}
The IREP robot from Vanderbilt is similar to the i-Snake—it features two cable-driven arms and a head that provides vision. The cable-driven arms have a cumbersome actuation housing, making transportation and deployment difficult. Limited force at the end effectors and lack of wrist dexterity limit the system. The system has demonstrated teleoperation, but only over a local area network. This would not be applicable to deploying surgical tools and technology in hard-to-reach or emergency areas. 21,22

Previous work has demonstrated the functionality of our insertable robots 23,24 These two-armed in vivo robots have performed over 75 in vivo procedures, including cholecystectomies and a partial colectomy in live porcine models. The robots have demonstrated the capability of manipulating tissue in all quadrants of the abdominal cavity in a live porcine model. 25 Although the robots have only been tested in the abdomen, it is possible that in these emergency scenarios the robot could be used for a more general purpose. Even if the robot is not inserted in the abdomen the surgeon still has control of the tools, and can have the person deploying the system orient the robot wherever is necessary. The robot is actuated by onboard motors and control boards, thus there is minimal cabling or housing that needs to be transported with the robot.

Surgical systems have been developed that have tele-operation capabilities. However, the systems that have more fully developed these features are bulky, and cannot be deployed cheaply, efficiently, or easily to emergency areas; whether that is a front-line or humanitarian aid to an affected area. The robotic system presented here has proven capable of surgical procedures, and the size and simplicity of the system make it adaptable to both external and internal procedures. By testing and implementing the proposed wireless system, the transportable robotic system can provide surgical care to remote areas with FST capability. With this system, U.S. military personnel could virtually deploy a surgical care system, U.S. military personnel could virtually deploy a surgical care from the military infrastructure, could reach remote locations, and provide immediate care from experts. The hypothesis being tested is that by controlling the current in vivo robot with off-the-shelf wireless controllers, a uniquely small and compact robotic surgery system can be deployed remotely. The wireless robotic system presented here, coupled with available video technology, could expand the capabilities of current evacuation protocols on a war or humanitarian front.

METHODS

The miniature in vivo surgical platform consists of a two-armed, gear-driven, multifunctional robot and a remote surgeon interface (Fig. 1). The surgical robot is normally attached to the control computer via a direct USB connection. This USB connection was easily made wireless using a pair of radio frequency (RF) transceivers, XBEE multipoint RF Modules (Digi International, Minnetonka, Minnesota). The video was transmitted on a dedicated set of radios.

The miniature in vivo robot consists of two, 4-DOF arms each composed of a 2-DOF shoulder, elbow, and wrist. Each arm is about six inches long. Various tools may be used for each end effector, including grasper, cautery, and suction/irrigation tools. Each joint is driven by on-board brushless DC motors and miniature gear trains and each link houses custom, modular motor controllers, reducing the cabling along each arm to four wires: a 12V DC power bus and a RS-485 differential serial bus. The serial bus is converted from RS-485 to USB and input into a miniature computer (RaspberryPi, 900 MHz quad-core ARM Cortex-A7 CPU, with 1 GB RAM, running Raspiian, Caldecote, United Kingdom) USB hub. An off-the-shelf RF transceiver was also connected to a USB port and a Python script was written to connect the two communication ports to each other. The remote robot is shown in Figure 2. Future studies will move toward removing the miniature computer from the system. The mobile robot platform only requires a 12V-3A and a 3V3-650 mA power supplies, both of which could be drawn from a single battery pack.

The remote surgeon interface consists of user input devices, a mobile laptop computer (Dell Precision M6500 running Windows 8.1 with Intel Core i7 processor at 1.87 GHz, and 10 GB RAM, Round Rock, Texas), a custom software package, and an off-the-shelf USB-RF transceiver, as shown in Figure 3. A Custom software architecture (University of Nebraska-Lincoln, Lincoln, Nebraska) is used to handle user input, robot control, and communication with the motor controllers. The robot is operated using off-the-shelf Geomagic Touch haptic controllers (3D Systems, Rock Hill, South Carolina). These motorized haptic devices provide a bidirectional communication interface between the controller and master computer, outputting absolute position coordinates and receiving force inputs from a suite of plug-ins. Force input is applied to produce the barriers of the robotic workspace and transmit forces from the actual robot to the hand of the operator. Scaling and clutching features were also used to gain finer control of small motions.

The latency of the robot control communications was tested by measuring the response time of the on-board motor...
controllers to a robot command from the custom control software. A timer was stared when a message was sent to the robot, and stopped when it received the reply from the motor control boards. The time was measured for 3,000 successful responses and averaged. Although a response time benchmark for “smooth” operation has not been quantified, the response time of the wired robot platform has been deemed to be responsive enough to perform surgical tasks by the surgeon-author during benchtop and in vivo tests.

Providing a real-time video connection that is synced with the hand controllers is vital to a successful procedure. Exploration of this component started with a transmitter/receiver pair of Partom 5.8GHz 1200mW (Shenzen Partom Technology Development Co., Ltd., Guangdong, China) radios designed for analog audio/visual transmission. The Partom module is a common first-person view video setup for use in radio-controlled aircraft. Mushroom antennas were used for both the transmitter and receiver, providing a gain of −3 dBi and a rated range of 3 to 4 km of open air transmission. This particular product also recommended a set of 14 dBi panel antennas for a range of more than 14 km. These units performed as expected and will be explored further for use in carrying the control commands as well.

The performance of the remote platform was compared to the performance of the completely wired local system. The same laptop computer was used for both cases. The video system was verified, but not range tested at this time.

**RESULTS**

The existing computer software was tested while communicating with the robot at several different baud rates. The baud rates of the USB-RF transceivers as well as the on-board robot control boards were varied. The time required for 3,000 successful responses was recorded. The recorded time is based on the quality of the wireless connection. The results of these tests are shown in Table I.

Wired connections consistently timed in at around 20 milliseconds per successful round trip. The current robot uses six motor controllers, resulting in about eight messages per second to each motor controller. This communication

<table>
<thead>
<tr>
<th>Baud Rate</th>
<th>Wired Response Time (milliseconds)</th>
<th>Remote Response Time (milliseconds)</th>
</tr>
</thead>
<tbody>
<tr>
<td>9,600</td>
<td>48.325</td>
<td>233.311</td>
</tr>
<tr>
<td>38,400</td>
<td>19.480</td>
<td>78.526</td>
</tr>
<tr>
<td>57,600</td>
<td>19.679</td>
<td>80.978</td>
</tr>
<tr>
<td>115,200</td>
<td>19.178</td>
<td>85.412</td>
</tr>
</tbody>
</table>
speed has been determined to provide smooth operation of the robot for the number of controllers used in the current kinematic configuration. The wireless setup used in this proof of concept varied from 78 milliseconds per successful message loop at a baud rate of 38,400 to 233 milliseconds at a baud of 9,600.

DISCUSSION

Through testing, several potential bottlenecks were identified and discussed.

The experiment described here used a Raspberry Pi Linux computer to handle the messaging between the robot and the USB-RF dongle. The overhead and slow speed of this computer was taken into consideration. The same code was run on a more powerful desktop running Windows 7 with an Intel Core i7-2600K CPU at 3.4 GHz and 8 GB RAM. This computer showed approximately a 15% improvement in the 1,000 message test, but still had overhead slowing it down. The next step will be to design dedicated circuitry to connect the robot directly to the radio module.

Although the primary robot control software is efficient enough for wired communication, additional improvements could be implemented to help lessen the effect of dropped packets. The same can be said for the embedded software on the robot control boards. Work will be continued to combine both video and robot control signals into the same RF transceiver. A spooling function or command buffer could be implemented for smoother control of the robot, but this smoother control would come at the expense of response time.

CONCLUSION

The introduction of a radio network to our current robotic system can improve the military’s ability to provide remote surgery care in situations where FST are available. Current technology is capable of teleoperation; however, the most robust systems are large and incapable of deployment to the emergency scenarios highlighted. Smaller robots have been developed, but not with satisfactory teleoperation. The system highlighted is capable of both. Wirelessly controlling the in vivo robot is an initial step in making a mobile surgical system; small and simple enough that it can be deployed in a wide variety of urgent cases and bringing expert surgeon control to remote locations. The robotic system requires minimal medical training, including administering anesthesia and making an incision. The minimal training makes the logistics of deployment easier, in situations where evacuation may not be possible. The collected data show that wireless control is possible, but the system is limited by hardware. Further development can mitigate these issues.

Our group has been able to control the surgical robot over a wireless network, demonstrating platform capability for telesurgery. Military capability in remote robotic guidance has already been amply demonstrated by the successful use of unmanned aerial vehicles, and numerous high-speed defense communication links and protocols already exist—all of which will ease the implementation of robotic telesurgery. This technology provides a viable solution to the lack of immediate surgical care in remote or inaccessible locales and can have a revolutionary effect on remote surgical care. The work shown here is a proof of concept showing that controlling the current in vivo robot with wireless off-the-shelf components is feasible, and leaves room for promising future work.

ACKNOWLEDGMENTS

The U.S. Army Medical Research Acquisition Activity, 820 Chandler Street, Fort Detrick, MD 21702-5014 is the awarding and administering acquisition office. This work was supported by the Office of the Assistant Secretary of Defense for Health Affairs under Award No. W81XWH-14-1-0058.

REFERENCES


Empirical Study on the Impact of a Tactical Biosurveillance Information Visualization on Users’ Situational Awareness

Valeriya V. Kettelhut, MD, PhD, MPH*; Trevor C. Vanschooneveld, MD†; James C. McClay, MD‡; David F. Mercer, MD, PhD§; Ann Fruhling, PhD∥; Jane L. Meza, PhD¶

ABSTRACT Decisions on antibiotic-resistant infection (ARI) prevention in dynamic health care settings should be agile and target the right process at the right time. Health information technologies can aid the recognition of high-risk situations for ARI transmission and timely facilitate operators’ situational awareness (SA) in various military and civilian health care locations or transport platforms. High SA is one of the significant predictors of better performance. The objective of this study was to evaluate the impact of the developed health information visualization (VIZ) on the users’ SA regarding situations when risks of ARI transmission and exposure are high. The enrolled 19 subjects assessed the proposed VIZ artifacts representing 1 scenario, compared the VIZ effectiveness against the currently employed local methods, and reported their SA (perception and comprehension) with the use of a pre- and post-self-rating questionnaire. The results showed that the VIZ significantly increased SA in the study subjects and revealed the importance of communicating the risk of exposure to ARIs. The VIZ enabled the participants to quickly acknowledge the high-risk individuals (super-spreaders), locations (hot spots), and biosafety (deficient infection prevention). The study concluded that SA-oriented technologies may be promising for promoting better infection prevention practices.

INTRODUCTION Hazards associated with antibiotic-resistant infections (ARIs) force health care practitioners (HCPs) to seek innovative approaches to slow the emergence of new ARIs and prevent their spread. Properties of health care settings, such as population density, prevalence of and proximity to diseased cases, clustering of contacts, repeated contacts, and contamination of personnel, environment, and equipment, are the significant contributing factors to the spread of ARIs. It is hypothesized that constantly changing patterns in spatial distribution and prevalence of infectious cases may compromise the effectiveness of infection prevention and control. Thus, decisions on infection prevention should be agile and target the right process at the right time. Health information technologies can aid the recognition of high risk for infection transmission situations and timely facilitate the communication of a “Common Operating Picture” in various military and civilian health care locations or transport platforms.

In dynamic environments, decision-making is highly dependent on situation awareness (SA), an essential attribute of high-reliability organizations. The SA concept in health care relates to the analysis of patient safety and health care quality issues. Several studies provided evidence that SA systems have been associated with better patient outcomes. Currently, many hospitals use infection surveillance information to increase staff awareness about patients with known ARIs by flagging them in electronic health record (EHR) systems. Literature lacks evidence on whether this practice is sufficient for decreasing the risks for infection transmission and improving patient outcomes.

Computerized platforms can recognize individuals who are at the most risk of exposure to ARIs in real time. Health data visualization (VIZ) can enable HCPs to gain a high level of understanding the risks quickly. For this, providing the right information in the right format is important. The challenges for an interface designer include understanding the actionable information and utilizing the effective VIZs that support perception and comprehension. By combining geographical data on health events with the location of hazards and preventive and therapeutic services, an information system can reveal spatial associations between exposures, services, and health outcomes.

Depicting infection surveillance data with maps has long been a standard approach to illustrate geographic clustering and regional differences in disease prevalence or incidence. Graphical structures like maps can immediately communicate essential points or critical cues. In health care, maps can also furnish large amount of EHR data and condense them into a single image that would alleviate informational overload caused by fragmented and granular EHR data.

The investigator developed the conceptual model (Fig. 1) and the VIZ (Fig. 2) to visualize the hospital spatial data linked to individual patients’ clinical-laboratory data for tracking
diseased patients’ locations, antibiotic administrations, contact data, and infection prevention intervention data.

The goal of the VIZ was to effectively increase SA of HCPs about the hospital areas at high risk for ARI transmission and subjects at high risk of exposure to ARIs. To ensure the efficiency of the VIZ, the investigator applied the SA-oriented system approach and the rules for image construction. It was expected that the VIZ would prompt HCP’s understanding of the local setting’s epidemiological situation and to enable HCPs to answer the questions, “At a given location, what are the risks of infection transmission?” and “At a given place, who is at a greater risk of exposure to ARIs?” The rules of image construction facilitate effective data VIZ by taking advantage of visual perception, whereas the SA approach helps identify critical cues that reduce informational overload and aid comprehension.

The objective of this study was to assess the impact of the VIZ on users’ self-reported SA. It was hypothesized that the novel VIZ would increase subjects’ SA, including perception (Level 1 SA) and comprehension (Level 2 SA): the effect would be greatest in those without access to all patient records due to role-based access, without prior knowledge about patient surveillance status, and with less experience.

METHODS

The study employed a quasi-experimental survey with pre-and postdesign. The survey included: 1) administration of the prequestionnaire followed by the introduction of the VIZ, 2) administration of the postquestionnaire, and 3) collection of the verbal feedback. The survey took place in April 2015.

The study setting chosen was a 50-bed medical-surgical unit at a Midwest teaching hospital. Participants included care technicians and nurses involved in the care in this unit, hospital infection preventionists, and clinical consultants.

The subject enrollment methods included: 1) a solicitation sponsored by the unit management group, 2) a solicitation sponsored by the Executive Director of Infectious Disease

FIGURE 1. The conceptual model of antibiotic-resistant infection transmission.

FIGURE 2. The visualization for real-time infection transmission risk assessment (TBioS: Tactical Biosurveillance User Interface).
and Epidemiology Department, and 3) through a verbal solicitation by the investigator. The sampling addressed the need to study individuals varied in their tasks and locations, including the front-line staff and infection preventionists, but used the same EHR-based hospital infection surveillance method. The investigator selected subjects in a conventional manner, obtained the informed consent, and stratified them into two groups, “unit-based staff” vs. “non-unit-based staff.” The “non-unit” group consisted of the clinical consultants and infection preventionists. The “unit-based” group included the staff only. The Institution Research Board approved this study (IRB number 171-15-EX).

A questionnaire measured subjects’ perception (Level 1 SA) and comprehension (Level 2 SA) of the infection transmission risk factors with the use of the VIZ vs. the current EHR-based data. The questionnaire included 1) seven SA-items in the pre-test (Q1–Q7) and post-test (q1–q7) to measure Level 1 SA and Level 2 SA, 2) one item in the pre- (Q8) and post- (q8) tests to measure a belief about the importance of knowing the risk of exposure to ARIs, 3) one multiple-choice item in the post-test (q6a) to measure subject’s direct performance, and 4) post-test items (q9–q12) to measure the subject’s perception of the VIZ and the training material usability. This questionnaire served as a self-rating instrument. Pre- and postmeasurements of each individual response were taken on a 5-point Likert scale, where 1 = lowest score and 5 = highest score.

The respondents reported their self-perceived SA levels about 1) patients who are carriers of ARIs, 2) locations of these patients on the unit, 3) type of ARI organisms present on a given day, 4) patients who were prescribed with antibiotics, 5) use of the infection prevention intervention, 6) locations where the risk for ARI transmission may be high, and 7) individuals at risk of exposure to ARIs.

The participants took approximately 30 minutes out of their work schedule to complete the study. The pretest took approximately 7 to 8 minutes. The pretest data measured the subjects’ baseline SA attributed to their current practice of the EHR use. Then, the investigator provided brief training to the group by introducing the VIZ and explaining the meaning of the VIZ’s visual artifacts and concepts with the use of the training material. Subsequently, the participants reviewed the VIZ that displayed the data on a given day. All participants then answered the post-test questionnaire at their pace, which took, on average, 10 minutes. At the end, the subjects provided verbal feedback on the VIZ usability, which was recorded and analyzed.

**Statistical Analysis**

The sample size was 19 participants. The Wilcoxon sign test compared the median of two dependent samples. The statistical program was SPSS v.21 (IBM, Armonk, New York).

The dependent variables included Level 1 SA, Level 2 SA, total SA, and usability of the VIZ. The independent variables included users’ experience (years of health care experience), role (nursing aid, nurse, and physician), and access to patients’ records in the EHR.

The SA measurements operational definitions:

1. A measurement of Level 1 SA is a sum of the self-perceived rating scores in the answers about patients who are carriers of ARI bacteria (Q1/q1), locations of these patients in the unit (Q2/q2), type of ARIs identified in these patients (Q3/q3), patients who receive antibiotics (Q4/q4), and patients who do not have chlorhexidine bathing in the previous 24 hours (Q5/q5).

2. A measurement of Level 2 SA is a sum of the scores in the answers about circumstances when the risk of ARI transmission increases (Q6/q6) and individual level of exposure to ARIs (Q7/q7).

3. A measurement of Total SA is a sum of Level 1 SA and Level 2 SA.

**RESULTS**

The study sample included 10 (53%) medical–surgical unit staff members and 9 (47%) nonunit-based consultants. There were 16 (84%) registered nurses (RNs), one (5%) resident (MD), and one (5%) nursing assistant (NA). The nursing group consisted of nine unit-based RNs, three nonunit palliative care RNs, and five nonunit infection preventionist RNs. Seven of 19 (37%) HCPs had less than 5 years health care experience and were the unit-based RNs. All participants had access to the EHR: the infection preventionist RNs have authority to access all patients’ records, consulting HCPs and unit bed-side RNs can access only their patients’ records, and the unit managers can access any patient record if a patient is bedded on their unit.

**Construct Validity**

The Cronbach’s α for the seven items was 0.891 and the corrected item-total correlation for each item was >0.30.

**Team SA Measurements**

The baseline median overall-team total SA EHR was lower than the median total SA VIZ (2.29 vs. 4.57) (Table I). The Wilcoxon signed ranks test showed a significant increase in the median overall-team total SAEHR vs VIZ (p < 0.001). The mean significantly increased the median Level 1 SA EHR vs VIZ (1.8 vs. 4.6, p < 0.001) and Level 2 SA EHR vs VIZ (2.0 vs. 4.5, p < 0.001).

After stratifying the overall-team SA score by the questionnaire items, the median SAEHR scores showed high variability for Q1, Q2, Q3, and Q7 (range 1–5) (Fig. 3) compared with the median SAVIZ scores, most of which were 5.00 (“ceiling effect”) (Fig. 4).

**Unit-Staff and Nonunit Staff SA Measurements**

After stratifying the subjects into unit staff and nonunit staff, the unit staff showed higher baseline median Level 1 SAEHR.
Level 2 $SA_{EHR}$ and total $SA_{EHR}$ scores than the nonunit staff’s scores (2.50, 2.50, 2.50 vs. 1.40, 1.00, and 1.29) (Figs. 5 and 6). The nonunit staff had the lowest baseline median Level 2 $SA_{EHR}$ score (1.00).

Overall, the median total SA scores significantly increased for each group after using the VIZ; although, the unit staff ($n = 10$) had a smaller increase in the median $SA_{EHR}$ ($1.00$) vs. VIZ ($4.36$, $p = 0.011$) than the nonunit staff ($n = 9$) median $SA_{EHR}$ vs. VIZ ($1.29$ vs. $4.86$, $p = 0.008$). The median Level 1 SA and Level 2 SA scores significantly increased in both groups. The greatest magnitude of the increase was in the nonunit Level 2 $SA_{EHR}$ vs. VIZ score ($1.50$ vs. $5.00$, $p = 0.001$), whereas the lowest magnitude of the increase was in the unit staff for the $SA_{EHR}$ vs. VIZ ($1.00$ vs. $4.00$, $p = 0.001$) score. The Wilcoxon signed ranks test showed 1) a significant increase in the median $SA_{EHR}$ vs. VIZ scores for Q3, Q4, and Q6; 2) no difference for Q1; and 3) insignificant difference for Q2, Q5, and Q7 in the unit staff. The nonunit staff had a significant increase in the median $SA_{EHR}$ vs. VIZ scores for all questionnaire items.

**Unit Novice vs. Unit-Experienced Staff SA Measurements**

The unit staff was stratified into the novice ($n = 7$) group (1–5 years in health care) and the experienced ($n = 3$) group (>5 years in health care). The novice group had a higher baseline median Level 1 SA, Level 2 SA, and Total SA than the experienced group (2.40, 2.59, and 2.43 vs. 1.50, 1.00, and 1.36). The experienced group tended to have a greater median Level 1 SA, Level 2 SA, and total SA scores (4.80, 4.80, and 4.79) when using the VIZ comparing with the

---

**TABLE I.** Median Overall-Team Situational Awareness (SA) Score (Electronic Health Record/EHR) Vs. Health Data Visualization (VIZ)

<table>
<thead>
<tr>
<th>Questionnaire Item (Median)</th>
<th>EHR (Pretest)</th>
<th>VIZ (Post-test)</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item 1</td>
<td>3.00</td>
<td>5.00</td>
<td>0.017</td>
</tr>
<tr>
<td>Item 2</td>
<td>2.00</td>
<td>5.00</td>
<td>0.001</td>
</tr>
<tr>
<td>Item 3</td>
<td>2.00</td>
<td>5.00</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Item 4</td>
<td>1.00</td>
<td>5.00</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Item 5</td>
<td>1.00</td>
<td>5.00</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Level 1 SA Score</td>
<td>1.80</td>
<td>4.60</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Item 6</td>
<td>1.00</td>
<td>4.00</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Item 7</td>
<td>2.00</td>
<td>5.00</td>
<td>0.001</td>
</tr>
<tr>
<td>Level SA 2 Score</td>
<td>2.00</td>
<td>4.50</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Total SA Score</td>
<td>2.29</td>
<td>4.57</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

---

**FIGURE 3.** Median overall-team situational awareness (SA) electronic health record score by items ($n = 19$).

**FIGURE 4.** Median overall-team situational awareness (SA) health data visualization score by items ($n = 19$).

**FIGURE 5.** Median unit staff situational awareness (SA) score electronic health record (EHR vs. health care visualization (VIZ) ($n = 10$).
Novice group (4.20, 4.50, and 4.28), a pattern similar to the nonunit staff.

**SA Direct Performance Measurements (q6a)**

The participants were asked to identify at least one location at high risk for infection transmission using the VIZ that displayed 1-day data. Seventeen individuals out of 19 (89.47%) provided their responses. All of these responders correctly (100%) identified one location. Two subjects (10.53%) did not respond, yielding the accuracy rate of 89.47%.

**Risk of Exposure to Infections Score (Occupational Hazard Q8–q8)**

The items Q8/q8 measured a belief about the importance of knowing a personal occupational exposure to ARIs. The Wilcoxon signed ranks test showed a significant increase in the median overall team Q8EHR vs. VIZ score (4.00 vs. 5.00, \( p = 0.03 \)). The baseline median score for Q8 was high (4.00), ranging from 2.00 to 5.00. After using the VIZ, the median score for q8 increased to 5.00 (\( p = 0.03 \)), ranging from 3.00 to 5.00.

**Perception about Usability of the VIZ and Training Material**

Fourteen subjects of 19 (70 %) perceived the VIZ as a quicker and easier way to be informed about the risks associated with ARIs than with the EHR-based current practice (q9). Eighteen subjects (94.73 %) perceived that the use of visual artifacts (q10) and infection prevention deficiencies (q11) persuasively alerted them about the presence of infection transmission hazards. Fifteen subjects (78 %) responded that the training material was easy to understand (q12). The team median score for q9 was 5.00 (range 3–5), q10 was 5 (range 3–5), q11 was 4.00 (range 2–5), and for q12 was 4.00 (range 2–5).

**Qualitative Feedback**

After taking the pre- and postsurveys, each group had a verbal discussion about the VIZ. The comments were analyzed and transcribed into the following themes: 1) the VIZ’s impact on clinical practice, 2) a need for the knowledge about occupational exposure to ARIs, and 3) the VIZ usability (Table II).

**DISCUSSION**

This study provided some empirical evidence on the effects of the ARI VIZs on the health care staff’s perception and comprehension of high risk for infection transmission situations. The study employed seven subjective SA measures and one objective SA measure via the pilot questionnaire. The nonunit group members had more than 5 years of experience. The unit group consisted of 7 of 10 (70%) unit-based members with less than 5 years in health care.

The study provided pilot evidence that the group significantly increased the total SA when the VIZ was used. When

---

**TABLE II. Qualitative Feedback on the Health Care Visualization (VIZ) Effects**

<table>
<thead>
<tr>
<th>Domain</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Impact on Clinical Practice</td>
<td>“People do not really understand the need to approach patient care from a “clean-to-dirty” perspective, when “clean” patients need to be taken care before approaching “dirty” patients. This interface would help to direct care planning.” “There is a challenge to make decisions how to better arrange the unit staff. Currently, the managers match patient acuity and staff skills level. The infection prevention module is one piece among many others the managers have to address.” “I see how this interface makes the infection prevention and control practice meaningful: patient chlorhexidine bathing clearly appears as a protection for the staff from being exposed to antibiotic-resistant infections. This information is very persuasive for us to do our best.”</td>
</tr>
<tr>
<td>Occupational Exposure</td>
<td>“We are often desensitized to our environment; this design makes total sense about the hospital environmental dangers.” I have no idea what is going on in a unit when I am called for a consultation.” “I need avoid exposure to influenza because of my health issue. I do not know where influenza cases are located. This design would help me protect myself better.”</td>
</tr>
<tr>
<td>VIZ Usability</td>
<td>“The visual data informs me in a much better way than the sign on the door.” “I feel this interface works better for me than the flag in the EHR.” “I would like to be notified via my Bluetooth when I am approaching a high-risk zone. I do not have time to check the monitoring screens.” “This map is a useful quick snapshot of all the patients: we do not see them lamped together as a group.”</td>
</tr>
</tbody>
</table>
strongly associated with the Hawthorne effect14 in real time may help to improve the compliance with this hexidine bathing occurred. The availability of this information situational awareness (SA) needs of other stakeholders within the local epidemiological context may also address tion are readily available in EHRs and easy to capture unlike the teams and control practices. The study showed that the VIZ increased negatively affect the compliance with infection prevention 1.00) about the receipt of chlorhexidine bathing, which may positively affect containment practice. The use of antibiotics is a considerable contributing factor to antibiotic resistance.11–13 The data on antibiotics administration are readily available in EHRs and easy to capture unlike accurate diagnoses of infections. Representing antibiotic data within the local epidemiological context may also address situational awareness (SA) needs of other stakeholders groups (e.g., antibiotics stewardship, risk managers).

The findings revealed a lack of awareness (median = 1.00) about the receipt of chlorhexidine bathing, which may negatively affect the compliance with infection prevention and control practices. The study showed that the VIZ increased the teams’ SA about the locations where deficiency of chlorhexidine bathing occurred. The availability of this information in real time may help to improve the compliance with this procedure. The literature showed that the availability of feedback on compliance with infection prevention guidelines was strongly associated with the Hawthorne effect14–16 resulting in better compliance.

The essential finding of this study was the detected improvement in comprehension after introducing the VIZ. The participants assessed their baseline SA regarding the high risk for infection transmission situations as relatively low (median = 2.00). The post-test results showed a significant increase of the team’s SA of such situations (median = 4.50). Seventeen out of 19 participants correctly showed the high-risk locations using the VIZ. The participants (n = 19) demonstrated high baseline scores (median = 4.00) for “a need to know their occupational exposure to ARIs” (range 2–5). After introducing the VIZ, the post-test scores significantly increased (median = 5.00, range 3–5; p = 0.033). This finding informs that awareness about the occupational risk of exposure to ARIs may add value to infection surveillance. Tracking and sharing the information about individuals’ exposure to ARIs may leverage compliance with infection prevention and promote innovative solutions. The verbal post-test feedback provided additional insight about the need to know the extent to which HCPs are exposed to ARIs. The participants communicated that with the VIZ-based information they recognized “hospital environmental dangers” leading to a prospect to “protect myself better.” It appears that the VIZ supports comprehension about the risk of exposure to ARI organisms via the explicit image indicating hotspots or areas of high risk for exposure. Overall, the participants endorsed that the VIZ-based information representation makes the infection prevention and control meaningful to them.

As the study showed, the less experienced staff members perceived their awareness at a much higher level when comparing to the more experienced staff. This can be explained by level of experience. The experienced group of HCPs, those who worked more than 5 years in health care, showed high consistency in their EHR vs. VIZ responses in spite of their membership status. It would be beneficial to further explore the correlations between the prevalence of novice staff and ARI rates.

The usability measures showed that 70% of the respondents perceived the VIZ as superior to the EHR in its capability to provide meaningful and easily understandable information on infection transmission risks and exposures. During the verbal post-test discussion, the participants characterized the VIZ as the better means to alert the staff about the infection risks than the current flagging method for carriers of ARI organisms in the EHR. The infection preventionists expressed their strong interest in the VIZ by asserting that the unit-population surveillance information displayed with the VIZ “makes lots of sense” for understanding the unit epidemiology. At the same time, few responders rated the VIZ-based usability equivalent to the EHR. It appears that the VIZ training material was somewhat challenging for the less experienced staff, which may require either more time for training or a different method of information representation. The duration of training (less than 10 minutes), level of experience, and complexity of the information could affect these findings.

The verbal feedback provided additional insight on the VIZ value-added contribution to practice. The infection preventionists sparked a discussion on tactics, which are considered the best practice, in application to the data displayed with the VIZ. For example, they described the need to approach patients in a particular order: starting patient care with noninfected patients and then approaching the infected patients. Another tactic concerned patient’s arrangement on the unit. VIZ of the patients’ infection states and their locations in the unit would enable the assessment of patient safety based on their location for timely adjustments. Finally, many participants asserted that the VIZ motivated them to practice “better protection” from ARIs.

The study has some strength and limitations. The aim of the VIZ was to enable HCPs to rapidly gain a high level of

Impact of Tactical Biosurveillance Information Visualization on Awareness
understanding regarding the infection transmission “hotspots” and individuals at high risk of exposure to ARI organisms. The strength of this development was its focus on the user-centered design. The study pilot evidence showed that the proposed VIZ has the potential to address the infection prevention and control challenges. The early evaluation of the VIZ permitted the use of self-rating technique for measuring SA. The main advantages of the self-rating techniques include the ease of use, rapid collection of data, and low cost. Although the investigator managed to acquire relatively rich information with this pilot evaluation, the future study needs to explore objective measures of SA with rigorous research design (e.g., randomization, case-control). A future study can test the following objective measures: time on task and degree of errors (accuracy). Situational Awareness Global Assessment Technique may be employed as the effective means of assessing both individuals and teams in a human simulation environment. Situational Awareness Global Assessment Technique will help better understand to what extent the VIZ aids or undermines HCP ability to perform.

This evaluation also took advantage of the survey design, including its economy, rapidity of turn-around in data collection, and ability to identify attributes of a population from a small group of individuals.

It needs to be acknowledged that the pre- and poststudy design, convenience sampling method, small sample size (n = 19), and single setting generally work as limitations for generalization of the results. A more rigorous study design would have a control group with randomized assignment for EHR vs. graphical VIZ. Another problem relates to the use of the SA self-rating methodology. The important problem of this technique is that users, including both experts and nonexperts, are not always aware of what they do not know. Therefore, the disadvantage of SA self-rating technique includes its subjective nature due to the possible influence of perceived performance. Finally, it is important to emphasize that the quality of EHR data and sources of data for risk factors (antecedents of events) can affect the accuracy of risk stratification and mislead the decision-making and resource allocation. Therefore, identification of the high-quality predictive models is critical.

In 2013, the World Economic Forum announced that ARIs became a serious public health threat. The study demonstrated how the SA-oriented approach increases SA in the infection prevention domain. First, the study showed that the EHR data have great potential for solving population-level problems, coordinating the distributed team members’ communication, and motivating the staff members to enhance their performance. The latent EHR data buried in nonstructured text need to be abstracted for meaningful use. The integrated displays may present the aggregated, deidentified population data to HCPs who, otherwise, do not have access to all patients’ records.

Second, the proposed VIZ showed that it can enable the team to achieve much higher SA in seconds. The VIZ of mapped data minimized users’ time for acknowledging spatial characteristics or locations of the high risk for infection transmission areas (transmission hotspots); reduced users’ needs for processing of population-level biosurveillance information; enabled the recognition of subjects at high risk of exposure (super-spreaders); and eliminated users’ need for searching critical information in the latent EHR text and nontext documents for population-level decision-making about minimizing exposure events.

**SUMMARY**

The study provided pilot evidence that the VIZ enabled users to identify daily infection burden by presenting the population level data; to identify high-risk locations for infection transmission, regarded as high priority for infection control services; and to identify individuals who are at high risk for exposure to ARIs, regarded as high priority for infection prevention services. The analysis of the empirical data and the SA measurements helped understand the scope of information that supports perception and comprehension. It appears that SA-oriented technologies may be promising for decreasing ARI rates and promoting cost-effective innovations in infection prevention by enabling HCPs to systematically detect situations where risk of infection transmission and exposure to infections increase, and naturally forcing the health care teams to seek better prevention.

**REFERENCES**

9. Endsley MR, Bolstad CA, Jones DG, Riley JM: Situation awareness oriented design: from user’s cognitive requirements to creating effective supporting...


Military Emergency Medical Service System Assessment: Application of the National Park Service Needs Assessment and Program Audit to Objectively Evaluate the Military EMS System of Okinawa, Japan

LCDR Elliot M. Ross, MC USN*; MAJ Stephen A. Harper, MC USA*; LTC Cord Cunningham, MC USA*; CDR Benjamin D. Walrath, MC USN*; CDR Gerard DeMers, MC USN†; Col Chetan U. Kharod, USAF MC*

ABSTRACT  Introduction: As part of a Military Emergency Medical Services (EMS) system process improvement initiative, the authors sought to objectively evaluate the U.S. military EMS system for the island of Okinawa. They applied a program evaluation tool currently utilized by the U.S. National Park Service (NPS). Methods: A comprehensive needs assessment was conducted to evaluate the current Military EMS system in Okinawa, Japan. The NPS EMS Program Audit Worksheet was used to get an overall “score” of our assessment. After all the data had been collected, a joint committee of Military EMS physicians reviewed the findings and made formal recommendations. Results: From 2011 to 2014, U.S. military EMS on Okinawa averaged 1,345 ± 137 patient transports annually. An advanced life support (ALS) provider would have been dispatched on 558 EMS runs (38%) based on chief complaint in 2014 had they been available. Over 36,000 man-hours were expended during this period to provide National Registry Emergency Medical Technician (EMT)-accredited instruction to certify 141 Navy Corpsman as EMT Basics. The NPS EMS Program Audit Worksheet was used and the program scored a total of 31, suggesting the program is well planned and operating within standards. Conclusion: This evaluation of the Military EMS system on Okinawa using the NPS program assessment and audit worksheet demonstrates the NPS evaluation instruments may offer a useful assessment tool for the evaluation of Military EMS systems.

INTRODUCTION

A 2007 consensus statement from the U.S. Metropolitan Municipalities’ Emergency Medical Services (EMS) Medical Directors titled Evidence-Based Performance Measures for Emergency Medical Services Systems: A Model for Expanded EMS Benchmarking identified “crude measures of stakeholder satisfaction and other anecdotal measures” as the primary evaluation source of EMS system performance.¹ A lack of standardized evidence-based metrics limits local EMS system analysis.² Several agencies, including the Commission on the Accreditation of Ambulance Services and National Park Service (NPS) and North Carolina College of Emergency Physicians, have created objective criteria to evaluate EMS systems. Current Department of Defense (DoD) instructions and manuals on Fire and Emergency Services direct usage of applicable national and state certification process, but include very little detail on EMS. Service-specific regulations and instructions are more detailed in EMS requirements, but do not include standardized checklist-type criteria in evaluating programs. The DoD’s nascent Prehospital EMS Working Group has not developed any evaluation criteria or process either.

The island of Okinawa is a critical base of operations in the execution of the joint forces Pacific Strategy, hosting over 55,000 active duty service members of all four services and their dependents on 877 square miles. U.S. Naval Hospital Okinawa (USNHO) and Marine Corps Installations Pacific work in collaboration with the U.S. Air Force to provide EMS on the island. The EMS ambulances are independent of the fire command and control but are housed at the fire stations. There are seven total EMS ambulance stations on the island. All the EMS responders are active duty military and National Registry Emergency Medical Technician (NREMT) Certified as EMT Basics. A board-certified Emergency Medicine physician working out of USNHO, the lone-receiving hospital for all DoD patients, provides online medical direction. The medical director for the system is a board-certified Emergency Medicine physician assigned in writing by the hospital commanding officer and is responsible for all duties of the Military Treatment Facility Medical Director as described in the Navy instruction
for Pre-hospital Emergency Medical Services for Naval Facilities (BUMEDINST 6320.94).

As part of a system process improvement initiative, the authors sought to objectively evaluate the DoD EMS system for the island of Okinawa. The authors used a program evaluation tool currently used by the NPS. To provide NPS managers with direction and guidance in establishing and managing emergency medical care programs, NPS-51, the NPS Emergency Medical Services Guideline, was released in November 1984. In October 1985, the National Association of State EMS Directors passed a resolution unanimously supporting the adoption of NPS-51 in all states and offered its assistance in its implementation. In November of the same year, the National Council of State EMS Training Coordinators unanimously supported this resolution. A second version of NPS-51 was released in January 1991. The most recent version, RM-51, was updated in 2009.

METHODS

This is an institutional review board-exempt project used for process improvement purposes. Authorization to release deidentified data to the primary investigator was provided by the Commanding Officer of USNHO.

In December of 2014, a comprehensive needs assessment in accordance with Chapter 4 of RM-51 was conducted. Said needs assessment identifies and evaluates (1) available internal and external resources, (2) EMS workload, (3) requirements for training and certification, (4) transport capabilities and response times, (5) location and capability of the local area’s medical facilities, (5) fiscal resources, (6) EMS communications, and (7) special considerations.

Data and documents were reviewed from the Okinawa EMS system for the past 4 years to conduct this needs assessment (Table I). To assess the program’s workload, a retrospective review of the Okinawa EMS run reports from 2011 to 2014 including total number of EMS transports and number of prehospital deaths (defined as death with loss of vital signs in the prehospital setting) by year was conducted. In addition, the percentage of transports where advanced life support (ALS) level response was indicated based on chief complaint was also reported for 2014. The ALS chief complaint categories included neurologic, cardiac, respiratory, and toxic ingestion, where a prehospital ALS intervention might be required. The 2014 data abstraction was conducted by the Medical Director as part of the 100% prehospital chart review conducted for internal system continuous quality improvement. Additionally, all prehospital deaths are reviewed in depth by the Medical Director as part of continuous quality improvement. Descriptive statistics were also performed on the 2014 run data to further assess system demographics.

The NPS EMS Program Audit Worksheet was used to get an overall “score” of our assessment. The worksheet consists of 20 categories rated between 0 and 5 based on document condition and system implementation (Table II, Figure 1). A well-planned system operating within standards would score from 0 to 35. A score of 36–60 indicates a need for improvements to be in compliance with standards, and a score of 61–100 reveals the need of major improvement to be in compliance (Figure 2).

After all the data had been collected, a joint committee of EMS board-certified physicians with representatives from the U.S. Army, U.S. Navy, and U.S. Air Force reviewed the findings and made formal recommendations. These recommendations were then used by the island EMS Director and Hospital Commander to drive the recommended changes for the EMS system.

RESULTS

The aforementioned needs assessment clarified the available resources on the island to assist in mutual aid situations and clarified the processes for which these requests can be made. When dealing with the host nation (Japanese) EMS system, many differences in procedures were identified and subsequently addressed in a forum with community hospital and EMS representatives. For example, although Japan has a robust prehospital medicine training program as well as advanced education in EMSs, the skill sets/scope of practice for advanced medical technicians in Japan are limited.

<table>
<thead>
<tr>
<th>TABLE I. Summary of Metrics from National Park Service Emergency Medical Services Program Needs Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specific Metrics Included</td>
</tr>
<tr>
<td>• Total Number of Patients by Type: Medical (Cardiac, Seizure, Stroke, Diabetic, etc.) and Trauma (Fractures, Soft Tissue, etc.)</td>
</tr>
<tr>
<td>• Number of Patients That Received ALS; Also Valuable to Identify Cases Where ALS Was Not Available But Would Have Been Appropriate</td>
</tr>
<tr>
<td>• Number of Patients That Received BLS</td>
</tr>
<tr>
<td>• Number of Patients Treated and Released at the Scene Due to the Minor Nature of Their Injury/Illness</td>
</tr>
<tr>
<td>• Total Number of Extended Care Patients (More Than 1 Hour With the Patient)</td>
</tr>
<tr>
<td>• Fatalities by Type</td>
</tr>
<tr>
<td>• Total Number of Patients Transported</td>
</tr>
<tr>
<td>• Method of Transport</td>
</tr>
<tr>
<td>• Average Time to Patient Contact by a Basic EMS Provider</td>
</tr>
<tr>
<td>• Average Time From BLS Provider to ALS Provider</td>
</tr>
<tr>
<td>• Average Time From Patient Contact to Arrival at Hospital; May be Helpful to Separate Remote Evacuations From Roadside to Gain a Meaningful</td>
</tr>
</tbody>
</table>
compared with the skill sets laid out by the NREMT as minimum practice standards for paramedics. Therefore, the NREMT-certified paramedic has a broader scope of practice when providing ALS care when compared to the Japanese ALS providers. This difference impacts decisions on the use of host nation ALS ambulances, in addition to the inherent limitations of communicating medical emergency information in a foreign language.

From 2011 to 2014, 1,345 ± 137 patients were transported annually. By year, the total volume was 1,383 in 2011, 1,148 in 2012, and 1,383 in 2013, and 1,467 in 2014. Prehospital deaths by year were 10, 5, 3, and 5, respectively, from 2011 to 2014. An ALS provider would have been dispatched based on chief complaint on 558 EMS runs (38%) in 2014 had they been available. Additional data from 2014 indicate that 11.4% of calls were for pediatric (age <18) patients and 40.3% were low-acuity patients released at the scene. Additionally, 31.5% of calls were related to trauma. Given the geography of the island, 19.9% of the ALS calls were greater than 1-hour transports. Nontrauma-related deaths accounted for 56.5% (13/23) of the total deaths (Figures 3–5).

There were clear areas for improvement within the system’s requirements for training and certification. The Navy responders first complete “A” school at the DoD Medical Education and Training Campus (METC) then report for duty on the island. If assigned to the Emergency Department, they must complete an intensive 6-week accredited National EMT Certification course sponsored by the USNHO Staff Education and Training Department. Local policy requires successful NREMT Certification before serving as an EMS first responder (in contrast to U.S. Navy policy, which does not require corpsmen [HM] to obtain certification before graduation from METC). This process of local certification was noted to be extremely manpower intensive. From January 2011 to December 2014, 141 EMT students participated in the local EMT certification course. Each course required approximately 240 hours of instruction. A total of 33,840 student man-hours and 2,400 instructor man-hours were used to provide the local EMT course. For each

### TABLE II. Summary of System Review Categories From the National Park Service Program Audit Worksheet

<table>
<thead>
<tr>
<th>Emergency Medical Services Program Audit Worksheet</th>
<th>Response Categories</th>
<th>Postresponse Categories</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preresponse Categories</td>
<td>Park EMS Needs Assessment</td>
<td>Completed Within Past 10 Years</td>
</tr>
<tr>
<td>Park EMS Needs Assessment</td>
<td>Completed Within Past 10 Years</td>
<td>Signed Park EMS Plan in Place</td>
</tr>
<tr>
<td>Completed Within Past 10 Years</td>
<td>Signed Park EMS Plan in Place</td>
<td>Personnel Identified</td>
</tr>
<tr>
<td>Signed Park EMS Plan in Place</td>
<td>Personnel Identified</td>
<td>Qualifications and Authorizations Identified</td>
</tr>
<tr>
<td>Personnel Identified</td>
<td>Qualifications and Authorizations Identified</td>
<td>Partnerships in Place and Current Education and Training</td>
</tr>
<tr>
<td>Qualifications and Authorizations Identified</td>
<td>Partnerships in Place and Current Education and Training</td>
<td>Integrated in Park EMS Program</td>
</tr>
<tr>
<td>Partnerships in Place and Current Education and Training</td>
<td>Integrated in Park EMS Program</td>
<td>Communications Strategies</td>
</tr>
<tr>
<td>Education and Training</td>
<td>Communications Strategies</td>
<td>Equipment, Facilities, and Response Strategies Established</td>
</tr>
<tr>
<td>Program Communications Strategies</td>
<td>Law- and Policy-Related Issues National Scope of Practice Standards</td>
<td>Law- and Policy-Related Issues National Scope of Practice Standards</td>
</tr>
</tbody>
</table>

FIGURE 1. Scoring criteria for system implementation and document condition from the National Park Service Program Audit Worksheet.
course offering, an average of 40% (or, 100 hours) of classes were taught by staff physicians, and the remaining 60% (140 hours) by nationally certified EMTs. A total commitment of 960 staff physician man-hours and 1,440 EMT man-hours was required to provide instruction for the 141 students. Overall, first attempt pass rate for the NREMT examination during this time was 77% (108/141). Another 8 students passed the certification exam on their second or third attempt with overall pass rate of 82%. We noted this manpower burden could be avoided if Navy HMs arrived at USNHO with their National EMT Certification.

Transport times to the receiving hospital were in excess of 45 minutes from the two northern most camps (representing 20.7% of the total call volume). Four of the remaining six camps had transport times greater than 30 minutes. No issues were noted with EMS scene response times. Interfacility transports were conducted using advanced medical personnel such as critical care nurses or physicians allowing for critical care transport capabilities but this was only for hospital-to-hospital transfers, and not for scene response.

The locations of the EMS assets across the island were assessed to ensure appropriate response capabilities and memorandums of understanding were obtained with host nation EMS services for responses in the most remote areas of the island. In addition, with the relatively high percentage of ALS calls, the joint EMS committee evaluated the ideal numbers and locations for future augmentation of the EMS system with ALS ambulances.

Two separate DoD dispatch centers are in use for coordination of emergency response. Neither dispatch center had emergency medical dispatch-trained operators to provide pre-arrival instructions such as bystander cardiopulmonary resuscitation instruction nor the ability to triage calls based on chief complaints. Recommendations were made to integrate the dispatch systems and improve online medical instruction.

Finally, the NPS EMS Program Audit Worksheet was used to assess the U.S. military EMS program on Okinawa. The program scored a total of 31, suggesting the program is well planned and operating within standards. See Table II and Figures 1 and 2 for description of audit worksheet content.

![FIGURE 2. Overall program rating scores from the National Park Service Program Audit Worksheet.](image)

![FIGURE 3. ALS call volume as percentage of total calls determined by Dispatch Chief Complaint.](image)
DISCUSSION

We found that the DoD EMS system in Okinawa scored well on the NPS audit worksheet and that the program assessment tool identified multiple specific areas for improvement. This is the first time an objective scoring tool such as the NPS audit worksheet has been applied to a DoD EMS system. Given the fluid nature and high turnover within the DoD medical department, having objective evaluation tools for EMS systems run by inexperienced EMS Medical directors is of significant value.

Other studies looking at EMS system evaluation point out a significant shortfall in many DoD EMS systems in the area of data collection and analysis. Comparative studies between EMS systems are difficult to perform due to the diverse nature of the systems. However, a robust outcomes-evaluation mechanism, as suggested in the “EMS Agenda for the Future,” is a necessity to provide feedback and focus process improvement efforts. The current military EMS system in Okinawa uses a paper-based emergency treatment record and has no searchable data repository. This has been a shortfall identified in military combat prehospital medicine as well and has only recently been addressed by forward thinking providers and leaders.

Recommendations from the joint committee on prehospital care reviewing this data suggested changing the DoD EMS system in Okinawa from a “basic life support (BLS) only” to a hybrid system using 1/3 ALS providers and 2/3 BLS providers strategically placed across the island to maximize ALS coverage. Thus, two of the six Navy EMS stations should be ALS. They also recommend the U.S. Air Base Kadena EMS station be upgraded to ALS level of care. These recommendations were formally presented to the Commanding Officer of USNHO and Commanding Officer for Marine Corps Installations Pacific.

Although the NPS tools were of value, many of the components are not present or relevant to military EMS systems. A modification of items contained in the NPS audit worksheet to address similar DoD-specific concerns will increase the quality of the assessment tool. Although the

FIGURE 4. Okinawa Military EMS call volume by station location.

FIGURE 5. Map of Okinawa Department of Defense operated ambulance station locations.
DoD is single federal entity, each service operates its own EMS system, using different protocols and different data collection mechanisms making comparisons, even with outcome data, very difficult. A universal program assessment tool may assist in moving these disparate DoD systems closer to a single accepted standard much like the NPS.

Our study has several limitations. First, this is primarily a process improvement initiative and not a formalized research process and therefore has the expected limitations of a retrospective review. Second, the assessment tools are not outcomes based, but are performance based, which does not allow for a full functional evaluation of current system processes. Finally, this is an assessment of a single EMS system and the results are specific to this system.12–14

CONCLUSION

In this evaluation of the DoD EMS system on Okinawa, we found that the NPS evaluation instrument may offer a useful assessment tool for the evaluation of Military EMS systems. Without standardized criteria to evaluate EMS systems, it is difficult to provide valid recommendations for medical leaders. The DoD should further develop joint service EMS system evaluation criteria across the enterprise using the NPS tool as a model.

ACKNOWLEDGMENTS

We would like to thank the National Park Service, Ken Phillips, Barry D. Smith, and LTC Will Smith for their significant contributions to this work. We would also like to thank the Prehospital Research and Innovation in Military and Expeditionary Environments (PRIME2) Research Group.

REFERENCES

ABSTRACT  Background: Navy Hospital Corpsmen (HMs) are the Navy equivalent to Emergency Medical Technicians (EMTs) both in-garrison and on the battlefield. In 2000, the Emergency Medical Services (EMS) Education Agenda for the Future highlighted the need for a single certification agency to provide consistent evaluation of entry level competence for each nationally recognized EMS provider level. Administered by the National Registry of EMTs (NREMT), National EMT Certification is currently utilized by 46 states, the District of Columbia, four territories, and six federal organizations as part of their processes for granting licensure. Unlike the Air Force (USAF) and Army (USA), the Navy (USN) does not require National EMT Certification to perform the duties equivalent to a civilian EMT. Our objective is to describe the number of USN HMs, USAF medics, and USA combat medics who have obtained National EMT Certification from 2007 through 2014. Methods: Results from all USN HMs, USAF medics, and USA combat medics who tested between January 1, 2007 and December 31, 2014 were queried from the NREMT database. Descriptive statistics were calculated based on a retrospective review of prospectively collected testing data. Results: During the study period, 89,136 Military Service Members received their EMT certification from the NREMT. The breakdown of the total and percent of total is; USA Combat Medics (n = 69,761; 78.3%), USAF Medics (n = 16,195; 18.1%), and USN HMs (n = 3,180; 3.6%). Approximately 4,000 HMs graduate yearly from the Department of Defense Medical Education and Training Campus at Fort Sam Houston, Texas and 253 HMs obtained certification in 2014. Conclusions: About 6.3% (253/4,000) HMs obtained National EMT Certification in 2014, which is a nationally recognized standard for entry-level competence utilized by civilian EMTs and other branches of the military. More information about those HMs that obtain certification may help Commanders maximize the number of HMs obtaining certification. Mandating National EMT Certification for HMs graduating from initial entry training would increase the numbers obtaining certification and bring them in line with USA, USAF, and national movement toward requiring certification for medical providers at all levels.

INTRODUCTION  Sailors selected for training as a Navy Hospital Corpsman (HM) report to Navy Medical Education and Training Command on Joint Base San Antonio-Fort Sam Houston (JBSA-FSH). Currently, most members of the U.S. Army (USA) and U.S. Air Force (USAF) complete their initial medical training at JBSA-FSH as well. U.S. Navy (USN) and USAF students actually share classroom didactics and instruction, whereas USA students receive a similar program of instruction but taught separately from the other two services. The course of instruction meets National Registry of Emergency Medical Technicians (NREMT) requirements for the National Emergency Medical Technician (EMT) certification examination.1

In 2000, the Emergency Medical Services (EMS) Education Agenda for the Future highlighted the need for a single certification agency to provide consistent evaluation of entry-level competence for each of the four nationally recognized EMS provider levels.2 National EMS Certification encompasses all four nationally recognized EMS provider levels, Emergency Medical Responder, EMT, Advanced EMT, and Paramedic certifications.3 In 2006, the Institute of Medicine formally concurred with this recommendation and further recommended that “states accept national certification as a prerequisite for state licensure and local credentialing of EMS providers.”4 Administered by NREMT, national EMS certification is currently utilized at one or more levels by 46 states, the District of Columbia, four territories, and six federal organizations as part of their processes for granting licensure to practice.5 It is worth noting that while some states describe the process at their level as certification, in reality it is more appropriate to be called licensure. USA and USAF require their medical personnel to obtain National EMT Certification to perform the duties equivalent to a civilian EMT. In order to obtain National EMT Certification, a candidate must successfully complete the computer adaptive cognitive exam and a psychomotor exam comprised of 5 to 10 state-specific skills. The USA Combat Medics and USAF Medics must obtain National EMT Certification in order to graduate from their Initial Entry Training medical programs. Obtaining the
National EMS certification helps to validate and ensure a competent workforce.2

Use of National EMS Certification is not limited to initial entry-level medical technicians in these three services. The USA recognizes the importance of such certification at other levels of care as well. In addition to certification of basic medical personnel to the EMT level, USA flight medics are required to obtain National Paramedic Certification.6 The U.S. Coast Guard, while not requiring National EMT Certification for initial training, does require certification for advancement to noncommissioned officer ranks.7 Although the USN does not mandate National EMS Certification at any level for enlisted military medics, they do appear to recognize the importance of certification in general as evidenced by the existence of an established funding source to pay for National EMS Certification testing through the Department of Navy Credentialing Opportunities On-line (DON COOL) program.8 Additionally, the USN requires EMS providers on USN installations to maintain National EMS Certification to treat and transport patients in a prehospital setting.9 Nevertheless, little is known about the actual number of HMs who obtain National EMT Certification. The objective of this study is to describe the number of USA, USAF, and USN personnel who obtained National EMT Certification between 2007 and 2014.

METHODS
In 2014, the NREMT database contained records for over 300,000 EMS professionals who possessed current National EMS Certification.10 Additionally, the NREMT collects and maintains data and demographics on all candidates who take a cognitive examination for National EMS Certification. These data are primarily collected and used in quality improvement activities. In 2014, over 130,000 computer-based examinations were administered by the NREMT.10 The NREMT database was queried for all USA, USAF, and USN candidates who earned National EMT Certification between January 1, 2007 and December 31, 2014. Descriptive statistics were calculated. All analyses were performed using STATA/IC 12.0 (StataCorp, College Station, Texas).

RESULTS
During the study period, 89,136 Service Members from the USA, USAF, and USN obtained National EMT Certification. The majority were USA Combat Medics (n = 69,761; 78.3%), followed by USAF Medics (n = 16,195; 18.1%), and then USN HMs (n = 3,180; 3.6%). Table I and Figure 1 display the number of personnel who obtained National EMT Certification for each year of the study period by branch of service. There was a step-wise decrease in the number of USN personnel that gained National EMT Certification between 2007 and 2014. In 2007, 758 USN personnel obtained National EMT Certification and this figure dropped to 253 by 2014. Approximately 4,000 HMs graduate on a yearly basis from the Department of Defense Medical Education and Training Campus at Fort Sam Houston, Texas.11 This means about 6.3% (253/4,000) of new HMs obtained their National EMT Certification in 2014. Since the USA and USAF require National EMT Certification, 100% of their graduates obtain the certification.

DISCUSSION
This study is the first to quantify the number of USA, USAF, and USN personnel who obtain National EMT Certification. A total of 69,761 USA Combat Medics, 16,195 USAF Medics, and 3,180 USN HMs obtained National EMT Certification from 2007 to 2014. Although National EMT Certification is requirement for medics in the USA and USAF, the USN does not mandate HM certification. Nevertheless, the USN as stated above has recognized the value of National EMT Certification. Despite available funding to support HMs who wish to obtain this certification, only 6.3% of HMs obtained National EMT Certification in 2014. The authors have heard many anecdotal reasons that the USN does not require EMT certification. While that is a very interesting discussion, finding written guidance against certification is difficult. In fact, many Navy publications either recommend or require EMT certification depending on unit or duty assignment.9,12

This study only looked at obtaining National EMT Certification and did not review initial and overall pass rates for

<table>
<thead>
<tr>
<th>Year</th>
<th>USA (Row %)</th>
<th>USAF (Row %)</th>
<th>USN (Row %)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>2007</td>
<td>10,676 (79.1)</td>
<td>2,071 (15.3)</td>
<td>758 (5.6)</td>
<td>13,505</td>
</tr>
<tr>
<td>2008</td>
<td>10,201 (82.6)</td>
<td>1,664 (13.5)</td>
<td>481 (3.9)</td>
<td>12,346</td>
</tr>
<tr>
<td>2009</td>
<td>10,966 (81.9)</td>
<td>2,008 (15.0)</td>
<td>412 (3.1)</td>
<td>13,386</td>
</tr>
<tr>
<td>2010</td>
<td>9,313 (74.5)</td>
<td>2,721 (21.8)</td>
<td>458 (3.7)</td>
<td>12,492</td>
</tr>
<tr>
<td>2011</td>
<td>8,257 (78.5)</td>
<td>1,936 (18.4)</td>
<td>323 (3.1)</td>
<td>10,516</td>
</tr>
<tr>
<td>2012</td>
<td>6,957 (74.6)</td>
<td>2,098 (22.5)</td>
<td>266 (2.8)</td>
<td>9,321</td>
</tr>
<tr>
<td>2013</td>
<td>7,478 (77.5)</td>
<td>1,943 (20.1)</td>
<td>229 (2.4)</td>
<td>9,650</td>
</tr>
<tr>
<td>2014</td>
<td>5,913 (74.7)</td>
<td>1,754 (22.1)</td>
<td>253 (3.2)</td>
<td>7,920</td>
</tr>
<tr>
<td>Total</td>
<td>69,761 (78.3%)</td>
<td>16,195 (18.2%)</td>
<td>3,180 (3.6%)</td>
<td>89,136</td>
</tr>
</tbody>
</table>
the certification exam. Students in the USA and USAF that did not pass the certification exam cannot graduate from their initial entry training and they are either diverted to another military occupation or separated from the service. Service Members typically receive three attempts to pass the test during their initial entry training. Since the USN does not require National EMT Certification, if an HM did not pass the test, they retain their military occupation but may receive a different duty assignment based on if their job required EMT certification. For instance, for an HM to work on an ambulance at OCONUS locations requires certification. Further research is needed to assess NREMT certification initial and overall pass rates of the respective services’ medical providers. This information would provide valuable insight as to the quality of military medic training provided by the Department of Defense initial medical education for enlisted personnel.

As with most retrospective studies, this study has several limitations. First, we assumed that the personnel that obtained EMT certification were actually medical providers who obtained the certification during their Initial Entry Training and not some other military occupation. The exact number of military personnel graduating from their Initial Entry Training medical programs is not known. Given that the number of HMs graduating in 2014 was approximately 4,000, it was estimated that 6.3% of graduating HMs obtained their EMT certification. For USA and USAF graduates, it can be assumed that 100% of those graduating from their Initial Entry Training medical programs obtained EMT certification as that is a graduation requirement to proceed on to their military occupation.

A further limitation of this study related to the Navy HMs is that the timing of when individuals took the test is not known. Presumably those obtaining certification did so after reporting to their first duty assignment. But even if that is true, the amount of time between graduating from medical training and taking the test is not known. The HMs attempting EMT certification several months after doing their training may be at a disadvantage and have lower pass rates resulting in fewer numbers obtaining certification. If the USN had a better understanding of when their HMs take the test and their overall pass rates, the USN could adopt quality improvement initiative to try and help those Sailors that are seeking certification.

CONCLUSIONS
A very low proportion of graduating HMs obtain National EMT Certification. More information about those who do obtain their certification may provide insight that Commands could use to maximize HMs obtaining certification. Mandating National EMT Certification would better standardize entry-level enlisted medic requirements across the Department of Defense. Such a mandate would also bring the entry-level military EMT on par with the national civilian goals set forth in EMS Agenda for the Future and the overall national movement toward requiring medical providers at all levels to be certified.

REFERENCES
ABSTRACT  Objective: To describe the use of template-based screening for risk of infectious disease exposure of patients presenting to primary care medical facilities during the 2014 West African Ebola virus outbreak. Methods: The Military Health System implemented an Ebola risk-screening tool in primary care settings in order to create early notifications and early responses to potentially infected persons. Three time-sensitive, evidence-based screening questions were developed and posted to Tri-Service Workflow (TSWF) AHLTA templates in conjunction with appropriate training. Data were collected in January 2015, to assess the adoption of the TSWF-based Ebola risk-screening tool. Results: Among encounters documented using TSWF templates, 41% of all encounters showed use of the TSWF-based Ebola risk-screening questions by the fourth day. The screening rate increased over the next 3 weeks, and reached a plateau at approximately 50%. Conclusions: This report demonstrates the MHS capability to deploy a standardized, globally applicable decision support aid that could be seen the same day by all primary care clinics across the military health direct care system, potentially improving rapid compliance with screening directives.

INTRODUCTION

During 2014, a large Ebola virus outbreak occurred in West Africa that quickly overwhelmed local resources and eventually resulted in an international response coordinated by the World Health Organization.1 Health systems initiated screening systems for patients who had traveled to affected areas.2 When an Ebola Virus Disease (EVD) case emerged within the United States in October of 2014, efforts for handling patients including screening patients by assessing risk of recent exposure to EVD were redoubled; EVD-related issuances by the U.S. Centers for Disease Control and Prevention peaked in October 2014.3 The U.S. Military Health System (MHS) responded in a number of ways, including aspects of the response within the direct care primary care arena.

Unrelated to the EVD outbreak itself, a standardized, team-based care delivery and documentation methodology had been generally adopted in primary care settings across the MHS starting in 2011. This methodology, called Tri-Service Workflow (TSWF), centers around standardized templates called Alternate Input Method (AIM) forms that function within AHLTA, DoD’s electronic health record (EHR). These AIM forms and their associated team-based workflows have become a standardized approach for delivering and documenting primary care across the MHS. The TSWF AIM forms feature standardized layouts, content, and prompts, all intended to support evidence-based, standards-driven, comprehensive care in a “patient-centered medical home” paradigm. Initiated in 2011, the TSWF templates became adopted across most primary care clinics and platforms of the Army, Navy, and Air Force, so that by late 2014, these templates were utilized in close to 80% of all primary care encounters (unpublished data). The extent of the use and standardized nature of these AIM forms make them well suited for rapid dissemination of system-wide tools and/or policies, particularly those relevant to primary care. For the first time since the development of TSWF, the MHS was able to capitalize on these TSWF features to facilitate rapid system-wide implementation of standardized screening protocols urgently advocated by national authorities. This article, then, describes the role of TSWF template management in supporting the implementation of MHS’ strategy for EVD exposure risk assessment strategy in primary care.

EHR templates in general offer excellent tools for dissemination of standardized processes, for prompting actual actions, for documenting actions and results, and eventual data collection to assess program effectiveness. Templates provide clinical decision aids in support of standard of care. Primary care clinicians have found that template use during patient encounters enables decision making, speeds information gathering, and provides reminders of important issues that need to be addressed with the patient during the encounter.4 Templates can also improve clinical documentation and prompt physicians to order procedures and tests in accordance to particular guidelines.5 This approach minimizes the burden of searching for or memorizing guidelines and instead embeds these guidelines into the electronic medical record system for easy, system-wide use.
The ease and speed at which templates can be modified makes them a valuable tool for disease outbreak or disaster response and data collection, particularly when standardized templates are widely used. Because changes to these forms do not require changes to the underlying system, alterations can be made quickly and inexpensively. The literature addressing how the use of templates can facilitate adoption of clinical processes and outcomes is mixed and likely represents the complexity of issues involved. However, there is an evidence that EHRs in general and template functionality in particular can have a positive impact on clinical actions and outcomes.\textsuperscript{6–8}

To summarize the relevant MHS response to the 2014, West African Ebola virus outbreak, an Ebola Working Group was initiated as an “ad hoc” entity to develop recommendations for the leadership. This group developed, and MHS implemented, an EVD risk-screening process for primary care settings to support early notification and early response to potentially infected persons. For primary care settings, three questions were developed based on recommendations from the Centers for Disease Control. The questions were expected to be asked of all patients presenting for care or their caregivers. On November 3, 2014, these evidence-based questions were inserted into the TSWF templates in the AHLTA EHR system; the updated template was immediately available all clinical staff with properly loaded templates. The three self-explanatory questions were as follows; their visual appearance in AHLTA is reproduced in Figure 1:

1. In the previous 21 days, has the individual resided or traveled to any of the following countries in West Africa: Liberia, Sierra Leone, Guinea, Nigeria, or any region where EVD transmission is active? (If YES, a text box is provided for Ebola Travel History details).
2. Has the patient visited anyone who might have Ebola or involved with body fluids, surroundings or remains of an Ebola patient during the past 21 days? (If YES, a text box is provided for Ebola Exposure History details).
3. Has the patient had a fever OR nausea/vomiting, diarrhea, abdominal/stomach pain, joint and muscle ache, severe headache, new skin rash, or unexplained bruising in the past 21 days? (If YES, a text box is provided for Ebola Symptom History details).

AHLTA users obtain routine access the TSWF AIM forms by loading them into their personal user’s list of favorite templates. The forms need to be loaded in a specific manner so that the link to the enterprise version is maintained, otherwise the user will have access to a “local” version that does not get automatically updated. When primary care users first obtain AHLTA accounts, these templates are correctly loaded by onsite clinical information systems personnel and users are instructed how to do this correctly. It is possible, however, for users perhaps unknowingly to save the template as a “local” template, breaking the link that provides automated updates. Periodic checks and record reviews can help identify users with an old template. There are no established data that describe the proportion of users with outdated templates, but anecdotal observation indicates that it is small, probably less than 5% of users.

In addition to the 3-item questionnaire, a banner was placed on the initial load screen of the form which clearly alerted users of the Ebola-screening policy. Answers to the screening questions emitted to the Subjective/Objective (S/O) section of the note, providing documentation that screening had been accomplished and the results. As is the case for other screening protocols, providers could easily review and access the documentation. MHS and Service-specific issuances were disseminated separately describing procedures to follow should screening indicate higher risk. Treatment facilities developed protocols appropriate for their settings. A description of these protocols is beyond the scope of this article.

METHODS
Clinical encounter data were extracted for purposes of evaluating the EVD risk-screening program effectiveness. Data for each primary care clinic encounter from October 20, 2014 to December 15, 2014, throughout the MHS direct care system were abstracted from the Comprehensive Ambulatory Patient

![FIGURE 1. Screenshot of TSWF-based EVD risk-screening tool. Reproduced for visual representation only; see body of text for the wording of the screening questions.](https://example.com/figure1.png)
Encounter Record data repository. Data from October 20 through November 2 were used to establish a baseline for the software terms representing the three EVD-screening questions.

For each encounter, information was compiled to indicate whether a TSWF AIM form was used, if any of the TSWF-based EVD risk-screening questions were documented, the clinic Medical Expense and Performance Reporting System (MEPRS) code, the Medical Treatment Facility (MTF) Defense Medical Information System code, the date of the encounter, and the MTF’s governing Service including National Capital Region Medical Directorate (NCRMD) considered as a separate Service. Comprehensive Ambulatory Patient Encounter Record data are reported the same way from all MHS fixed facility direct care facilities and are considered to reflect actual or “true” screening documentation; hence no systematic reporting bias is expected. Only primary care specialties were assessed because the TSWF AIM forms are designed for use in these clinical areas. Data were compiled into Microsoft Excel compatible files; Excel was used to perform simple frequencies and rate calculations. Rates of TSWF AIM form use and TSWF EVD risk assessment documentation among TSWF AIM form users were calculated.

RESULTS
Among 341 MTFs that reported primary care encounters, 328 (96%) had encounters that utilized TSWF AIM forms for encounter documentation (Fig. 2). 317 (93%) had encounters utilizing TSWF-based EVD risk assessment documentation during this time period. Most of those with 0% TSWF-based EVD risk assessment screening were small isolated clinics with low encounter counts (data not shown). The software terms utilized for the EVD risk-screening questions were rarely used before November 3, 2014 (Fig. 3).

During the period of November 3, 2014 to December 15, 2014, 76.5% of direct care primary care encounters utilized TSWF AIM forms for documentation (Table I). Among encounters documented using TSWF templates, 40.9% of all encounters showed use of the TSWF-based EVD risk-screening questions by day 4 of implementation (Fig. 3). The screening rate increased dramatically the first 4 days, then the increased slowed somewhat over the next 3 weeks and reached a plateau at approximately 50% by the end of the first month of implementation. Interestingly, there was a significantly lower rate of screening during weekends, demonstrated in Figure 3.

TSWF AIM form use rates varied by Service and by primary care specialty. Navy clinics demonstrated the highest use rates at 79.1%, closely followed by the Air Force at 78.2%, the Army at 75.4%, and NCRMD was significantly lower at 54.5% (Table I). Family Medicine clinics had the highest TSWF AIM form use at 82.1% of encounters, with General Internal Medicine clinics the lowest at 58.2% (Table II).

TSWF-based EVD risk-screening rates among TSWF AIM form users varied by Service (Table I). The Army’s 61.6% was highest, followed by Air Force, NCRMD, and Navy.

TSWF-based EVD risk-screening rates among TSWF AIM form users also varied across primary care specialties (Table II). Clinics categorized as “Other primary care” utilized the tool the most at 59.5% of TSWF-documented encounters, while pediatrics had the lowest utilization rates at 16.7%.

DISCUSSION
TSWF-based EVD risk screening was widely and rapidly adopted in primary care settings across the MHS. The addition of these screening questions to the TSWF AIM forms was intended as an adjunct to Service and MTF policy implementation, as a visual reminder inserted into the daily workflow of primary care clinical documentation. It was not possible to train each primary care AHLTA user on use of the screening tool during this time period, nor was specific training apparently necessary for widespread adoption. The extent local MTF-based training specific to adoption of the TSWF-based tool is not known.

Adoption rates between military Services varied for reasons that are not entirely clear. During the time period described in this study, EVD risk assessment was a high visibility national and international issue. Risk-screening recommendations from national and international authorities were prominently distributed through many channels including national new media. Although policies for EVD risk screening and assessment were disseminated throughout the West African Ebola outbreak of 2014, the use of the TSWF-based EVD risk assessment tools was not specifically mandated by MHS or the Services. Some MTFs instituted a paper-based risk assessment method (personal observation) or other local tracking mechanisms that might have been intended to accommodate local operational or State reporting requirements. It is
beyond the scope of this report to describe the various risk assessment methodologies in place during this time period.

Adoption rates of the TSWF-based EVD risk-screening tool varied across primary care specialties (Table II). Explanations for this finding are speculative. It would stand to reason that adoption of TSWF EVD risk assessment would be highest in specialties with the highest TSWF utilization; however, this was not observed. Pediatrics had the highest utilization of TSWF forms (95.6%) occurs yet by far the lowest use of TSWF EVD risk screening (15.6%). The low rate of EVD risk screening in Pediatrics could be due to a perceived low pretest probability in this age group on the part of Pediatric staff. It is not known whether the lower EVD risk-screening documentation rates in Pediatrics and Flight/Undersea Medicine is due to a conscious determination of lower risk for these patients or a gap in actual screening of patients in these settings.

The reason for the drop in TSWF screening rates observed during weekends is unknown and a discussion of causes is speculative. Most but not all MHS primary care clinics are closed on weekends. It is possible that a higher proportion of staff without TSWF training work on weekends, weekend staff lacked familiarity with the guidelines, or different screening methods were used such as paper-based screening.

Among TSWF form users, the uptake of using the TSWF EVD risk-screening method was prompt and sustained over this time period. A saturation effect was reached within a very few days. Reasons for nonuse of the TSWF-based risk-screening tool in these primary care settings is not known; however, as observed above, use of paper EVD risk screening and travel history forms was observed at some MTFs. It is not known if a paper-based method or other non-TSWF methods were systematically implemented at certain MTFs or across certain commands. Although a paper-based screening method has the advantage of local flexibility, it also has the liability of the difficulty of collecting accurate implementation data. Entering screening results into a structured data

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Encounter Count</td>
<td>TSWF-Using Encounters (%)</td>
<td>TSWF-Using Encounters With TSWF-Based Ebola Screening (%)</td>
</tr>
<tr>
<td>NCRMD</td>
<td>40,358</td>
<td>24,348 (60.3)</td>
</tr>
<tr>
<td>Navy</td>
<td>256,068</td>
<td>202,622 (79.1)</td>
</tr>
<tr>
<td>AF</td>
<td>335,159</td>
<td>262,167 (78.2)</td>
</tr>
<tr>
<td>Army</td>
<td>539,314</td>
<td>406,805 (75.4)</td>
</tr>
<tr>
<td>Totals</td>
<td>1,170,899</td>
<td>895,942 (76.5)</td>
</tr>
</tbody>
</table>

*Probable under count because a technical problem identifying encounters that was TSWF-documented from 4 AF MTFs.*

source like the EHR allows for metric generation as soon as the data flows allow.

These data are not to be construed as lack of EVD risk screening for encounters not documented using TSWF methods. TSWF form users and nonusers alike may have used other methods for screening such as a paper form. The presentation of these data is not intended to demonstrate the full extent of EVD risk screening in the MHS, only the extent of TSWF-based EVD risk screening using the structured data entry method provided in the EHR template.

A possible source of data contamination includes mixed use of software terms. The EVD risk-screening questions were essentially prepopulated text fields that could potentially have been used for documenting other aspects of care. Before November 3, only very small counts of these text fields were detected. It is possible but not likely that use of these specific text fields could have increased independent of the EVD assessment. A separate analysis of MTF-level use of these specific text fields before November 3, 2014, showed very low counts, zero, or close to zero with a maximum of 0.25% of all encounters on a daily basis (data not shown other than Fig. 3).

One potential source of low adoption of TSWF-based EVD risk screening is the way the TSWF AIM forms are loaded into a user’s template favorites list in AHLTA. It is not possible to calculate the number of users who have saved “local” versions of the templates. Anecdotal observations indicate that a small minority of users do this.

TSWF AIM forms are a specific form of template management used in DoD, and these forms offer a unique way to study the effects of template management. These forms were intended as enterprise-wide solutions, and while not specifically mandated across the MHS, these forms were in wide use across the MHS at the time of this study as noted above. The design of TSWF AIM forms addresses several template management challenges noted in the literature, including speed of dissemination, EHR compatibility, and a lack of template uniformity. If implemented without addressing these challenges, template management suffers from major limitations. Technical solutions must support template placement across a health care delivery enterprise without burden placed on the part of the providers. Like-wise, the template must be compatible with the EHR. TSWF meets each of these criteria, save for standardized use. In addition, governance supports dissemination and use of standard templates that meet organization goals and that are used across the health care delivery enterprise.

TSWF template management still faces key challenges. Standardized templates require consensus to be achieved among medical departments of the military Services, a process that could be time consuming. These kinds of questionnaires require maintenance to ensure that they are evidence-based and temporally relevant. Active governance is required to ensure that proper questions are asked. These questions must be carefully aligned to the organization’s goals to avoid overburdening workers.

CONCLUSIONS
The major contribution demonstrated here is not the absolute rate of TSWF AIM form use or use of TSWF-based EVD risk-screening documentation, but rather the demonstration of a viable, widely and rapidly adopted clinical decision support tool on a massive scale. The utility of this dissemination method is limited at the present time to conditions or circumstances that are applicable across the entire MHS: all or nothing.

Adoption of the TSWF-based EVD risk-screening tool would likely have been higher had its use been formally or officially directed. Template management is an easy-to-develop and rapidly implementable tool for primary care teams to collect, document, and utilize patient data immediately in clinical decision making, in this case during a disease outbreak of international significance.

While this report shares the results of a program effectiveness assessment, the findings open doors to potential future research. A comparison could be undertaken of the success of appropriate policy implementation by routine communication channels versus dissemination via template adjustments. An assessment of local variation of practice to better understand the differences in MTF-level adoption, the influence of demographic strata, and the reason for the 50% plateau of sustained rate of TSWF-based EVD risk screening. These questions would help policy makers determine effective methods for disseminating those clinical action policies that are
appropriate to EHR documentation. Moving forward, template management will continue to play a critical role in assisting the MHS in meeting its needs for responsive, system-wide clinical decision support.

REFERENCES

Variation in Postinjury Antibiotic Prophylaxis Patterns Over Five Years in a Combat Zone

Col Bradley A. Lloyd, USAF MC†; COL Clinton K. Murray, MC USA‡; William Bradley, MS†‡; Faraz Shaikh, MS†‡; Deepak Aggarwal, MSE, MSPH†‡; M. Leigh Carson, MS†‡; David R. Tribble, MD, DrPH†

ABSTRACT In 2008, a clinical practice guideline (CPG) was developed for the prevention of infections among combat casualties and was later revised in 2011. We evaluated utilization of antimicrobials within 48 hours following injury in the combat zone over a 5-year period (June 2009 through May 2014) with regard to number of regimens, type of antimicrobial, and adherence to the 2011 CPG. The study population consisted of 5,196 wounded military personnel. Open fractures and skin and soft-tissue injuries were the most frequent injuries. Closed injuries had the highest overall compliance (83%), whereas open fractures and maxillofacial injuries had significant improvement in compliance from 2009–2010 (34 and 50%, respectively) to 2013–2014 (73 and 76%, respectively; p < 0.05). Part of the improvement with open fractures was a significant reduction of expanded Gram-negative coverage (61% received it in 2009–2010 compared to 7% in 2013–2014; p < 0.001). Use of Gram-negative coverage with maxillofacial injuries also significantly declined (37–12%; p = 0.001). Being injured during 2011–2014 compared to 2009–2010 was associated with CPG compliance (p < 0.001), while high injury severity scores (≥ 10) and admission to the intensive care unit in Germany were associated with noncompliance (p < 0.001). Our analysis demonstrates an increasing trend toward CPG compliance with significant reduction of expanded Gram-negative coverage.

INTRODUCTION

The administration of antimicrobials as part of the immediate care of trauma patients is a standard practice to prevent subsequent infections. In general, the majority of clinical practice guidelines (CPGs) related to infectious complications resulting from traumatic injuries have been developed in the civilian setting.1–8 Furthermore, review of the effectiveness of these guidelines has also been focused on civilian populations.2,9–11 As combat trauma presents with differences in injury mechanism, severity of injury, and timing of treatment, the Department of Defense (DoD) Joint Trauma System (JTS) convened an expert consensus panel to adapt these civilian guidelines for the military setting. The first CPG for the prevention of infections associated with combat-related injuries was published in March 2008.12 The CPG outlined appropriate postinjury antimicrobial prophylaxis choices as determined by injury pattern and specifically recommended cefazolin or clindamycin (i.e., Gram-positive coverage) for open fractures (including Type IIb and IIIc) and maxillofacial injuries. Shortly thereafter, the JTS released internal guidance in March 2010 based on civilian guidelines,1,4 which was contradictory to the 2008 published CPG, as it recommended expanded Gram-negative coverage (levofloxacin, aminoglycosides, or other Gram-negative coverage) in addition to the Gram-positive coverage for contaminated fractures or certain maxillofacial injuries.

In an attempt to standardize recommendations and incorporate evidence-based findings related to dosing regimens, a revised CPG was published in 2011 by an expert panel of DoD trauma and infectious disease physicians,13 which was later validated by a corresponding JTS document in 2012.14 In this revised CPG, addition of Gram-negative coverage for postinjury prophylaxis in the setting of open fractures and maxillofacial injuries was not recommended and the number of choices for antibiotic prophylaxis with penetrating abdominal injuries was significantly curtailed (i.e., only cefazolin/metronidazole or ertapenem were recommended as appropriate prophylactic antibiotic choices).15 The goal of the expert consensus panel was to significantly decrease broad-spectrum antibiotics use in the postinjury antibiotic prophylaxis time frame because of concern that use of overly broad antibiotics for prophylaxis may increase the risk of multidrug-resistant organism acquisition and infection, unnecessarily expose patients to potential toxicities (e.g., aminoglycoside-induced kidney injury), and possibly lead to less emphasis on quality surgical debridement.15,16 In two prior publications, we evaluated antimicrobial prophylaxis compliance with the 2008
CPC over a 6-month and 1-year period among injured United States (U.S.) military personnel medically evacuated to Landstuhl Regional Medical Center (LRMC).\textsuperscript{17,18} Although the 6-month analysis was comprehensive at the time, methodological limitations restricted applicability of the results. Therefore, a revised methodology based on Abbreviated Injury Scale (AIS) codes for injury categorization, rather than International Classification of Disease (9th edition) diagnostic codes, and improved antibiotic regimen coding were applied to a subsequent one-year cohort, resulting in rates of overall CPC compliance of 75%.

Despite the improved methodology, the determination of “CPC compliant” in both publications allowed for subjects who had received expanded Gram-negative coverage for open fracture and maxillofacial injuries because of the conflicting recommendations of the 2008 CPC and 2010 JTS guidance. For example, in the 1-year analysis, 48% of patients with open fractures who were considered “compliant” with the 2008 CPC received Gram-negative coverage in addition to the standard Gram-positive regimen. Similarly, 27% of “compliant” maxillofacial injury patients received Gram-negative coverage.\textsuperscript{17,18} Furthermore, in both the 6-month and 1-year analyses,\textsuperscript{17,18} vancomycin was allowed as a substitution for Gram-positive prophylaxis despite not being recommended by either the 2008 CPC or 2010 JTS guidance.

Examination of data from the 1-year analysis suggested that a longer term assessment of antimicrobial use trends was needed. Currently, the overarching observational cohort project, the DoD—Department of Veterans Affairs Trauma Infectious Disease Outcomes Study (TIDOS), has collected data for over 5 years. During this period, use of broad-spectrum antibiotics for prophylaxis was discouraged as embodied in the 2008 and 2011 CPGs. Nevertheless, the antibiotic stewardship goal of reducing the use of Gram-negative coverage in the setting of postcombat-trauma antibiotic prophylaxis was hindered by varied clinical practice patterns of trauma surgeons, high turnover of trauma team members, and the fast-moving practice environment unique to trauma care in a war zone (e.g., several transfers of care within the first 72 hours postinjury). In the analysis herein, we evaluate the administration of postinjury prophylactic antibiotics vis-à-vis the 2011 CPC recommendations, which embody the antibiotic stewardship goals throughout the study period. Additionally, we examine the use of expanded Gram-negative coverage across injury types and duration of prophylactic antibiotic use over a 5-year period.

METHODS

Study Population and Data Collection

Trauma patients were eligible for inclusion in the analysis if they were active-duty personnel or DoD beneficiaries, at least 18 years of age, and sustained deployment-related injuries in either the Iraq or Afghanistan combat theaters requiring medical evacuation to LRMC (Germany) between June 1, 2009 and May 31, 2014. These data were collected as part of TIDOS, which is an ongoing observational cohort study assessing the short- and long-term infectious complications related to deployment-related traumatic injuries.\textsuperscript{19} For the analysis, data were obtained from the DoD Trauma Registry (DoDTR)\textsuperscript{20} and supplemented by the TIDOS infectious disease module. This study was approved by the Infectious Disease Institutional Review Board of the Uniformed Services University of the Health Sciences (Bethesda, Maryland).

Injury Characterization and Classification

Injury data obtained from the DoDTR were standardized into AIS-defined codes\textsuperscript{21} using an injury coding software system, Tri-Code (Digital Innovations, Inc., Forest Hill, Maryland). In brief, AIS is an anatomic based injury severity scoring system, which allows for the categorization of distinct injury types (e.g., blunt force and penetrating trauma) by specific body regions. Trauma patients were classified into one of the five established injury categories\textsuperscript{18} based on their pattern of injury as well as their antibiotic prophylaxis requirement in a stepwise fashion using the highest level of antibiotic coverage required. Patients not meeting an injury criteria requiring antibiotic prophylaxis were placed in the “closed” injury category. Because of the limited numbers, patients meeting the criteria for a penetrating central nervous system injury were excluded from the analysis.

Antimicrobial Prophylaxis Compliance

As previously described,\textsuperscript{18} antibiotic use was determined via prospective review of the medical records. Compliance was defined as receipt of approved antibiotics in accordance with the 2011 CPC recommendations.\textsuperscript{13} Unlike the previous two publications that evaluated compliance over 6 and 12 months,\textsuperscript{17,18} expanded Gram-negative coverage in open fracture and maxillofacial injuries was not deemed compliant. Antibiotic regimens, including broad-spectrum Gram-negative coverage, were classified as previously described\textsuperscript{18} and a new class (expanded Gram-negative coverage) was developed, which included regimens with (e.g., piperacillin/tazobactam and meropenem) and without (i.e., levofloxacin) anaerobic coverage. Although not available in the United States, intravenous amoxicillin–clavulanate was the first-line choice for postinjury antimicrobial prophylaxis by the British military, in accordance with their published guidelines.\textsuperscript{22} As a portion of wounded U.S. military personnel were treated by coalition forces in the combat zone, use of amoxicillin–clavulanate as a substitute for cefazolin and cefazolin/metronidazole (in the case of penetrating abdomen injury) was deemed to be compliant. Furthermore, use of vancomycin with any injury category was not considered compliant. This more stringent evaluation was chosen to better assess trends toward compliance with the 2011 CPC, which embody the expert panel’s antibiotic stewardship goals.
As in our previous publications, adherence was assessed in the immediate period following injury out to 48 hours to account for the potential of documentation omissions and multiple transitions of care associated with combat trauma care/medical evacuation. Our analysis also examines duration of antimicrobial use specific to injury patterns and use of specific antimicrobials (i.e., ciprofloxacin/levofloxacin, vancomycin, meropenem/imipenem, and aminoglycosides).

**Statistical Analysis**

Categorical variables across the study years were assessed using Fisher’s exact and χ² tests for trends. Nonparametric tests were used to compare overall continuous variable distributions. Data were also examined on a per study year basis: 2009–2010, 2010–2011, 2011–2012, 2012–2013, and 2013–2014 (June through May of subsequent year). Logistic regression was also used to examine potential predictors for antimicrobial prophylaxis compliance (related to 2011 CPG) in a univariate and multivariate analysis. Statistical analysis was performed with SAS version 9.3 (SAS, Cary, North Carolina). Significance was defined as p < 0.05.

**RESULTS**

**Study Population**

A total of 5,196 military personnel sustained injuries during the study period (June 2009 through May 2014). The patients were predominantly men (98%) who sustained blast injuries (55%) in support of military operations in Afghanistan (90%). The proportion of blast injuries remained fairly consistent during the first 3 years of the analysis (56–59%); however, as military operations slowed during the last 2 years of the analysis, it declined to 48% during 2012–2013 and 39% in 2013–2014 (p < 0.001). The proportion of wounded personnel with gunshot wounds also increased from 16% during the first year to 21% in 2013–2014 (p = 0.005). Similarly, injury severity score (ISS) varied over the analysis period with the highest scores during 2010–2011 (median: 10; interquartile range: 5–22) and the lowest during the last year of the analysis (median: 8; interquartile range: 4–17; p < 0.001). Furthermore, the proportion of patients with a shock index >1.5 (i.e., heart rate/systolic blood pressure) significantly increased from 2.2% in 2009–2010 to 3.5% in 2011–2012 and then declined to 2.8% in 2013–2014 (p = 0.015). There was no significant difference in the proportion of patients admitted to the LRMC intensive care unit (ICU) over the study period (26–31%; p = 0.194).

Infections were observed in 5% of wounded personnel admitted to LRMC. As with injury severity, the proportion of patients with infections varied over the years with an increase from 3.4% in 2009–2010 to 7.1% in 2011–2012, followed by a decline to 2.3% in the final year (p < 0.001). Furthermore, 2,445 (47%) patients transferred to a participating military hospital in the United States, of which 30.0% had an infectious complication. There was a significant difference in the proportion of patients with an infection at the U.S. hospitals over the study period with an increase from 27% in 2009–2010 to 36% in 2012–2013 (p = 0.001).

**Injury Patterns**

Injuries requiring prophylaxis for open fractures were the most common injury pattern accounting for 37% of the study cohort followed by patients with skin and soft tissue (25%) and closed (24%) injury patterns (Table I). There was a significant difference in the pattern of injuries across the study years (p < 0.001). Open fractures peaked during 2010–2011 (42%) and then declined to 30% in the last study year. In contrast, closed injuries declined from 24% in the first year to 19% during 2010–2011 and then increased to 33% in the last year of the analysis. Penetrating abdominal injuries increased from 4% in 2009–2010 to 5% between 2010 and 2013, whereas maxillofacial injuries declined from 11% in the first year to 7% in the last year of the analysis. Lastly, the proportion of skin and soft-tissue injuries remained fairly constant across the years, ranging from 23 to 27%.

**Antimicrobial Use Patterns**

Between 2009 and 2011, 58 to 61% of patients with open fractures received expanded Gram-negative coverage; however, the proportion significantly decreased to 12, 8, and 7% in 2011–2012, 2012–2013, and 2013–2014, respectively (p < 0.001; Figure 1). A similar pattern was observed for maxillofacial injuries with the proportion of expanded

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Open Fractures</td>
<td>1,922 (37.0)</td>
<td>401 (34.9)</td>
<td>631 (41.8)</td>
<td>499 (36.3)</td>
<td>285 (35.1)</td>
<td>106 (30.0)</td>
</tr>
<tr>
<td>Skin and Soft Tissue</td>
<td>1,306 (25.1)</td>
<td>306 (26.7)</td>
<td>376 (24.9)</td>
<td>322 (23.4)</td>
<td>215 (26.5)</td>
<td>87 (24.7)</td>
</tr>
<tr>
<td>Closed†</td>
<td>1,252 (24.1)</td>
<td>275 (24.0)</td>
<td>290 (19.2)</td>
<td>361 (26.3)</td>
<td>208 (25.7)</td>
<td>118 (33.4)</td>
</tr>
<tr>
<td>Maxillofacial</td>
<td>470 (9.1)</td>
<td>123 (10.7)</td>
<td>134 (8.9)</td>
<td>123 (9.0)</td>
<td>65 (8.0)</td>
<td>25 (7.1)</td>
</tr>
<tr>
<td>Penetrating Abdomen‡</td>
<td>246 (4.7)</td>
<td>43 (3.8)</td>
<td>78 (5.2)</td>
<td>70 (5.1)</td>
<td>38 (4.7)</td>
<td>17 (4.8)</td>
</tr>
<tr>
<td>Total</td>
<td>5,196</td>
<td>1,148</td>
<td>1,509</td>
<td>1,375</td>
<td>811</td>
<td>353</td>
</tr>
</tbody>
</table>

All indicated time periods begin in June and end in May. Data are expressed as number (percentage). Injury patterns were significantly different across the years; p < 0.001. *Injury patterns based on Abbreviated Injury Scale (2005-Military) codes. †Defined by no evidence of open wounds. ‡Defined by injury to a hollow viscus except for hematomas, anal injuries, and “no perforation partial thickness.”

---

Variation in Postinjury Antibiotic Prophylaxis Patterns
Gram-negative coverage decreasing from 36–37% between 2009 and 2011 to 12–25% in the later period (p = 0.001).

Along with the use of expanded Gram-negative coverage, specific antibiotics evaluated included vancomycin, carbapenems, and fluoroquinolones. In general, very little vancomycin was prescribed during the study period. Patients with penetrating abdominal injuries received the largest proportion (7%) during 2009–2010. With regard to injury patterns, there were no significant differences in the use of vancomycin over the study years. The largest proportion of carbapenem usage (6%) was associated with penetrating abdominal injuries during 2013–2014; however, there was no statistical difference in the receipt of the antibiotic over the study years for any of the injury patterns. Use of levofloxacin and ciprofloxacin did significantly decrease for each injury pattern over the study period (p < 0.001) with the largest reduction being from 51–55% with open fractures (2009–2011) to 2–5% (2011–2014). For penetrating abdominal injuries, 42% were prescribed levofloxacin/ciprofloxacin in 2010–2011, which decreased to 8% in 2012–2013 and zero in 2013–2014.

Cefazolin use in open fractures and soft-tissue injuries requiring prophylaxis varied widely during the study years (Table II). Specifically, 95% of patients with open fractures received cefazolin; however, the proportion that received the antibiotic for more than 5 days significantly reduced from 51% in 2009–2010 to 23–28% between 2011 and 2014 (p < 0.001). Duration of cefazolin use in soft-tissue injuries also significantly decreased during the study years (p = 0.018), with 16% receiving the antibiotic for more than 5 days in 2009–2010 compared to 8–10% between 2011 and 2014.

Adherence to Antimicrobial Prophylaxis Recommendations
Using the 2011 CPG as the reference, antimicrobial adherence ranged from 34% with penetrating abdominal injuries to 83% with closed injuries (Table III). Compliance with guidelines related to closed injuries significantly improved over the study years from 78% in the first year to 90% in the last (p = 0.006). In addition, maxillofacial injuries showed improvement with an increase from 50% compliant in 2009–2010 to 76% in 2013–2014 (p = 0.042). Open fracture compliance also improved from 34% in the first year of the analysis to 77% in 2012–2013 and then declined slightly to 73% in the final year (p < 0.001). Furthermore, skin and soft-tissue injuries had an increase in compliance from 58% in 2009–2010 to 69% in 2011–2012 and then decreased to 61% in 2013–2014 (p = 0.013). Finally, penetrating abdominal injuries improved from 16% in 2009–2010 to 47% in 2011–2012 and then declined to 34 to 35% for the final 2 years of the analysis (p = 0.019). The overwhelming reason for noncompliance related to penetrating abdominal injuries was administration of antimicrobial regimens not consistent with the CPG rather than not prescribing prophylaxis.

Predictors of Antimicrobial Prophylaxis Compliance
Time period (injuries sustained during 2011–2014 as compared to the earlier years), admission to LRMC ICU, volume of blood product transfusion within 24 hours, age at injury, shock index, and ISS were examined for an association with CPG antimicrobial prophylaxis compliance in a univariate analysis (Table IV). Because of the high correlation with injury severity and LRMC ICU (p < 0.001), injury pattern was not included in the model. Admission to the LRMC ICU, shock index (>1), and ISS (≥10) were significantly associated with noncompliance (p < 0.001), while age at injury and time period were significantly associated with compliance (p = 0.032 and <0.001, respectively). Using stepwise selection, only ISS, LRMC ICU admission, and time period were retained in the final multivariate analysis (Table IV). As with the univariate analysis, ISS (≥10) and LRMC ICU admission were associated with noncompliance (p < 0.001), whereas being injured in the time period of
2011–2014 was associated with antimicrobial prophylaxis compliance ($p < 0.001$).

DISCUSSION

In the years leading up to the publication of the 2011 CPG, DoD combat-related antibiotic prophylaxis recommendations (e.g., 2008 CPG and 2010 JTS guidance) were contradictory. This discord was a reflection of not only disagreement in clinical practice between the JTS trauma team members, but also the controversy that existed in the civilian literature regarding the use of Gram-negative coverage in the setting of Type III open fractures.\textsuperscript{1,2,6,12–25} The primary goal of the 2011 CPG and follow-on internal 2012 JTS guidance was to standardize care and reduce variability in practice in the fast-moving and austere environment which complicates combat trauma care.\textsuperscript{13,14} Overall, our analysis of antimicrobial usage patterns over a 5-year period highlights an increasing trend toward compliance (in accordance with the 2011 published CPG) from 2009–2010 to 2013–2014. Notably, there was a marked reduction in the administration of expanded Gram-negative coverage and broad-spectrum antibiotics that coincided with the release of the CPG, which was published and disseminated downrange to military treatment facilities, Landstuhl Regional Medical Center, and U.S. facilities caring for combat casualties. Furthermore, the participation of senior DoD trauma leaders in the development of the guidelines, along with infectious disease physicians, most likely led to improved acceptance of the recommendations.

Over the study period, compliance with the 2011 CPG improved in the period around June 2010 to June 2012, and these improvements were sustained throughout the study period. Additionally, and even though injury severity (which has been shown to be a negative predictor of CPG compliance) declined over the years, the multivariate analysis identified time period as a significant predictor of compliance (odds ratio: 0.40; 95% confidence interval: 0.35–0.45). Use of broad-spectrum antibiotics such as meropenem was very low, as was the use of vancomycin. These results provide further evidence of the effectiveness of the CPG and add to the medical literature supporting the use of standardized guidelines to enhance antibiotic stewardship efforts. Of note, a recent assessment of compliance with recommendations at trauma centers found that for every 10% increase in compliance, there was an associated 14% decrease in mortality risk.\textsuperscript{26} Furthermore, adherence to surgical antibiotic prophylaxis guidelines among trauma patients was associated with a lower proportion of surgical site infections, duration of antibiotic usage, and total hospitalization.\textsuperscript{27} Therefore, a

### TABLE II. Usage of Cefazolin by Injury Pattern with Regards to Duration, N (%)\textsuperscript{a}

<table>
<thead>
<tr>
<th>Cefazolin Usage</th>
<th>Open Fractures\textsuperscript{a} (N = 1,922)</th>
<th>Skin/Soft Tissue (N = 1,306)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Received Cefazolin</td>
<td>1,825 (95)</td>
<td>1,054 (81)</td>
</tr>
<tr>
<td>2009–2010\textsuperscript{b} Total Injuries</td>
<td>401</td>
<td>306</td>
</tr>
<tr>
<td>Duration of Cefazolin Use</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤ 48 hours</td>
<td>94 (23)</td>
<td>112 (37)</td>
</tr>
<tr>
<td>3–5 days</td>
<td>87 (22)</td>
<td>92 (30)</td>
</tr>
<tr>
<td>&gt;5 days</td>
<td>205 (51)</td>
<td>48 (16)</td>
</tr>
<tr>
<td>2010–2011\textsuperscript{b} Total Injuries</td>
<td>631</td>
<td>376</td>
</tr>
<tr>
<td>Duration of Cefazolin Use</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤ 48 hours</td>
<td>199 (32)</td>
<td>152 (40)</td>
</tr>
<tr>
<td>3–5 days</td>
<td>154 (24)</td>
<td>118 (31)</td>
</tr>
<tr>
<td>&gt;5 days</td>
<td>254 (40)</td>
<td>45 (12)</td>
</tr>
<tr>
<td>2011–2012\textsuperscript{b} Total Injuries</td>
<td>499</td>
<td>322</td>
</tr>
<tr>
<td>Duration of Cefazolin Use</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤ 48 hours</td>
<td>173 (35)</td>
<td>148 (46)</td>
</tr>
<tr>
<td>3–5 days</td>
<td>182 (36)</td>
<td>88 (27)</td>
</tr>
<tr>
<td>&gt;5 days</td>
<td>115 (23)</td>
<td>26 (8)</td>
</tr>
<tr>
<td>2012–2013\textsuperscript{b} Total Injuries</td>
<td>285</td>
<td>215</td>
</tr>
<tr>
<td>Duration of Cefazolin Use</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤ 48 hours</td>
<td>109 (38)</td>
<td>88 (41)</td>
</tr>
<tr>
<td>3–5 days</td>
<td>82 (29)</td>
<td>54 (25)</td>
</tr>
<tr>
<td>&gt;5 days</td>
<td>73 (26)</td>
<td>20 (9)</td>
</tr>
<tr>
<td>2013–2014\textsuperscript{b} Total Injuries</td>
<td>106</td>
<td>87</td>
</tr>
<tr>
<td>Duration of Cefazolin Use</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤ 48 hours</td>
<td>38 (36)</td>
<td>33 (38)</td>
</tr>
<tr>
<td>3–5 days</td>
<td>30 (28)</td>
<td>21 (24)</td>
</tr>
<tr>
<td>&gt;5 days</td>
<td>30 (28)</td>
<td>9 (10)</td>
</tr>
</tbody>
</table>

\textsuperscript{a}Duration of cefazolin use is significantly different across the study years ($p < 0.001$ for open fractures and $p = 0.018$ for open skin and soft tissue).

\textsuperscript{b}Period starts in June and ends in May.

### TABLE III. Antimicrobial Prophylaxis in Combat-related Trauma: Adherence to Published CPG\textsuperscript{a}

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Open Fractures</td>
<td>1,922</td>
<td>1,027 (53.4)</td>
<td>138 (34.4)</td>
<td>214 (33.9)</td>
<td>379 (76.0)</td>
<td>219 (76.8)</td>
<td>77 (72.6)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Skin and Soft Tissue</td>
<td>1,306</td>
<td>810 (62.0)</td>
<td>177 (57.8)</td>
<td>217 (57.7)</td>
<td>221 (68.6)</td>
<td>142 (66.1)</td>
<td>53 (60.9)</td>
<td>0.013</td>
</tr>
<tr>
<td>Closed</td>
<td>1,252</td>
<td>1,035 (82.7)</td>
<td>213 (77.5)</td>
<td>232 (80.0)</td>
<td>302 (83.7)</td>
<td>182 (87.5)</td>
<td>106 (89.8)</td>
<td>0.006</td>
</tr>
<tr>
<td>Maxillofacial</td>
<td>470</td>
<td>264 (56.2)</td>
<td>62 (50.4)</td>
<td>68 (50.8)</td>
<td>78 (63.4)</td>
<td>37 (56.9)</td>
<td>19 (76.0)</td>
<td>0.042</td>
</tr>
<tr>
<td>Penetrating Abdomen</td>
<td>246</td>
<td>83 (33.7)</td>
<td>7 (16.3)</td>
<td>24 (30.8)</td>
<td>33 (46.9)</td>
<td>13 (14.2)</td>
<td>6 (35.3)</td>
<td>0.019</td>
</tr>
</tbody>
</table>

\textsuperscript{a}Based on 2011 CPG.\textsuperscript{11} All indicated time periods begin in June and end in May. \textsuperscript{b}Percentages are based on the total number of injuries for the different patterns per years (provided in Table I). \textsuperscript{c}$p$ value compares number compliant across study years.
continued focus on improving compliance with antibiotic stewardship goals and decreasing variability is justified.

Our study is limited to an extent by our methodology, which relies on retrospective review of prospective data abstracted from charts by trauma research nurses and populated into the DoDTR. Although we mitigated this issue in our last analysis by using AIS codes to improve injury characterization, there is still a possibility that a small number of subjects may have had their injuries misclassified. Methodology constraints notwithstanding, in this analysis, we applied more stringent assessments of adherence and still showed dramatic improvements in CPG compliance over the course of 5 years. In addition, we have further improved our methodology by delineating antibiotic regimens as those containing expanded Gram-negative coverage, which will allow us to take our research to the next logical step— studying infectious outcomes in those that did and did not receive expanded Gram-negative coverage.

**REFERENCES**


**TABLE IV.** Examination of Predictive Factors Related to Antimicrobial Prophylaxis Compliance

<table>
<thead>
<tr>
<th>Predictive Factor</th>
<th>Univariate Odds Ratio (95% Confidence Interval)</th>
<th>p Value</th>
<th>Multivariate Odds Ratio (95% Confidence Interval)</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time Period</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2009–2010</td>
<td>Reference</td>
<td>&lt;0.001</td>
<td>Reference</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>2010–2014</td>
<td>0.37 (0.33–0.42)</td>
<td>&lt;0.001</td>
<td>0.40 (0.35–0.45)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>ICU Admission at LRMC</td>
<td>1.92 (1.70–2.17)</td>
<td>&lt;0.001</td>
<td>1.51 (1.27–1.81)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Age at Injury</td>
<td>0.99 (0.98–1.00)</td>
<td>0.032</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blood Product Requirements 24 Hours Postinjury</td>
<td>0.509</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1–9 units</td>
<td>Reference</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10–20 units</td>
<td>0.98 (0.75–1.27)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;20 units</td>
<td>1.17 (0.88–1.55)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shock Index&lt;sup&gt;a&lt;/sup&gt;</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0–1</td>
<td>Reference</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1–1.5</td>
<td>1.46 (1.19–1.79)</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;1.5</td>
<td>1.77 (1.28–2.46)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ISS&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
<td>&lt;0.001</td>
<td></td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>0–9</td>
<td>Reference</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10–15</td>
<td>1.87 (1.59–2.20)</td>
<td>1.66 (1.38–1.98)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16–25</td>
<td>1.71 (1.45–2.03)</td>
<td>1.32 (1.08–1.62)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥26</td>
<td>2.40 (2.07–2.79)</td>
<td>1.68 (1.36–2.08)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>Using injury severity as a predictor, statistical power is 0.547. <sup>b</sup>Shock index is defined as heart rate/systolic blood pressure. ISS is an overall score of severity based on anatomical regional values.

**ACKNOWLEDGMENTS**

We are indebted to the Infectious Disease Clinical Research Program Trauma Infectious Disease Outcomes Study team of clinical coordinators, microbiology technicians, data managers, clinical site managers, and administrative support personnel for their tireless hours to ensure the success of this project. Support for this work (IDCRP-024) was provided by the Infectious Disease Clinical Research Program (IDCRP), a Department of Defense program executed through the Uniformed Services University of the Health Sciences. This project has been funded by the National Institute of Allergy and Infectious Diseases, National Institute of Health, under Inter-Agency Agreement Y1-AI-5072, and the Department of the Navy under the Wounded, Ill, and Injured Program.


Emergency Department Wounds Managed by Combat Medics: A Case Series

CPT Steven G. Schauer, MC USA*†‡; COL James A. Pfaff, MC USA (Ret.)†

ABSTRACT  Background: Combat medics are an integral part of their unit helping to conserve the fighting strength. Minor wounds are a common problem in the deployed settings that affect a soldier’s ability to partake in operations. The medics often manage wound care, there is very little data on the outcomes. Methods: Cases were acquired as part of a quality assurance project providing training feedback to medics on wound management. Laceration management is delegated to the medic at the direction of the provider. Follow-up included a series of short questions regarding wound outcomes: infection, revision, and cosmetic outcome (extremely satisfied = 1, unsatisfied = 5). Chart review was used when direct follow-up with the patient was not available for the remainder of the wounds. Results: The project period was from May 2014 to June 2015. During this time there were 30 wound repairs documented. Direct contact follow-up was available for 57% of the encounters, the remainder was via chart review. The location of the wounds were as follows: facial 5, head/neck 0, upper extremity (excluding hand) 3, hand 16, lower extremity 5, and trunk 1. The average wound length was 2.98 cm (range, 0.8–8.0 cm). No wounds became infected. No wounds required revision. The average cosmetic rating was 1.8 (95% confidence interval = 1.48–2.12). Conclusions: In this series of wounds closed by medics in the emergency department no complications or revisions were necessary. Further research is needed to determine if this can be extrapolated to other military settings.

INTRODUCTION

Background

Traumatic lacerations are common reason for presentation to the emergency department with over 12 million visits yearly in the United States. Most of the lacerations occur on the face, scalp, and hands of young males, similar to most of the combatant soldiers. Wound care accounts for 5 to 20% of Emergency Department (ED) malpractice claims indicating that it is relatively high risk procedure and thus competence in wound management is an essential skill.

Wound management is a complex skill that cannot be taught via a short, predeployment course. It requires ongoing training and skills maintenance. Recent data suggests that Soldiers, dependents and retirees support the use of combat medics in their clinical care in the ED and as a part of the overall medical treatment facility (MTF) mission. Given that our unique patient population supports the integration of combat medic skills training into routine care it is plausible that the garrison clinical setting could be utilized as a training platform for obtaining and maintaining this skill.

Importance

The U.S. military is rapidly transitioning into a garrison-based military status and thus skills sustainment will become ever more challenging to maintain readiness for the next conflict. We must continue to explore ways to keep the medics actively engaged in clinical-based health care for combat readiness. Wound management is both a straightforward and relatively complex skill that medics can obtain and maintain in the MTF setting in preparation for their next deployment to the austere, combat setting.

METHODS

Case Acquisition

These cases were acquired through as part of a quality assurance project at the Bayne-Jones Army Community Hospital, Department of Emergency Medicine. In support of the OTSG guidance the medics were provided training on wound closure and on completion of the training which included obtaining satisfactory skills under direct observation, they were permitted to clean and close wounds under indirect supervision with patient permission. Wound closure by the medics was at the discretion of the supervising provider.

The medics were required to fill out a standardized form collecting patient contact information, a description of the
wound, and the closure techniques. The project lead would then follow up the patients at a later date. Phone or e-mail follow-up was used. Patients were interviewed regarding wound infections, complications requiring revision, and cosmetic outcomes (1 = extremely satisfied, 2 = very satisfied, 3 = somewhat satisfied, 4 = slightly satisfied, and 5 = not at all satisfied). When unable to reach the patient via phone or e-mail their records were reviewed. Records were reviewed for evidence of infection by reviewing notes from the time of the initial visit and pharmacy records were reviewed for filling of antibiotic prescriptions. Additionally, records were reviewed for documentation of wound revision or replacement of sutures. When available, records indicating removal of sutures were reviewed for evidence of complications. Feedback was then provided to the medic based on the follow-up obtained.

**Data Analysis**

Case data were aggregated and analyzed using Microsoft Excel, version 14, Redmond, Washington. Descriptive statistics were used.

**RESULTS**

The project took place from May 2015 to June 2015. During that time 30 wounds were managed by the medics in the ED. Phone or e-mail follow-up was available on 57% of the wounds, chart review follow-up was conducted on the remainder. Follow-up was conducted an average of 56 days after the initial visit (range, 8–133 days). Table I outlines the locations of the wounds.

Active duty soldiers accounted for 77% of patients included in the project. The average wound length was 2.98 cm (range, 0.8–8.0 cm). None of the wounds became infected or had to be revised. Of the patients available for phone or e-mail follow-up the average cosmetic score was 1.8 (95% confidence interval = 1.48–2.12, range 1–3).

**DISCUSSION**

“Train as we fight” is a time honored adage for military training. While it is difficult to replicate battlefield conditions there are battlefield tasks that can be extrapolated to peacetime medical care. Wound care is one of these. The *Soldier’s Manual and Trainer’s Guide*, military occupational specialty 68W, Health Care Specialist Skill Levels 1, 2, and 3 identifies a multitude of skills that combat medics should possess with wound preparation and repair of one of these. As stated in the letter from U.S. Army Medical Command OTSG, the clinical environments at the MTFs are a prime environment for sustaining these skills. Medics are often required to function independently or in a setting with remote supervision while deployed. Minor wounds are a common occurrence that may temporarily cause removal of a soldier from combat operations. However, proper wound care can return them to duty immediately or almost immediately without having to transfer them to a higher level of care.

This case series highlights a relatively simple task, yet of significant skill and importance, to far-forward care that should be considered as we move forward into a more garrison-based status as the combat operations continue to draw down in Afghanistan. The U.S. Military and coalition forces have seen an unprecedented volume of combat operations in the Joint Theatres throughout the Iraq and Afghanistan campaigns. However, it is unlikely that we are going to see such large and continual operations during the military careers of most of the service members that are currently on active duty. Thus, we must make maintenance of clinical skills an on-going priority during the drawdown in order to maintain readiness for the next military conflict. Furthermore, the ability of medics to be a primary purveyor of wound care could improve independent licensed medical providers ability to optimize their time for other types of patient care. This would be a force multiplier in every sense of the word.

Finally, and importantly, the paucity of current literature on medic skills development and sustainment may be a reflection on the more recent emphasis placed on skills sustainment that was not seen with previous conflicts. Research into this training and sustainment is urgently needed as we move into a garrison-based military status and must maintain these skills in preparation for the next fight. This is the first project that we are aware of that documents medic suturing as both effective and competent in the proper clinical setting.

**CONCLUSIONS**

In this series of wounds closed by medics in the emergency department no complications or revisions were necessary. Further research is needed to determine if this can be extrapolated to other military settings.

**REFERENCES**

Pertussis Outbreak Among Soldiers During Basic Training: The Need for Updated Protocols

Basheer Halhal, MD*1; Yuval Glick, MD*1; Inbal Galor, MD*; Ankory Ran, MD*;
CDR David J. Bacon, MSC USN†; Elon Glassberg, MD*

ABSTRACT  Pertussis is a highly contagious, vaccine preventable upper respiratory disease. The incidence of the disease has been rising in the past few decades. During the winter of 2015, an upper respiratory outbreak occurred in one of Israel Defense Forces basic training bases in northern Israel. Following the detection of the first primary cases, a suspected outbreak investigation was initiated in conjunction with more rigorous clinical and laboratory testing efforts to include specific antibody enzyme-linked immunosorbent assay assays and polymerase chain reaction to diagnose pertussis. Initially, 1,596 soldiers were surveyed clinically using a questionnaire and physicians’ interviews for upper respiratory disease symptoms. A total of 158 soldiers were further evaluated and 38.6% (61) of those were diagnosed as having pertussis (with laboratory evidence). Based on the protocol that we developed during the course of this outbreak, a postexposure prophylaxis was given to every soldier for whom there was a high level of suspicion for infection and met the inclusion criteria for the postexposure prophylaxis protocol. The effects of the postvaccination waning immunity among a vaccinated population were demonstrated, thus the need of maintaining a high index of suspicion of *Brodetella pertussis* as a causative agent during respiratory diseases outbreaks in young soldiers.

INTRODUCTION

Pertussis is a highly contagious, vaccine preventable upper respiratory disease; caused by the fastidious gram-negative coccobacillus *B. pertussis*. The disease’s meaning in Latin is “intense cough.” Similarly, the Chinese name for the pertussis is “the 100-day cough,” reflecting the classical most distinguished characteristic of the disease: prolonged, intense cough.

The disease usually manifests as a prolonged cough with the classical symptoms, to include inspiratory whoop, paroxysmal cough, and posttussive emesis. Before introduction of a pertussis vaccine, most patients suffering from the disease were 10 years of age or less. Since the introduction of a vaccine, infected patients tend to be older, which has caused a change in disease manifestation with classical symptoms appearing in less than 50% of adults infected. This change has made diagnosing the disease more challenging and inadvertently has led to an under diagnosis of the disease and unreported cases.

*Brodetella pertussis* infects only humans, without any known animal or environmental reservoir. Pertussis infection spreads through droplets. The incubation period for *B. pertussis* ranges from 1 to 3 weeks, but usually lasts 7 to 10 days. Since the incubation period is longer than most viral upper respiratory diseases, outbreaks of pertussis are detected later in the disease process. Of the household members exposed to a person suffering from pertussis, ~33% will develop symptoms.

The Centers for Disease Control and Prevention guidelines for case definition of pertussis as adopted by the Israeli ministry of health (MOH) are the following: “In the absence of a more likely diagnosis a cough illness lasting ≥ two weeks with one of the following symptoms, will be clinically diagnosed as Pertussis: Paroxysms of coughing, OR Inspiratory "whoop," OR Posttussive vomiting.”

Since the organism is fastidious, it survives in respiratory secretions for only a few hours, and requires special conditions for its growth. The Israeli MOH guidelines for laboratory diagnosis are specific IgG levels >70 IU/mL (positive IgA is considered insufficient for diagnosis and follow-up test is recommended). According to the Israeli MOH guidelines, it is mandatory to report any cases of pertussis to the public health departments.

Bacterial culture for pertussis is most useful during the first 2 weeks of cough and before antibiotic use. Polymerase chain reaction (PCR) may be effective at diagnosing pertussis 2 to 4 weeks after onset of cough, whereas the use of serologic assays are most useful in the 2 to 8 weeks following cough onset.
Since 1957, the pertussis whole cell vaccine was included as part of the Israeli national vaccination plan. In 2002, introduction of the acellular vaccine replaced the use of the whole cell vaccine. Originally, health care providers administered the vaccine at 2, 4, 6, and 12 months of age. Since 2005, a booster was added to the schedule for children between the ages of 7 to 8 and since 2008 an additional booster was added at 13 to 14 years of age. Although 93% of Israeli born were vaccinated as children, the rates of pertussis infections have been rising since 1999. A similar trend has been documented in the United States and Western Europe. The increase in cases has been considered to be caused by a short duration of protective immunity produced by the acellular vaccine.

In 1993, Arav–Boger et al examined 533 serum specimens taken from Israeli Defense Forces (IDF) new recruits before joining the military and at discharge, demonstrating an almost 13% of soldiers seroconverted to being positive for anti to B. pertussis antibodies during their 3 years army service (based on IgG). It is worth mentioning that during those years, there were no booster doses given routinely against pertussis after 12 months of age. Rendi–Wagner et al analyzed 1,982 serum specimens in Israel between the years 2000 and 2001. The estimation yielded an infection incidence rate of 2.448 per 100,000 population compared to an annual incidence of reported pertussis of 5.6 per 100,000 for the same period. The findings indicates that despite a high vaccination coverage rate (>93%), there is still a considerable circulation of B. pertussis, particularly in adolescents and elderly in Israel.

Postexposure prophylaxis (PEP) is available and warranted for individuals who have been in close contact with a pertussis patient. The antibiotic regimen of prophylaxis is similar for the treatment’s regimen: Azithromycin, Erythromycin or Clarithromycin. To our knowledge, no widely acceptable regimen: Azithromycin, Erythromycin or Clarithromycin. To our knowledge, no widely acceptable approved protocol for dealing with pertussis outbreaks.

We report a pertussis outbreak during the winter of 2015 that occurred in one of the IDF's’ basic training bases, affecting 1,600 soldiers and resulting in 61 soldiers who were clinically and serologically diagnosed with pertussis. Did we develop a clinical protocol?

**METHODS**

Since the training recruits are generally healthy and without chronic conditions a series of questions were developed to identify acute respiratory illness (Table I). According to the infectious diseases and epidemiological experts—a recruit that answered yes to two or more questions was referred for medical examination by a physician.

Only soldiers determined by the physician to be clinically suspected for pertussis infection were treated with Azithromycin 500 mg once, followed by 250 mg Azithromycin for the 4 following days, in addition to 5 sick days.

**RESULTS**

In January 2015, two confirmed cases of pertussis infection were detected among the population of one subunit in a basic training base, using specific IgA and IgG enzyme-linked immunosorbent assay (ELISA) assays. Subsequently, the epidemiology section of the IDF medical corps headquarters was informed and “suspected outbreak” was declared.

---

**TABLE I.** List of Question Posed by Health Care Worker to Soldier Visiting the Infirmary With Respiratory Complaints

<table>
<thead>
<tr>
<th>Question</th>
<th>Yes/No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Have You Suffered From Cough during the Past 2–4 Weeks?</td>
<td></td>
</tr>
<tr>
<td>Has This Cough Resulted in Vomits?</td>
<td></td>
</tr>
<tr>
<td>Did You Have Fever During the Last 2 Weeks?</td>
<td></td>
</tr>
<tr>
<td>Did You Come in Contact with Anyone Who Suffered From Persistent Cough?</td>
<td></td>
</tr>
<tr>
<td>Has Anyone You Know Suffer from Pertussis Lately (Past 3 Months)?</td>
<td></td>
</tr>
</tbody>
</table>

All suspected cases that had persistent cough lasting more than 2 weeks, were referred to serology tests and some for PCR (due to the difficulty in obtaining a sample and transporting it in field circumstances) from the nasopharynx. The PCR was tested for *Bordetella pertussis*, *Mycoplasma*, *Pneumococcus*, *Streptococcus type A*, *Chlamidia pneumonia*.

In addition the entire population on the base was informed of the outbreak in order to increase awareness to pertussis cases among them, due to the nature of the crowdedness and decreased hygienic conditions in training bases.

As additional cases from other units continued to appear in active surveillance, it became apparent that the outbreak is not contained and the need for more comprehensive and prompt approach is needed. Without an internationally accepted approved protocol for dealing with pertussis outbreaks, a local protocol was developed with the help of senior epidemiologists: Suspected cases were considered as those suffering from cough more than 3 days, instead of more than 2 weeks. This was in order to increase sensitivity of detection, in an attempt to contain the outbreak, at the cost of more false positive cases. Two parameters were used to determine whether PEP treatment should be given to the entire members of a subunit (each made up of about 150 soldiers): “(1) The percentage of soldiers suspected of having pertussis in that particular sub-unit and (2) The number of cases with definitive diagnosis (defined as specific IgG antibodies titer > 70 IU/ML).”

The PEP treatment was given to the entire population of a subunit if three or more of the members were diagnosed as confirmed cases, or at least 30% of its members were considered as suspected. In subunits that did not meet the criteria, Individual suspected cases were still treated and isolated, even if it was decided not to treat their entire subunit.

In addition, every soldier with clinical symptoms and positive IgA antibody was treated as a suspected case and was given Azithromycin treatment and was referred to a subsequent serology test.
Active measurements were taken in order to detect and isolate those who were suspected to be infected. All the soldiers in the subunit had undergone an examination by a physician and surveyed using an epidemiologic questionnaire. The questionnaire contained questions aimed to identify any clinical symptoms of the disease, according to Centers for Disease Control and Prevention’s guidelines. After additional soldiers from two other subunits presented with similar symptoms, active surveillance measures were taken in all three subunits, which indeed detected several more soldiers suffering from prolonged cough, whooping cough, or vomiting after cough.

Overall, 1,596 soldiers from 11 subunits answered the questionnaire (Table II). Of those, 158 soldiers were determined to be “highly suspected” for pertussis infection, thus an ELISA assay for specific pertussis’ IgA and IgG antibodies was performed. Treatment was given without waiting for results, based on clinical and epidemiologic diagnosis. We referred 40 highly suspected cases including positive serology to PCR testing to detect respiratory viruses and respiratory bacteria including pertussis (Fig. 1).

Of these 158 soldiers, 38.6% (61) were found to be serologically positive (pertussis’ IgG > 70 IU/mL). All PCR results were negative for respiratory viruses and bacteria.

According to the protocol, PEP treatment was given to the whole subunit when there were three or more proved cases, or more than 30% were clinically suspicious. Accordingly, 932 soldiers from seven subunits were given PEP after they met the prophylaxis policy. This protocol proved to be affective. After implementation of the protocol the outbreak was contained.

Following the adaption of the PEP protocol, fewer cases were discovered sporadically, and the outbreak was finally contained within 2 weeks (spanning 5 weeks from the first to the last person). No hospitalizations or fatalities occurred during the outbreak.

DISCUSSION

Pertussis still represents a significant cause of respiratory tract infections in Israel, despite a well-established national vaccine immunization program. In contrary to other reports, this current outbreak occurred in a healthy, fully vaccinated 18 to 19 year old cohort, who had received a fifth dose of the vaccine at the age of 14.

In fact, since the 1970s, there has been a steady increase in report rates not only in Israel, but also in Western Europe and the United States.7,11

Over the recent years, several outbreaks were reported in Spain, the United States, and Japan, mainly among adolescents.12 we described an outbreak that took place in a close and crowded population contained in one of IDF’s basic training bases. To our knowledge, this represents one of the largest pertussis outbreaks described in an otherwise healthy population of young adults.

Introduction of the whole cell vaccine in the 1940s led to a sharp decline in the incidence of the disease in the developed world. However, during the past three decades, there has been a steady rise in pertussis reports in the United States, as well as in Israel, mainly among adolescents and adults. Hence, the current outbreak emphasized the importance of maintaining a high index of suspicion for pertussis, even among fully vaccinated population. In addition, this serves as another example for the waning immunity effect after the Tdap vaccination (which has been reported to be as high as 30% after 3 years).13

The symptoms presented by immunized patients, especially adolescents and adults, are often nonclassical. A prolonged cough may be the only manifestation of pertussis. Indeed, most of the soldiers mainly suffered from prolonged cough, without any other symptoms, a fact that strongly emphasizes and supports the need for a high index of suspicion for pertussis.

When dealing with an outbreak, prompt suspicion, isolation, and diagnostic measures are crucial for outbreak containment. Accordingly, in the setting of a pertussis outbreak, we chose to lower the bar and used a threshold of 3 days of cough instead of 2 to 3 weeks after confirmation of the first few cases, in the cost of false positives and over treatment with PEP. It is our belief that this contributed significantly to prompt control of the outbreak that was achieved within 5 weeks of the first case to the last one diagnosed.

### TABLE II. Results of the 1,596 Soldiers That Were Provided the Questionnaire

<table>
<thead>
<tr>
<th>Unit</th>
<th>N (Number of Soldiers)</th>
<th>Suspected (Number of Soldiers)</th>
<th>% Suspected</th>
<th>Positive Serology (Number of Soldiers)</th>
<th>Positive/Suspected (%)</th>
<th>Received PEP</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>145</td>
<td>18</td>
<td>12.4</td>
<td>9.0</td>
<td>50</td>
<td>y</td>
</tr>
<tr>
<td>B</td>
<td>144</td>
<td>14</td>
<td>9.7</td>
<td>7.0</td>
<td>50</td>
<td>y</td>
</tr>
<tr>
<td>C</td>
<td>170</td>
<td>28</td>
<td>16.5</td>
<td>4.0</td>
<td>14.3</td>
<td>y</td>
</tr>
<tr>
<td>D</td>
<td>100</td>
<td>5</td>
<td>5.0</td>
<td>2.0</td>
<td>40</td>
<td>n</td>
</tr>
<tr>
<td>E</td>
<td>113</td>
<td>39</td>
<td>34.5</td>
<td>19.0</td>
<td>48.7</td>
<td>y</td>
</tr>
<tr>
<td>F</td>
<td>150</td>
<td>8</td>
<td>5.3</td>
<td>4.0</td>
<td>50</td>
<td>y</td>
</tr>
<tr>
<td>G</td>
<td>147</td>
<td>6</td>
<td>4.1</td>
<td>1.0</td>
<td>16.7</td>
<td>n</td>
</tr>
<tr>
<td>H</td>
<td>144</td>
<td>9</td>
<td>6.3</td>
<td>1.0</td>
<td>11.1</td>
<td>n</td>
</tr>
<tr>
<td>I</td>
<td>104</td>
<td>11</td>
<td>10.6</td>
<td>3.0</td>
<td>27.3</td>
<td>y</td>
</tr>
<tr>
<td>J</td>
<td>97</td>
<td>12</td>
<td>12.4</td>
<td>9.0</td>
<td>75</td>
<td>y</td>
</tr>
<tr>
<td>K</td>
<td>282</td>
<td>8</td>
<td>2.8</td>
<td>2.0</td>
<td>25</td>
<td>n</td>
</tr>
</tbody>
</table>
In an attempt to confirm diagnosis, ELISA was used as it is the most available serological test in the IDF setting. Also, PCR was used in an attempt to elevate the diagnosis rate and rule out other causes (viral and bacterial). Unfortunately, all PCR results were negative. This is probably due to the poor quality of sample collection by the medics on the training base and the elongated transport period between the collection of samples and the arrival to the laboratory, which may have further lowered its sensitivity. Furthermore, some of the samples were taken more than 3 weeks from the onset of the cough. We recommend that in future outbreaks, samples should be collected as soon as symptoms occur and pertussis is suspected, by well-trained and informed medical personal and then promptly delivered to a laboratory with a higher level of diagnostic capabilities. It should be emphasized that treatment before tests results is at the discretion of the health care provider.

The decision when to or not administer the PEP to the population of an entire subunit played a crucial part of the disease outbreak management. The PEP is recommended

![FIGURE 1. Diagnosis and treatment. Graphic flow chart of decision-making protocol developed during the outbreak investigation.](image)

![FIGURE 2. Suggested protocol.](image)
after close contact to an infected person. Since all soldiers in every subunit (in similar to students of educational facilities such as boarding schools) might be in close contact with an ill soldier, we had to develop a protocol to define when PEP treatment is needed to the entire subunit. We tried to find the balance between prompt PEP to the maximal number of contacts, in an attempt to prevent additional cases and maintain the subunit fit for duty, and between over treatments. The fact that treatment protocol and PEP protocol are identical is important and sometimes helped in decision making.

In an effort to prevent a much larger pertussis outbreak an attempt to vaccinate the entire base was made. Unfortunately, vaccines were not available due to a worldwide shortage at that time. The shortage of vaccination against pertussis was a significant factor in outbreak management, and complicated the management and efforts for containment due to our inability to vaccinate healthy individuals to stop the spread of the outbreak.

This article has a few limitations. As mentioned previously, this report is based on the clinical data available that was retrospectively analyzed. Data were not always collected in a perfectly organized matter, causing gaps in collection of critical information that would have enabled the forming an epidemiologic curve as it was occurring in real time. In addition, due to the fact that it is a descriptive article and not a trial, we do not have a comparison group. The impact of the protocol implementation can be evaluated only in a before—after design and not in comparison to other outbreaks. The same protocol was later used in the response for two additional smaller pertussis outbreaks in the IDF. The prompt implementation led to quicker containment and less soldiers infected (unpublished data).

This outbreak emphasizes the importance of maintaining high index of suspicion in pertussis, even in young, healthy, and fully vaccinated adults, with prolonged cough. Prompt recognition of an outbreak and quick implementation of epidemiologic measures, such as active surveillance and post-exposure treatment for contacts, are crucial for outbreak response. Recognizing pertussis outbreaks is even more challenging since the disease manifestation is often not obvious, the incubation period is long, and laboratory diagnosis in not 100% sensitive or specific. The basic manifestation of pertussis is prolonged cough, which often results in a late diagnosis and late epidemiologic measures. Therefore, once diagnosis is made, prompt response and protocol implementation are key parameters in increasing the sensitivity for diagnosis, PEP and containment of the outbreak (Fig. 2).

We recommend that vaccination policy in Israel and implementation of a pertussis vaccine booster doses at age 18 be revised and updated periodically to answer re-emerging public health challenges like pertussis outbreaks.

Additional controlled clinical studies are necessary to develop and validate protocols for dealing with mass outbreaks of pertussis.

REFERENCES

Comparative Susceptibility of Different Mouse Strains to Liver-Stage Infection With Plasmodium berghei Sporozoites Assessed Using In Vivo Imaging

Qigui Li, PhD; Lisa Xie, PhD; Diana Caridha, PhD; Norma Roncal, MS; Qiang Zeng, MS; Jing Zhang, BS; Ping Zhang, BS; LTC Mark Hickman, MS USA; LTC Lisa Read, VC USA

ABSTRACT  Background: The liver stages of Plasmodium parasites are important targets for the discovery and development of prophylactic drugs. Methods: A real-time in vivo imaging system was used to determine the level of luminescence measured from firefly luciferase expression by sporozoites developing in hepatocytes in different strains of mice. Results: The luminescence values (photon counts/sec) measured from the anatomical liver location in the untreated mice infected with 10,000 Plasmodium berghei sporozoites were 8.15 × 10^5 for C57BL/6 Albino, 2.12 × 10^5 for C3H/HeNCrL, 0.91 × 10^5 for C57BL/6 WT, 0.28 × 10^5 for BALB/c, and 0.16 × 10^5 for ICR/CD-1 mice. This data suggests that the C57BL/6 Albino strain is most susceptible to luminescent photon, mainly because the less light scattering and absorption from deeper tissues and the skin in the strain of mouse. The photon count observed in black C57BL/6 wild type mice was shown to be 88.83% lower compared to C57BL/6 Albino mice. Although the highest growth rate of sporozoites in hepatocytes was found for C57BL/6 wild type mice in this study, the black skin of this mouse significantly reduced parasite-associated bioluminescence. Conclusions: The minimal light scattering and absorption and also enhanced susceptibility to liver infection of C57BL/6 Albino mice makes this strain preferable sensitivity for discovery and development of prophylactic antimalarial drugs.

INTRODUCTION

The exoerythrocytic (EE) stage of the malaria parasite in the host liver is the stage of development occurring between the sporozoite and the erythrocytic stages, which results from the release of merozoites from the hepatocytes to the blood. On invasion of a hepatocyte, sporozoites mature and multiply into trophozoites and subsequently into liver schizonts that contain thousands of merozoites. This initial phase, referred to as the EE stage, lasts about 48 hours in murine models and is clinically silent, whereas pathology is associated with the subsequent blood-stage infection by asexual parasite forms. Despite overall similarities, various inbred mouse strains show different susceptibilities to sporozoite infection by murine Plasmodium species. BALB/c mice show greater susceptibility for Plasmodium yoelii compared to C57BL/6 wild type (WT) mice, whereas Plasmodium berghei are more infective against C57BL/6 WT mice. Therefore, C57BL/6 WT mice are highly susceptible and BALB/c mice are relatively resistant to liver-stage infection by P. berghei sporozoites.

Varying degrees of susceptibility and resistance to malaria between different mouse strains are determined by host genetic variants, which are known to confer resistance to P. berghei liver-stage infection and control malaria severity in mice. Accordingly, a hypothesis was subsequently tested regarding the susceptibility of different mouse strains to malaria, which is reflected in an attenuated parasite load that may be a consequence of a protective immune response in the host. Clinical data have accumulated over the past 60 years to indicate that genetic factors can influence the onset, progression, severity of disease, and ultimate outcome of malaria infection in man. Similarly, this genetic determinant of susceptibility to malaria is complex, multigenic, and its analysis by genetic linkage and association studies, as well as by candidate gene testing, has revealed important interactions between host genes and the malaria parasite in mice.

Quantitative analysis of liver-stage burden is necessary for the study of drug interventions that target liver-stage parasites. Bioluminescence imaging (BLI) has long been used to detect and monitor pathogens in vitro using plate-based luminometers. The development of in vivo imaging systems, which can detect, image, and quantify luminescent and fluorescent signals from anesthetized animals, now permits real-time imaging of bioluminescent Plasmodium parasites in vivo. By using bioluminescent strains of P. berghei and a real-time imaging instrument, we have successfully conducted efficacy testing for drug development of novel 8-aminoquinoline derivatives. BLI is a highly sensitive tool for visualizing malaria parasite development, metastatic spread, and response to therapy. Although BLI has proven to be a useful tool due to its apparent simplicity and ease of implementation, few rigorous studies have been presented to assess the innate susceptibility.
of various mouse strains to *P. berghei* infection nor have studies been conducted to assess the reduced bioluminescence observed as a consequence of light scatter by deeper tissues and light absorption by pigmented molecules (e.g., hemoglobin, melanin, and other pigmented macromolecules) in colored animals. The problem of decreased bioluminescence in black mice is a known limitation for quantitative optical imaging of small animals, as skin pigmentation is a significant variable affecting BLI. This problem should be considered when establishing mouse models and conducting experimental studies with colored mice.

The necessity of using different mouse strains, which vary in efficacy to match various projects, has become more a larger issue given the requirement to conduct different studies assessing immune response, drug metabolism, and drug toxicity in a variety of mouse strains. The innate response against malaria infection in various strains of mice could interfere with the study of the immune response generated by invading sporozoites. Mouse strains with greater sensitivity to *P. berghei* liver-stage infection are also needed to reduce the number of sporozoites required to achieve nonvariant infections. Therefore, five strains of mice (C57BL/6 WT, C57BL/6 Albino, C3H/HeNCrl, BALB/c, and ICR/CD-1) purchased from the same vendor, Charles River Laboratories (Frederick, Maryland) were selected to assess infection susceptibility to *P. berghei* sporozoites. The apparent susceptibility of different strains of mice to *P. berghei* liver-stage infection measured with BLI can be impacted by the absorption of luminescent photons by the pigmented skin and hair of colored mouse strains. A number of variables were used to assess susceptibility to include BLI measurements, mortality, the influence of gender on infection, and response to drug treatment.

### METHODS

#### Animals and Sporozoite Inoculation Procedure

Age-matched, young adult (7-week-old) male or female mice were purchased from Charles River Laboratories. The four inbred mouse strains (C57BL/6 Albino, C57BL/6 WT, C3H/HeNCrl, and BALB/c) and the one outbred strain (ICR/CD-1) were chosen for in vivo imaging assessment of liver-stage infection based on their wide ranging use in malaria research and the known differences in the haplotype of the major histocompatibility region. All animal studies were performed under Institutional Animal Care and Use Committee approved protocols, which detail the experimental procedures and designs as well as the number of animals used. All animal use, care, and handling were performed in accordance with the current *Guide for the Care and Use of Laboratory Animals*. Sporozoites isolated from the same batch of mosquitoes were inoculated into various strains of mice on the same day to control for biological variability in sporozoite preparations. Each mouse was inoculated intravenously in the tail vein with approximately 10,000 to 100,000 sporozoites suspended in 0.1 mL volume on day 0 in different experiments.

#### Sporozoites Isolation and Viability Assessment

Firefly luciferase-expressing *P. berghei* ANKA sporozoites were obtained from laboratory-reared female *Anopheles stephensi* mosquitoes from the Department of Mosquito Biology, Walter Reed Army Institute of Research, Silver Spring, Maryland. The mosquitoes were maintained at 18°C for 19 to 22 days after feeding on *P. berghei* malaria-infected Swiss ICR/CD-1 mice. Salivary glands were extracted from malaria-infected mosquitoes and kept on ice in Roswell Park Memorial Institute medium with 1% mouse serum. To ensure that inoculated sporozoites were viable following the isolation procedure, they were stained with a vital dye containing fluorescein diacetate (50 mg/mL in acetone) and ethidium bromide (20 μg/mL in phosphate-buffered saline [Sigma Chemical Co., St. Louis, Missouri]) and counted in a hemocytometer. The viability of sporozoites ranged from 86 to 100%.

#### Susceptibility of Different Mouse Strains to Sporozoite-Induced Liver Infection

The infectivity of the sporozoites was estimated by determining the minimal number of sporozoites, inoculated intravenously, necessary to produce an infection in 100% of the mice per strain, as previously described. Five mice per strain were each inoculated with 10,000 sporozoites. Beginning on day 1 (24 hours) postinoculation, in vivo imaging studies of bioluminescence activity from luciferase-expressing *P. berghei*-infected mice were performed using a Perkin Elmer In vivo Imaging System (IVIS) Spectrum (Hanover, Maryland). Mice were evaluated at 24, 48, and 72 hours post sporozoite inoculation to determine liver- and blood-stage malaria infection. Starting on day 6 until day 30 after infection, the same mice were analyzed for blood-stage infections by quantitation of malaria parasites by flow cytometry (FCM). All FCM analyses were carried out with a FC500 MPL flow cytometer (Beckman Coulter, Fullerton, California).

#### Susceptibility Differences between Male and Female Mice to *P. berghei* Sporozoites

The infectivity of the viable and isolated sporozoites was estimated by determining the minimal number of sporozoites, inoculated intravenously, necessary to produce an infection in 100% of the mice per gender in four mouse strains. Five mice per gender for each strain were each inoculated with 10,000 sporozoites for four strains (C57BL/6 Albino, C3H/HeNCrl, and C57BL/6 WT) or 100,000 sporozoites for an infection resistant strain ICR/CD-1. Beginning at 24 hours postinoculation, in vivo imaging studies of bioluminescence activity from luciferase-expressing *P. berghei* sporozoites-infected mice were performed using the IVIS Spectrum. Mice were evaluated at 24, 48, and 72 hours post sporozoite inoculation to determine liver- and blood-stage malaria infection.2
IVIS Spectrum

In vivo imaging studies of bioluminescence activity from luciferase-expressing *P. berghei*-infected mice were performed using a Perkin Elmer IVIS Spectrum. Mice were evaluated at 24, 48, and 72 hours postsporozoite inoculation to determine liver- and blood-stage malaria infection. Mice received 150 mg/kg luciferin (Gold Biotechnology, St. Louis, Missouri) intraperitoneally in a volume not to exceed 150 μL. Three minutes postluciferin administration, the mice were anesthetized with inhaled isoflurane. The mice were then positioned ventral side up in the IVIS Spectrum on a 37°C platform. The mice continue to receive isoflurane through nose cone delivery. The camera exposure times utilized were 1 or 5 minutes for the 24-, 48-, and 72-hour time points with f-stop = 1 and the large binning setting was chosen. Quantitative analysis of bioluminescence emitted from whole bodies or region of intensity (ROI) were determined by measuring the luminescence signal intensity in photons per second using the ROI settings of the Living Image 3.0 software (Caliper Life Sciences, Inc., Hanover, Maryland). The ROI, expressed in total flux of photons, was set to measure the abdominal area at the location of the liver from whole body imaging.

Standard Drug Assay in High Susceptible Mouse Strains

Tafenoquine (TQ) and 4-methyl-primaquine (4-Me-PQ) were used as positive control compounds. The control antimalarial drugs were administered orally on days −1, 0, and 1 with respect to sporozoite inoculation. At 24, 48, and 72 hours postsporozoite infection, all inoculated mice were tested using the IVIS Spectrum instrument (Perkin Elmer, Hanover, Maryland). In addition, blood-stage infections were measured by FCM. Positive and negative controls were used for the IVIS calibration in each test. Susceptible strains of C57BL/6 Albino, C3H/HeNCrI, and C57BL/6 WT mice were challenged intravenously with 10,000 luciferase-expressing *P. berghei* sporozoites extracted from the salivary glands of heavily infected *Anopheles stephensi* mosquitoes. In vivo imaging assessments were performed at 24, 48, and 72 hours after infection. At day 6 to 30 after infection, the same mice were analyzed for blood-stage infection by determination of the course of parasitemia. Blood samples (3 μL each) for parasitemia determinations were collected on day 6 after inoculation and every other day thereafter until positive parasitemia was demonstrated or for 30 days, then twice weekly for another 4 weeks if negative for parasites.

Data Analysis

The calculated parameters derived from in vivo imaging experiments, causal prophylaxis activity, sporontocidal activity, parasite clearance, causal cure, delays in patency, and time to recrudescence were calculated as described previously. The minimum effective dose in 100% of all animals infected was defined as the lowest dose, which cured all animals in an experimental group during the first 30 days of the follow-up period after sporozoite inoculation. The data were generally found to fit a normal distribution. Means and standard deviations of photon measurement were also calculated. Blood-stage infection was defined as two positive parasitemia determinations made by FCM on blood samples taken on 2 consecutive days. The animals with negative parasitemia blood samples were monitored until day 30 after inoculation. Coefficients of variation were calculated as a percentage by dividing the standard deviation by the mean value. Statistical analysis was conducted with Microsoft Excel software by using a Student *t* test for dependent samples to compare the means of paired and unpaired samples between treatment groups.

RESULTS

The apparent susceptibility of five different mouse strains, both male and female, with varying skin pigments to *P. berghei* liver-stage infection was conducted by using BLI.

Susceptibility of Different Mouse Strains to *P. berghei* Sporozoite Infections

The susceptibility to infection of five different mouse strains, both inbred and outbred, to liver infections with *P. berghei* was determined by assessing bioluminescent signals from the liver at 24 and 48 hours, and the whole body at 72 hours after inoculation with intravenous sporozoites (Table I). The lowest observed intensities of bioluminescent signal from the liver region were induced by a minimal sporozoite inoculum of 10,000 in the least susceptible ICR/CD-1 mouse strain, and therefore an inoculum of 10,000 sporozoites was used in all mouse strains tested to make susceptibility comparisons. The data obtained showed that the highest susceptibility was found in C57BL/6 Albino mice, and the mean luminescence value (photon counts) collected from the livers in untreated control mice was 8.15 × 10^5^ photons/second (coefficients of variation [CV] = 16.20%) at 24 hours. The lowest sensitivity was found in ICR/CD-1 mice who demonstrated bioluminescent signals of 0.16 × 10^5^ photons/second (CV = 35.14%) following inoculation with 10,000 sporozoites.

At 48 hours, the highest apparent susceptibility, with mean photon counts of 136.92 × 10^5^, was found in C57BL/6 Albino mice (CV = 31.58%). Similar to the results obtained at 24 hours, the lowest sensitivity was shown in ICR/CD-1 mice who showed a bioluminescent signal of 1.03 × 10^5^ photons/second (CV = 42.05%) at 48 hours. At 72 hours, *P. berghei* liver merozoites will enter the blood and invade erythrocytes after the rupture of liver schizonts at 48 hours post inoculation. Consequently, at 72 hours, the mean luminescence signals could be detected only in the whole bodies of untreated mice (Fig. 1A–C, VC groups). The mean luminescence signal of 523.03 × 10^5^ photons/second (CV = 28.40%) observed in C57BL/6 mice was the
Comparative Susceptibility of Different Mouse Strains

| TABLE I. | Luminescence (ROI Intensity: Photons/Sec) and Parasite Growth Rates During Liver-Stage (24 and 48 Hours Postinoculation) and Blood-stage (72 Hours After Inoculation) Following Inoculation With 10,000 Firefly Luciferase-Expressing P. berghei Sporozoites Intravenously in five Different Mouse Strains (n = 25–45)³.

<table>
<thead>
<tr>
<th>Mouse Strains</th>
<th>Liver Stage</th>
<th>Growth Rate in Liver</th>
<th>Blood Stage</th>
<th>Growth Rate in Blood</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>24 hoursᵃ</td>
<td>48 hoursᵇ</td>
<td>48 hours/24 hours</td>
<td>72 hoursᵃ</td>
</tr>
<tr>
<td>C57BL/6 Albino</td>
<td>8.15 × 10⁵ (16.20)</td>
<td>136.92 × 10⁵ (31.58)</td>
<td>16.7 (24.61%)</td>
<td>523.03 × 10⁵ (28.40)</td>
</tr>
<tr>
<td>C3H/HeNCrl</td>
<td>2.12 × 10⁵ (32.07)</td>
<td>21.67 × 10⁵ (40.30)</td>
<td>10.18 (38.21)</td>
<td>96.32 × 10⁵ (36.33)</td>
</tr>
<tr>
<td>C57BL/6 WT</td>
<td>0.91 × 10⁵ (17.92)</td>
<td>27.64 × 10⁵ (27.12)</td>
<td>30.35 (18.51)</td>
<td>287.93 × 10⁵ (23.00)</td>
</tr>
<tr>
<td>BALB/c</td>
<td>0.28 × 10⁵ (46.45)</td>
<td>2.12 × 10⁵ (35.08)</td>
<td>7.59 (36.63)</td>
<td>5.31 × 10⁵ (46.50)</td>
</tr>
<tr>
<td>ICR/CD-1</td>
<td>0.16 × 10⁵ (35.14)</td>
<td>1.03 × 10⁵ (42.05) ***</td>
<td>6.44 (36.80)</td>
<td>3.96 × 10⁵ (33.44)</td>
</tr>
</tbody>
</table>

p Values

| C57BL/6 Albino/C3H | <0.001 | <0.001 | <0.01 |
| C57BL/6 Albino/WT | <0.001 | <0.001 | 0.494 |
| C3H/HeNCrl/WT | <0.001 | 0.089 | <0.01 |

ᵃMean values of photons measured and the parasite growth rate in liver and blood stages in different strains at 24, 48, and 72 hours after intravenous inoculation with P. berghei sporozoites.ᵇHours after intravenous inoculation with 10,000 P. berghei sporozoites.ᶜTwo mice at 24 hours and one mouse at 48 hours did not show infection, but they were infected at 72 hours, showing a delayed infection after inoculation.

The Impact of Colored Mice on Determination of Infection Susceptibility

Bioluminescent in vivo imaging is a highly sensitive tool for visualizing malaria parasites in anti-malarial drug discovery and development studies; however, bioluminescence is scattered by deeper tissues and absorbed by pigmented molecules (e.g., hemoglobin and melanin) in colored animals.¹³,¹⁴ In this study, the bioluminescent signal was significantly decreased in the colored C3H/HeNCrl mice (grey color) and the colored C57BL/6 WT mice (black) in comparison to the C57BL/6 Albino (white) mice. In the liver-stage infections conducted in this study, the mean luminescence values (photon counts) collected from the liver location of C57BL/6 Albino white mice were 8.15 × 10⁵ at 24 hours. This value is 3.87-fold and 8.93-fold significantly higher (p < 0.001) than that of grey C3H/HeNCrl (2.12 × 10⁵) and black C57BL/6 WT (0.91 × 10⁵) mice after infection with 10,000 sporozoites (Table I). The bioluminescence intensity was reduced 73.92% and 88.83% from the liver in C3H/HeNCrl grey and C57BL/6 WT black mice, respectively. At 48 hours, the mean luminescence observed from the liver location of C57BL/6 Albino mice was 136.92 × 10⁵ photons per second, which is 6.26-fold and 4.92-fold significantly higher (p < 0.001) than bioluminescence observed from C3H/HeNCrl grey (21.67 × 10⁵) and C57BL/6 WT black (27.64 × 10⁵) mice, respectively (Table I). The intensity was similarly reduced 84.17% and 79.81% from the liver locations in C3H/HeNCrl grey and C57BL/6 WT black mice, respectively. This data suggests that the grey and black skin and hair of the two strains of mice have a significant impact on light absorption, which in turn reduces the bioluminescent signal observed.

Sporozoite-Induced Immune Protection in Different Mouse Strains

Whole sporozoite inoculation is known to induce partly or fully protective antimalarial immune responses in rodents and in man. Complete sterile protection can be also obtained in mouse strains after immunization with P. berghei sporozoites.⁵ In the liver, different rates of change in bioluminescence signal at 48 hours were observed when compared to the signal observed at 24 hours in the various mouse strains tested. The luminescent intensity (ROI) in the liver region of C57BL/6 WT mice was increased 30.35-fold, suggesting that the fastest rate of signal change, likely associated with parasite growth, of sporozoites was in this mouse strain. The growth rate of signal change was mildly increased by 16.70-fold in C57BL/6 Albino mice, 7.59-fold in BALB/c mice, and 6.44-fold in ICR/CD-1 mice (Table I; Fig. 2). The growth rate of signal change results suggests that the fastest maturation of sporozoites was observed in C57BL/6 WT mice, which appear to have very little innate immunity to infection. The more moderate rate of signal change was observed in C57BL/6 Albino and C3H/HeNCrl mice, who appear to have modest immune protection, and the least change in signal rate was observed in BALB/c and ICR/CD-1 mice suggesting these strains of mice have the highest level of innate immunity to infection.⁵,⁹,¹² (Fig. 2).
Figure 1. Representative in vivo bioluminescent images of C57BL/6 Albino (A, white hair), C57BL/6 WT (B, black hair), and C3H/HeNCrl (C, grey hair) at different time points after injection of 10,000 sporozoites. Rainbow images show the relative levels of luminescence ranging from low (blue), to medium (green), and to high (yellow/red). Luminescence levels (photons/sec) of livers in whole mice at 24-, 48- and 72-hour time points following intra-gastric dosing daily for 3-consecutive-day on days −1, 0, 1, treated with 4-Me-PQ or TQ at the ED_{100} dose (each photo left) and vehicle control (VC, right) after sporozoite infection intravenously at day 0. Normally, *P. berghei* sporozoites reside in the mouse liver for 44 to 52 hours postinfection (n = 5).
The Impact of Mouse Gender on Infections by *P. berghei* Sporozoites

In order to determine the extent to which gender affects sensitivity to infections of *P. berghei* sporozoites, male and female mice from malaria susceptible strains (C57BL/6 Albino, C3H/HeNCrl, and C57BL/6 WT) and a malaria resistant strain (ICR/CD-1) were infected with different numbers of *P. berghei* sporozoites. Although the female mice seem to be more susceptible than male mice showing mean photon increases of 15.52 to 24.94% at 24 hours, 4.93 to 38.06% at 48 hours, and 9.15 to 43.92% at 72 hours, no statistically significant differences in infection susceptibility were found between males and females in either the resistant or the susceptible mouse strains (Fig. 3). Similar results were also found with regards to the endpoints of mortality and parasitemia as no statistically significant differences between males and females were observed. The C57BL/6 Albino and C57BL WT mice both demonstrated early deaths ranging from 8.4 to 9.0 and 7.6 to 12.4 days, respectively, and late mortality was observed in C3H/HeNCrl and ICR/CD-1 mice ranging from 14.8 to 15.8 and 13.4 to 14.2 days, respectively (Fig. 4).

**ED_{100} in Prophylactic Effect of TQ and 4-Me-PQ in Three Different Inbred Mouse Strains**

Analysis of the inhibition of in vivo liver-stage development by TQ and 4-Me-PQ, which were used as positive controls was assessed using BLI in three susceptible mouse strains (C57BL/6 Albino, C3H/HeNCrl, and C57BL/6 WT). Mice were treated with daily doses of TQ at 1.5, 3.0, 5.0, and 10.0 mg/kg body weight for 3 days and daily doses of 4-Me-PQ at 2.5, 5.0, and 10.0 mg/kg for 3 consecutive days (Table II, Fig. 1A-C, treated groups). All mice treated with 3 mg/kg of TQ and 5 mg/kg of 4-Me-PQ demonstrated either no or very minimal liver-stage luminescence during the observation period from 24 to 48 hours after infection. None of these animals developed blood-stage infection, and the 3 mg/kg dose of TQ and the 5 mg/kg dose of 4-Me-PQ are clearly the ED_{100} dose. Given the 3 mg/kg dose of TQ required to achieve an ED_{100} compared to the 5 mg/kg dose of 4-Me-PQ, TQ is definitely more potent for causal prophylaxis than 4-Me-PQ in these three inbred mouse strains (Fig. 1A–C, treated groups, and Table II).

**DISCUSSION**

The susceptibility of various strains of mice to infection with *Plasmodium* sporozoites has previously been assessed by the numbers of sporozoites required to cause a liver infection determined by real-time polymerase chain reaction (RT-PCR) methods to determine infection using dissected livers. Previous reports have shown differences in susceptibility among various mouse strains are independent of the ability of the sporozoites to invade hepatocytes. These differences are known to confer resistance to *P. berghei* liver-stage infections, and these differences are reflected by genetic variation in the host mice.
The outbred strain ICR/CD-1 was the most difficult mouse strain to infect, requiring a challenge dose of 50,000 sporozoites per mouse in order to achieve a 100% infection rate by bioluminescent imaging. ICR/CD-l and BALB/c mice were clearly more resistant to liver infection than other mouse strains. Variations in infection susceptibility may be due to simple variations in the ability of sporozoites to invade or develop in the hepatocytes of different “atypical” host species, considering the fact that the mouse is not the natural host of *P. berghei*. Other possibilities not fully evaluated in this study include variations in the ability of the tissue merozoites to invade erythrocytes and develop into subsequent erythrocytic stage parasites; however, given the similar pre-patent periods between the strains studied suggests that this is not the case. Relatively minor protective mechanisms may be all that are necessary to confer complete protection in this mouse strain.

Bioluminescent in vivo imaging studies have shown that the colored skin and hair of animals can absorb luminescent photons originating from deeper tissue and thus significantly reduce the observed bioluminescence, which may provide a different outcome when compared to data derived from other, more invasive assessment tools of infection such as RT-PCR. In addition, we have also observed differences in susceptibility between mice of the same strain purchased from different vendors to infection by sporozoites (Li, unpublished data). Therefore, this study was conducted with all five strains purchased from the same vendor.

Transmission of light through mammalian tissues is affected by both scattering and absorption. Scattering occurs at membranes in the cell and organelle. Absorption of light is highly tissue dependent, and it is generally determined by the presence of hemoglobin, melanin, and other pigmented macromolecules. The blue bioluminescence (480 nm) of *Renilla* luciferase, which is more strongly absorbed by pigmented molecules and is scattered by tissue, makes this reporter less suitable for BLI compared to firefly luciferase, which emits green light (562 nm). Accordingly, the use of firefly luciferase expressing *P. berghei* sporozoites greatly enhances the sensitivity of BLI measurements. Data derived from C3H/HeNCrl mice shows that the pigmented grey skin of C3H/HeNCrl mice contributed to an average loss of signal of 73.92 to 84.17%. The heavily pigmented black skin of C57BL/6 WT mice caused an average photon loss of 79.81 to 88.83% for liver-stage infection (24–48 hours) compared to C57BL/6 Albino white mice (Fig. 2). It must be noted that the genetic background of C57BL/6 WT and C57BL/6 Albino is the same. Therefore, the variability in the

**FIGURE 3.** Effects of gender on parasite luminescence in infected mice show varying susceptibility to infection at 24, 48, and 72 hours measured by in vivo imaging after infection with 10,000 *P. berghei* sporozoites in male and female C57BL/6 Albino, C57BL/6 WT, and C3H/HeNCrl mice and with 100,000 sporozoites in male and female ICR/CD-1 mice (n = 5).

**FIGURE 4.** Effects of gender on survival of mice after infection with 10,000 *P. berghei* sporozoites were observed in male (closed circles) and female (open squares) animals. Survival was slightly higher in males than in females, but not significantly different (n = 5).
degree of signal loss observed was most likely due to differences in melanin.

Furthermore, during in vivo BLI and other whole-body optically based imaging measurements, light may be scattered multiple times, depending on its wavelength and path-length, which renders the emitted light highly diffuse and reduces spatial resolution. In addition, the tissues from different strains of mice have different optical properties. Light is transmitted more efficiently through the tissues of white or hairless mice because melanin absorbs substantial amounts of light, which is scattered more by dark fur. Despite these limitations caused by light scattering and absorption, significant amounts of emitted bioluminescent light can still be detected externally.

In terms of susceptibility to infection, there are also large differences among mouse strains reflected in the actual number of sporozoites required to achieve nonvariant infections. These differences may be due to variations in protective immune responses in mice. The C57BL/6 mice are highly susceptible to development of \textit{P. berghei} but are also difficult to protect from malaria. In contrast, BALB/c mice are poorly susceptible to \textit{P. berghei} infection but are more easily protected from \textit{P. berghei} infection. Therefore, protection appears to be inversely correlated to the susceptibility of the host to sporozoite infection. The fastest growth rate of signal change observed in the liver and blood-stage regions of interest was observed in C57BL/6 WT mice that are known to be highly susceptible to infection by \textit{P. berghei}, but in this study, this strain was not the most usable model due to the black skin of the C57BL/6 WT mouse. Using RT-PCR or other methods, the C57BL/6 WT mice have been shown to be the most susceptible to liver-stage infection after inoculation of \textit{P. berghei} sporozoites, given this data, one might conclude this strain is the best choice for experimentation.

Although a lower growth rate of signal change was found in C57BL/6 Albino mice, this strain exhibited the highest apparent susceptibility assessed by BLI due to their white skin, and they are therefore the most tractable strain for bioluminescent assays. Therefore, both the genetic background and less light scattering and absorption of mouse strains are contributed to the susceptibility of strains. However, in the present study, the minimal light scattering and absorption is more important factor to increase the susceptibility of C57BL/6 Albino.

Susceptibility to malaria differs between females and males, and this sexual dimorphism may have important implications when assessing candidate malaria vaccines and drugs. However, little is known about the mechanisms mediating these sexual differences. In some studies of human and animal malaria infection, the prevalence and intensity of parasitic infections is higher in males than females. Sex differences in exposure as well as susceptibility to parasites may contribute to sex-based differences in the intensity and prevalence of parasites. Females typically have higher immune responses than males. Furthermore, recent research has shown susceptibility and mortality rates in rodents with blood-stage malaria are higher in males than females. In general, females produce more intense humoral and cell-mediated immune responses than males. Although males are more susceptible than females to many parasites, there are parasites such like malaria for which males are more resistant than females.

No statistically significant differences in susceptibility, mortality, and parasitemia were found between males and females in the resistant ICR/CD-1 mouse strain or in the susceptible C57BL/6 Albino, C3H/HeNCrI, and C57BL/6 WT mouse strains (parasitemia data not shown). However, increased resistance to infection in males was noticed in all mouse strains tested in this study. The gender related differences in resistance differed among the four strains, but males were always more resistant than females. These data suggest that differences in gender are not an experimental influence affecting susceptibility of infection or resistance to \textit{P. berghei} sporozoites in the 4 strains tested.

By using bioluminescent in vivo imaging and associated other data such as mortality, the data suggests that one sensitive mouse strain, the C57BL/6 Albino, is superior. Two other inbred strains (C3H/HeNCrI and C57BL/6 WT) provide acceptable levels of susceptibility to \textit{P. berghei} infection, and two mouse strains (the inbred BALB/c and the outbred ICR/CD-1) were quite resistant to \textit{P. berghei} sporozoite infection. As bioluminescent signals are readily absorbed by pigmented skin and hair, the C57BL/6 Albino (white) strain was chosen at WRAIR as the best mouse strain possible for conducting BLI liver-stage infection studies to avoid problems associated with signal attenuation due to dark skin and hair pigments.
REFERENCES

Lateral Flow Rapid Test for Accurate and Early Diagnosis of Scrub Typhus: A Febrile Illness of Historically Military Importance in the Pacific Rim

Chien-Chung Chao, PhD†; Zhiwen Zhangm, MS†; Giulia Weissenberger, BS†; Hua-Wei Chen, PhD†; Wei-Mei Ching, PhD†

ABSTRACT Scrub typhus (ST) is an infection caused by Orientia tsutsugamushi. Historically, ST was ranked as the second most important arthropod-borne medical problem only behind malaria during World War II and the Vietnam War. The disease occurs mainly in Southeast Asia and has been shown to emerge and reemerge in new areas, implying the increased risk for U.S. military and civilian personnel deployed to these regions. ST can effectively be treated by doxycycline provided the diagnosis is made early, before the development of severe complications. Scrub Typhus Detect is a lateral flow rapid test based on a mixture of recombinant 56-kDa antigens with broad reactivity. The performance of this prototype product was evaluated against indirect immunofluorescence assay, the serological gold standard. Using 249 prospectively collected samples from Thailand, the sensitivity and specificity for IgM was found to be 100% and 92%, respectively, suggesting a high potential of this product for clinical use. This product will provide a user-friendly, rapid, and accurate diagnosis of ST for clinicians to provide timely and accurate treatments of deployed personnel.

INTRODUCTION

Early detection, accurate diagnosis, and timely treatment of rickettsial disease are imperative to increase our countermeasures during outbreaks. Scrub typhus (ST) is one of the most common rickettsial diseases and is caused by Orientia tsutsugamushi, an obligate intracellular gram-negative bacterium.1,2 The disease is characterized by fever, headache, myalgia, rash, pneumonia, meningitis, and, in some cases, disseminated intravascular coagulation that may lead to circulatory failure.3 Eschar at the mite bite, occurring in around 50% of patients, and lymphadenopathy can be clinical clues to differentiate ST from other rickettsial infections. If the patient is left untreated, the disease can cause up to 50% mortality.4 Reports from India documented 12 to 17% case fatality rate in recent years.5,6 In 2012, the deadly disease spread to 10 different districts and more than 140 cases were reported. At least 14 people died of this disease.7,8 In 2015, a large scale ST epidemic (126 cases) with high case fatality rate occurred in Nepal.9 ST cases have been primarily reported from the rural population in the past, but recently it is being increasingly detected in people residing in expanding cities across India.10

Historically, ST was a leading cause of morbidity and mortality in the Asia-Pacific theaters during World War II. Among U.S. Army personnel, 5,663 cases were reported, including 243 deaths.11 It was again the leading cause (20–30%) of “Fever of Unknown Origin” during the Vietnam War.12,13 With two known cases among U.S. troops and eight cases among U.N. personnel occurred during the Korean War,14 ST has been described as a potential “war stopper” on the Korean Peninsula. The mites localize in grassy “scrub” vegetation; infested and uninfested areas can coexist separated only by several yards. Soldiers become infected when looking for “cover” behind vegetation or digging “foxholes.” There have been two recent outbreaks among U.S. Marines during their training at Camp Fuji, Japan. There were nine cases in 2000 and eight cases in 2001.15 Between 2002 and 2011, Defense Medical Epidemiology Database records nine ambulatory cases and four hospitalizations due to ST. Traditional geographic distribution of the disease occurs principally within an area of about 13 million square kilometers in the eastern hemisphere, and one billion people are at risk every year.16 Review of the 1,425 publications compiled through a comprehensive literature search in December 2015 for all articles related to ST since year 2000 revealed that more than 40% are from Southeast Asia. Many U.S. military laboratories (Naval Medical Research Unit 2 [NMRC-2], NMRC-A, Armed Forced Research Institute of Medical Sciences, Walter Reed Army Research Unit Nepal, U.S. Army Medical Materiel Center-Korea) are located (Fig. 1) in this region to study strategically relevant global health infectious diseases. Recent reports for the presence of ST in the Middle East...
(United Arab Emirates), Africa, and South America (Chile)\textsuperscript{17–19} have suggested that the distribution of ST may be beyond the traditional Asian-Pacific region, which implies the increased risk for U.S. military and civilian personnel deployed to these newly emerged regions (Fig. 2). Early diagnosis of the disease is essential to ensure that patients receive prompt, proper treatment and reach full recovery in a timely manner. The infection is easily treated with doxycycline.

\begin{figure}
\centering
\includegraphics[width=\textwidth]{ Origins_of_Publications}
\caption{Recent publications for ST and the military importance of the disease. The incidence has increased during the past several years in the Pacific Rim. Those areas that have military facilities are indicated: NMRU-2-PP, Naval Medical Research Unit No. 2, Phnom Penh; WARUN = Walter Reed Army Research Unit Nepal; ROK = Republic of Korea; AFRIMS = Armed Forces Research Institute of Medical Sciences.}
\end{figure}

\begin{figure}
\centering
\includegraphics[width=\textwidth]{ FIGURE_2}
\caption{Current geographic distribution of Orientia. World map detailing the presence of selected Orientia (as indicated by symbols) by continent/country. (From Luce-Fedrow et al.\textsuperscript{18})}
\end{figure}
(100 mg twice daily for 7 days or azithromycin (Z-pack; 500 mg then 250 mg). Field-deployable, point-of-care diagnostic assays are in urgent need to provide rapid, sensitive, and specific diagnostic information from clinical samples for the identification of O. tsutsugamushi infection in the battlefield or resource-limited clinical sites.

Many of the symptoms presented by ST patients are similar to other febrile diseases, such as malaria, leishmaniasis, dengue fever, and leptospirosis. Diagnosis of ST infection has relied on the detection of O. tsutsugamushi-specific antibodies during the acute and convalescent phase of the disease, and the gold standard serological assay is the indirect immunofluorescence assay (IFA). The culturing of Orientia from patient sample is associated with very low sensitivity due to the fastidious nature of the pathogen. The modern molecular technology has the potential to detect Orientia DNA during the acute phase of illness. Unfortunately the molecular assay sensitivities are only 35 to 56% when compared to the retrospectively confirmatory assays, such as fourfold increase of antibody titers in convalescent sample versus acute sample. In the study conducted by Zhang et al, 102 of 104 confirmed ST patient sera contained IgM to O. tsutsugamushi detected by IFA, but only 36 patients’ blood were positive by polymerase chain reaction amplification. The low sensitivities for molecular assays reflect the need of sensitive immunoassays to cover the dynamic spectrum of diagnostic positivity. The molecular assays are complimenting rather than replacing serological tests. Without the confirmation of the infection, patients are often treated by clinicians with a combination of antibiotics empirically. Currently, there is one commercially available rapid lateral flow assay for ST in the United States. The Scrub Typhus CareStart tests manufactured by AccessBio (Somerset, New Jersey). It uses a mixture of antigens from 3 Orientia strains. The company initiated the process for U.S. Food and Drug Administration (FDA) clearance under the impression that no prospective study would be required. Unfortunately the predicate device was discontinued, and consequently the FDA required prospective study for clearance. The company expressed no interest in continuing to pursue FDA clearance. Recently, another company, InBios (Seattle, Washington), made a second-generation product Scrub Typhus Detect (ST Detect) with broader reactivity. ST Detect uses a set of proprietary recombinant antigens from 4 of the most prevalent Orientia serotypes (Kato, Gilliam, Karp, and TA763). The Military Infectious Disease Research Program Area L has been focused on the research and development of candidate Rapid Human Diagnostic Devices for infectious diseases of military relevance and their transition to advanced development. For this purpose, we established the standard operation procedure for the IFA using whole-cell antigen from the same 4 Orientia strains and evaluated the performance of ST Detect using prospectively collected samples from Thailand, which is an endemic area for ST. The results from this study demonstrated that the product ST Detect IgM has met the requirements of a field-deployable, point-of-care diagnostic assay for the early diagnosis of ST.

MATERIALS AND METHODS

Patient Samples
The performance of the new ST Detect rapid tests was evaluated using 249 prospectively collected febrile patient sera. The patient sera were collected from patients with undifferentiated fever between 2011 and 2013 under the protocol “Development of the clinical policy for the diagnosis of scrub typhus and other rickettsial infections in Thailand,” which was approved by institutional review board of Faculty of Medicine, Siriraj Hospital, Mahidol University, Bangkok, Thailand (COA Si 204/2011). The study subjects were informed that their samples would be stored for future use and tested for the presence and identification of pathogens. The evaluation of ST Detect using these clinical samples was based on an in-house IFA (described in section “In- House IFA Assay”) approved by institutional review board of NMRC (PJT-14-03). Table II describes the population in this study with definitive diagnosis.

ST Detect Rapid Tests
The rapid tests were manufactured by InBios. There are two lateral flow immunoassays that detect ST-specific IgG and IgM separately for presumptive diagnosis of ST. These tests were donated by the company to the U.S. government to evaluate the performance of the rapid tests. The testing procedure requires no specialized equipment, takes less than 5 minutes and results are ready in 15 minutes (Fig. 3). The test is performed according to manufacturer’s instructions. About 5 to 10 μL of sera or whole blood is added to the test strip in the area beneath the arrow (sample pad area). The test strip is placed in a test tube or a plastic well (holds 150–200 μL of liquid), followed by adding three-to-four drops (90–120 μL) of Chase Buffer. Results should be recorded at 15 minutes.

In-House IFA Assay

Whole-Cell Antigen Preparation
The seeds of 4 strains (Karp, Kato, Gilliam, and TA763) were stored in Snyder’s 1 buffer (0.22 M of sucrose, 3.6 mM of KH₂PO₄, 8.6 mM of NaHPO₄, and 4.9 mM of glutamic acid, pH 7.4) at −80°C until use. Antigens were propagated by inoculating irradiated L929 cells (3,000 rads) with a 10-fold dilution of the seed in a flask (162 cm²). The infected cells were harvested 6 to 8 days postinfection when about 80 to 90% of the cells were infected with Orientia as determined by acridine orange staining. The cells were harvested from each flask and pooled into a 250-mL bottle and centrifuged
at 8,000 rpm or 110,000 rcf for 30 minutes in a refrigerated high-speed centrifuge (Sorvall RC-5C Plus, Thermo Scientific, Waltham, Massachusetts). The cell pellets were resuspended and washed once in phosphate-buffered saline (PBS; pH 7.2) followed by another centrifugation at 8,000 rpm as mentioned earlier. The resulting cell pellet was resuspended in sterile PBS at the ratio of 1:4 (one flask of cells to 4.0 mL of PBS) and aliquots were stored in a −80°C freezer.

Slide Preparations

The whole-cell antigen from each strain was diluted to contain 1,000 organisms per ×500 microscopic field in 5 μL solution dispensed onto a regular glass slide. An equal volume of each of the four antigens were pooled and mixed thoroughly in a vortex mixer. About 5.0 μL of the antigen mixture was deposited into each well (4 mm × 12-well) of Teflon-coated the slides. Similarly, normal L929 cells were used as negative control on the same slide. The antigen slides were air-dried in the biosafety cabinet (BSC), fixed in cold acetone for 10 minutes in the refrigerator and air-dried quickly in the BSC. All slides were stored at −20°C in an air-tight plastic slide box with a small bag of desiccants and sealed with tape.

Fluorescence Staining

Before fluorescence staining, the antigen-coated slides were warmed to room temperature quickly in the BSC to avoid condensate. To determine the titers of positive sera, each serum were serially diluted in PBS at two-fold dilutions from 1:40 to 1:6,400. Ten microliter of each diluted serum was added into one well. ST positive and negative sera were included in the same slide as controls. The slides were incubated at 37°C for 30 minutes inside a humidified chamber and then rinsed three times (5 minutes each) with PBS in a Copland jar. The washed slides were quickly dried using cool air from a hair dryer. While the slides were drying, the secondary antibody (fluorescein isothiocyanate-labeled rabbit antihuman IgM or IgG) was diluted in PBS. The diluted secondary antibody (10 μL) was added into each well, incubated at 37°C for 30 minutes, and washed as before. The slides were air-dried at room temperature and mounted with buffered glycerol mounting medium.

### TABLE I. Sensitivity and Specificity of InBios ST Detect Test Using Retrospective Samples

<table>
<thead>
<tr>
<th>Evaluation Sitea</th>
<th>IgG Sensitivity (%)</th>
<th>IgG Specificity (%)</th>
<th>IgM Sensitivity (%)</th>
<th>IgM Specificity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NMRC (200)b</td>
<td>99.1</td>
<td>85.9</td>
<td>94.2</td>
<td>83.3</td>
</tr>
<tr>
<td>Thailand (166)c</td>
<td>86.3</td>
<td>96.8</td>
<td>90.2</td>
<td>85.5</td>
</tr>
<tr>
<td>China (115)d</td>
<td>90.0</td>
<td>100.0</td>
<td>94.0</td>
<td>100.0</td>
</tr>
</tbody>
</table>

aIFA test was used as the gold standard to determine the performance of ST Detect test. The numbers in parentheses indicate the number of samples tested per site. bSamples are considered IFA positive if their IgG or IgM titer was equal or greater than 40. cData summarized from Silpasakorn et al.28 dData summarized from Zhang et al.24

![FIGURE 3. Assay procedure for ST Detect.](image-url)
and a cover glass. All stained slides were kept in a slide holder to prevent exposure to light until viewing under an ultraviolet microscope in the dark room.

Reading the Slides
Slides were examined at ×400 magnification under an Olympus (Waltham, Massachusetts) ultraviolet/light microscope (Olympus BX53F). A bright apple green fluorescent color seen on the stained Orientia organisms individually or as clusters intracellulary designated the sample as positive.

Quality Control
The negative control serum and the positive control with different dilutions were included in daily runs. The negative control should have no distinct and characteristic staining of the Orientia. A total of 16 ST-positive samples with various endpoint IFA titers (40, 80, 160, 320, 640, 1,280, 2,560, and 5,120, two of each titer) were used to ensure the quality of each batch. The maximum variation tolerated is a two-fold difference of the endpoint titer in those 16 positive controls. Two blinded operators read the same slide independently to determine the titer without knowing the other’s results.

RESULTS AND DISCUSSION
On the basis of communications between FDA and InBios for the preparation of the 510k package for ST Detect, the FDA has accepted IFA as the gold standard for ST serological assays. The IFA requires fluorescence microscopes and experienced technical staff. In addition, data vary among different laboratories. For this reason, we described the details of our IFA assay using the mixture of four prevalent strains established in our laboratory. We have demonstrated the reproducibility of our IFA by testing the same positive sera at different dilutions as described above. The ST Detect IgM and IgG rapid test use a mixture of recombinant 56-kDa

| TABLE II. Description of Clinical Samples Included in the Enzyme-Linked Immunosorbent Assay Evaluation\(^a\) |
| --- | --- | --- |
| Diagnosis | No. of Samples | Criteria (No. of Samples) |
| ST | 78 | Acute Sample IgG or IgM >400 (59) |
| | | Convalescent Sample IgG or IgM >400 (2) |
| | | 4-Fold Increase in IgG or IgM Titer (6) |
| | | PCR Positive for 47- or 56-kDa Gene (5) |
| | | PCR Positive and 4-Fold Increase in Titer (6) |
| Dengue | 11 | Nonstructural Protein1, IgG, or IgM Rapid Test Positive |
| Murine Typhus | 2 | Single IFA IgG and/or IgM Titer >400, or a 4-Fold Increase in IgG or IgM Titer |
| Leptospirosis | 17 | Single IFA IgG/IgM Titer >400 or a 4-Fold Increase in Titer of IgG or IgM, or Positive PCR for LipL32 |
| Other Bacteremia | 3 | Bacteria Culture Positive |
| Coinfection (ST and Others) | 8 | Acute Sample IgG or IgM >400 (7) |
| | | 4-Fold Increase in IgG or IgM Titer (1) |
| Unknown | 130 | Negative by All Above |

\(^a\)Samples were determined for the cause of infection using the criteria listed.

| Table III. Performance of ST Detect Using Prospectively Collected Samples From Thailand (\(n = 249\)) Based on In-House IFA Titers |
| --- | --- | --- | --- | --- | --- |
| IFA ≥40 as Positive | IFA ≥80 as Positive | IFA ≥160 as Positive | IFA ≥320 as Positive | IFA ≥640 as Positive |
| --- | --- | --- | --- | --- | --- |
| **IgM** | | | | | |
| Sensitivity | 85.2 | 90 | 98.4 | 100 | 100 |
| Specificity | 98.8 | 96.1 | 95.2 | 92.2 | 89.4 |
| PPV | 97.2 | 90.1 | 87.3 | 78.9 | 70.4 |
| NPV | 92.3 | 96.1 | 99.4 | 100 | 100 |
| Accuracy | 94.4 | 94.4 | 96 | 94 | 91.6 |
| **IgG** | | | | | |
| Sensitivity | 85.2 | 90 | 98.4 | 100 | 100 |
| Specificity | 98.8 | 96.1 | 95.2 | 92.2 | 89.4 |
| PPV | 97.2 | 90.1 | 87.3 | 78.9 | 70.4 |
| NPV | 92.3 | 96.1 | 99.4 | 100 | 100 |
| Accuracy | 94.4 | 94.4 | 96 | 94 | 91.6 |

PPV, positive predictive value; NPV, negative predictive value.
proteins from the same four most prevalent strains as the antigen, which has been shown to have a broader reactivity than any ST assay developed before. 27 Previously, the prototype was evaluated in China, Thailand, and our laboratory with sensitivities and specificities >80% (Table I). 24,28 Additionally, we also evaluated the performance of an earlier version of this product using archived ST patient sera with different IFA titers. Five to nine samples for each IFA titer were randomly selected. A total of 45 archived positives with different IFA titers and 57 negatives including healthy individuals and patients with other diseases were tested. The tests were more than 96% sensitive and 100% specific for IgG and IgM. The rapid test could detect positive samples as early as day 4 after onset of fever (Ching WM, Chao CC, Chen HW, unpublished data). In this study, we evaluated the second version of the prototype with optimized detergent amount for the construction of ST Detect by using 249 prospectively collected samples from a febrile study conducted in Thailand (Table II). The test can be performed in a resource-limited setting without the need of fluorescence microscope and extensive training of the technical staff. Visual interpretation of the test strip is more definitive than IFA. The sensitivity and specificity for the ST Detect IgM were calculated using different IFA titer cutoffs (Tables III and IV). With the IFA titer cutoff value at 1:320, the sensitivity and specificity for the ST Detect IgM is 98% and 95%, respectively. We did two-fold serial dilutions starting from 1:40. The titer of 1:320 was chosen because in most of the publications titers of 1:400 or higher was determined to be proper as the cutoff value. The sensitivity and specificity for the ST Detect IgG is 74% and 77%, respectively, using the same IFA cutoff value. This study shows promising results in support of Detect IgM for utilization as a field-deployable, point-of-care test for a common infection of military significance (Table V). This is particularly relevant in a clinical setting as the detection of IgM antibodies is crucial to diagnose acute infection in endemic areas where higher IgG of a single specimen could be attributed to prior infection. The high value of negative predictive value (NPV) for the IgM test (Table III) is of importance in endemic areas to reduce the unnecessary empirical treatment of patients with similar clinical symptoms as those of dengue or malaria. In contrast, the performance of the rapid test for IgG is less satisfactory. A possible explanation for this is that these samples are from the first bleed from each patient. Most of these patients seek medical attention in less than 5 days after onset of fever, thus the lower level of IgG antibodies may not be detectable by either IFA or strips. The low specificity of IgG could be due to background level from previous infections in endemic areas. Further improvement of IgG detection will be necessary.

Field-deployable, point-of-care diagnostic assays for the detection of ST specific IgM and IgG antibodies will provide rapid, sensitive, and specific diagnostic information from clinical samples for the identification of O. tsutsugamushi infection in the battlefield or resource-limited clinical sites. The product made by InBios can be performed in a resource-limited setting without the need of fluorescence microscope and extensive training of the technical staff. ST Detect IgM has reached a technology readiness level 6 and is ready to move forward for a large-scale prospective clinical evaluation.

### Table IV. 2 × 2 Table for Sensitivity and Specificity Calculation of IgM and IgG Using Various IFA Titers

<table>
<thead>
<tr>
<th>IFA Titer</th>
<th>Positive</th>
<th>Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>≥40</td>
<td>69</td>
<td>12</td>
</tr>
<tr>
<td>≥80</td>
<td>64</td>
<td>7</td>
</tr>
<tr>
<td>≥160</td>
<td>62</td>
<td>1</td>
</tr>
<tr>
<td>≥320</td>
<td>56</td>
<td>0</td>
</tr>
<tr>
<td>≥640</td>
<td>50</td>
<td>0</td>
</tr>
</tbody>
</table>

### Table V. Performance Characteristics of InBios ST Detect IgM Test for Role of Care 1 or 2 Medical Facilities

<table>
<thead>
<tr>
<th>Assay Parameter</th>
<th>Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>94%</td>
</tr>
<tr>
<td>Turnaround Time</td>
<td>&lt;15 Minutes</td>
</tr>
<tr>
<td>Compatible Sample Types</td>
<td>Serum or Plasma or Whole Blood</td>
</tr>
<tr>
<td>Required Sample Volume per Test</td>
<td>10–50 μL</td>
</tr>
<tr>
<td>Steps Required per Test</td>
<td>Two Steps</td>
</tr>
<tr>
<td>Shelf Life</td>
<td>At Least 12 Months</td>
</tr>
<tr>
<td>Storage Condition</td>
<td>Field Temperature</td>
</tr>
</tbody>
</table>

ACKNOWLEDGMENT

This work was financially supported by Naval Medical Research Center, research work unit 6000.RAD1.J.A0310.

REFERENCES


Skin Regeneration Using Dermal Substrates that Contain Autologous Cells and Silver Nanoparticles to Promote Antibacterial Activity: In Vitro Studies

Michael A. J. Zieger, PhD*; Manuel Ochoa, MS†; Rahim Rahimi, MS†; Gonzalo Campana, MD*; Sunil Tholpady, MD, PhD‡; Babak Ziaie, PhD†; Rajiv Sood, MD§

ABSTRACT We hypothesized that the addition of silver nanoparticles (AgNP) to a dermal substrate would impart antibacterial properties without inhibiting the proliferation of contained cells. Our in vitro model was based on the commercial substrate, Integra. The substrate was prepared by simple immersion into 0 to 1% suspension of AgNP (75 or 200 nm diameter) followed by rinsing for 20 minutes and sterilization under an ultraviolet C lamp. A total of 10^7 human adipose stem cells per cubic centimeter were injected and after 1 hour, 6 × 10^5 keratinocytes/cm^2 were seeded and cultured for up to 14 days. Constructs were evaluated using a metabolic assay (WST-1), and hematoxylin and eosin and immunoperoxidase staining. Bactericidal activity was measured using a log reduction assay against bacteria that are prevalent in burns. The presence of AgNP did not significantly change the metabolic activity of constructs after 14 days of culture, and the distribution of cells within the substrate was unchanged from the controls that did not have AgNP. Antibacterial activity of Integra containing AgNP (75 nm diameter) was concentration dependent. In conclusion, the addition of AgNP to the dermal substrate suppressed bacterial growth but did not significantly affect cell proliferation, and may represent an important property to incorporate into a future clinical skin regeneration system.

INTRODUCTION
Up to 20% of combat casualties suffer significant burn injuries.1 The standard of care for deep burn wounds is skin grafting, with a piece of split-thickness skin taken from a healthy donor site. When burns cover a large area of the body, the area of healthy skin available for grafting is often insufficient. Fortunately, skin can be harvested more than once from the same donor site because the epidermis regrows over time from proliferating keratinocytes that migrate from the hair follicles and sweat glands. The dermis, however, cannot regrow, so skin grafts are taken with as little dermis as possible. Acellular dermal matrices (ADM) have been developed for use with thin skin grafts to repair deep burn wounds. ADMs are placed onto the freshly excised wounds and skin grafts are then applied on top 2 to 3 weeks later. This time interval allows for migration of dermal cells into the matrix and generation of a vascular network before the application of the skin graft.

A variation of the standard approach to repairing deep burn wounds could potentially accelerate healing and involves seeding the ADM with autologous cells at the time of its surgical placement onto the wound. Mesenchymal stem cells derived from bone marrow or adipose tissues (adipose stem cells [ASCs]) have been shown to accelerate wound healing by secreting trophic factors into the wound2,3 or by differentiating into epidermal and dermal cells.4–6 Preclinical in vivo studies have shown that treating cutaneous wounds with ASCs increases new blood vessel formation7–9 and accelerates reepithelialization8 and wound closure.10,11 ASCs could be easily isolated from liposuction fat obtained from the patient at the time of the initial wound debridement surgery. In addition to ASCs, the ADM could be seeded with autologous keratinocytes obtained by the enzymatic digestion of cutaneous punch biopsies. This approach could potentially reconstitute the epidermis over the wound using isolated keratinocytes rather than split-thickness skin grafts.12 It would maximize the epidermal coverage derived from limited donor sites, avoid the morbidity associated with skin grafting, and eliminate the delays in waiting for the regrowth of donor sites before additional skin graft harvesting.

A major complication of deep burn wounds and large surface area wounds is the development of bacterial infections.13 Wound infections lead to delayed healing, longer hospital stays, and higher mortality.14 Service members in particular have a higher rate of infection-associated mortality following combat-related burns because wound excision and skin grafting is often delayed until they are evacuated to a burn center.15 The current methods used to control bacterial wound infections involve the use of topical antimicrobial agents such as silver nitrate soaks, silver sulfadiazine cream, and silver ion-release dressings, which demonstrate rapid and broad-spectrum antibacterial activity.16 Despite being clinically safe, silver antimicrobials are potentially cytotoxic, which may lead...
to impaired vascularization and epithelialization of the wound. In the present pilot study, we investigated the potential of regenerating skin using dermal substrates that contain human ASCs, keratinocytes, and silver nanoparticles (AgNP). Integra (Plainsboro, New Jersey) wound regeneration matrix was treated with AgNP, and the resulting antibacterial properties of the construct and the effect of the AgNP on ASC and keratinocyte proliferation was measured in vitro.

**MATERIALS AND METHODS**

**Isolation and Culture of Human ASCs**

Human ASCs were isolated from human adipose tissue samples obtained by liposuction. Briefly, samples were digested in 0.1% (w/v) collagenase P (Roche Diagnostics, Indianapolis, IN) for 25 minutes at 37°C. Cells were then freed by vigorous mixing and centrifuged at 300g for 5 minutes to separate the stromal vascular fraction from the fat cells. The stromal vascular fraction pellet was resuspended in growth medium (EGM-2MV; Lonza, Walkersville, MD), pipetted sequentially through 500-, 297-, and 70-μm-sized mesh and cultured in T-75 flasks at 37°C in an atmosphere of 5% CO₂. The growth medium was changed every 48 hours, and cells were passaged when 80 to 90% confluent. Cells were used for experiments between passages three and six.

**Isolation of Human Keratinocytes**

Human keratinocytes were isolated from discarded bits of split-thickness skin used in skin grafting. The pieces (2–3 mm wide) were incubated overnight in 1 mg/mL Dipase II (PluriSTEM; EMD Millipore, Billerica, Massachusetts) at 4°C. The epidermis was then separated from the dermis, minced and incubated in 0.25% trypsin-EDTA (Gibco, Grand Island, NY) for 15 minutes at 37°C. The trypsin was neutralized with KGM-Gold medium (Lonza, Walkersville, MD), pipetted sequentially through 500-, 297-, and 70-μm-sized mesh and cultured in T-75 flasks at 37°C in an atmosphere of 5% CO₂. The growth medium was changed every 48 hours, and cells were passaged when 80 to 90% confluent. Cells were used for experiments between passages three and six.

**Culture of ASC/Keratinocytes in Integra**

Integra wound matrix (Integra Life Sciences, Plainsboro, NJ) was cut into 8-mm-diameter discs using a cork bore and equilibrated with growth medium in a 48-well plate. Proliferating ASCs were trypsinized, resuspended in growth medium, and 5 × 10⁵ cells per disc (50 μL of 10⁵ cells/mL) were injected using a 25-gauge needle. One hour was allowed for ASCs to attach to the matrix, and then 3 × 10⁵ freshly isolated keratinocytes (6 × 10⁵ cells/cm²) were pipetted onto the surface. Each construct was cultured in 1 mL of keratinocyte growth medium at 37°C with a medium change every 48 hours. After 7 days, the constructs were transferred to the 12-mm inserts of a 12-well Transwell plate (3.0-μm polycarbonate membrane; Costar, Corning, NY) and cultured for an additional 7 days in the same medium. Wells and inserts were filled with 0.8 and 0.4 mL medium, respectively, to maintain an air–liquid interface at the surface of the constructs to promote keratinocyte differentiation.18

**Measuring Cell Proliferation in Integra**

The proliferation of cells in the Integra constructs was measured using the WST-1 assay (Roche, Indianapolis, IN). Briefly, discs of Integra were seeded with ASCs or keratinocytes and cultured as described earlier. The constructs were collected over time, washed with Hanks balanced salt solution (HBSS) containing calcium and magnesium, and then incubated in 240 μL of the WST-1 reagent (1:10 dilution with HBSS) for 4 hours at 37°C. Duplicate absorbance readings of the reaction medium were made at a wavelength of 450 nm and were corrected with background absorbance readings and a reference wavelength of 630 nm using the equation 

\[ \text{Absorbance} = (A_{450nm} - A_{630nm})_{\text{sample}} - (A_{450nm} - A_{630nm})_{\text{blank}}. \]

In some experiments, the Integra discs were preloaded with 0, 0.25, 0.5, or 1.0% AgNP (200 nm diameter), seeded with 10⁷ ASC/cm², cultured for 14 days at 37°C, and then measured with the WST-1 reagent.

**Loading Integra with AgNPs**

The AgNPs (75 or 200 nm; aqueous suspensions) were purchased from nanoComposix (San Diego, CA). The particles were resuspended by vigorous vortexing for 1 minute immediately before use. One square centimeter pieces of Integra were inverted onto sterile gauze to remove free water and then immersed into the AgNP suspension for 1 minute. Particles that remained unbound to the matrix were washed free by vortexing the samples in Dulbecco’s phosphate-buffered saline for 20 minutes. Samples were then sterilized by exposure to an ultraviolet C lamp and subsequently seeded with ASCs or used to measure bactericidal activity.

**Assessment of AgNP Deposition by Scanning Electron Microscopy**

Samples of Integra (5 × 5 mm²) were equilibrated in deionized water, and upon removal, the liquid in the matrix was wicked away using a laboratory wipe (Kimwipes; Kimberly-Clark, Roswell, Georgia). Next, samples were immersed in a 0.5% solution of AgNP (200 nm diameter) for 1 minute. Excess AgNP solution was wicked away from the AgNP-loaded Integra sample, and the sample was then allowed to dry completely at room temperature. The Integra was then imaged via scanning electron microscopy (SEM). For comparison, a sample of Integra without AgNP was also dried and imaged in the same manner.

**Measuring Bactericidal Activity of Integra–AgNP**

The log reduction assay was used to measure the ability of Integra containing AgNP to kill bacteria. Bacterial cultures of *Staphylococcus aureus* and *Pseudomonas aeruginosa* (ATCC 25923 and 27853) were recovered from frozen stock
Skin Regeneration Using Antibacterial Dermal Substrates

by overnight growth in tryptic soy broth at 37°C, and Klebsiella pneumoniae (ATCC 31488) was recovered by overnight growth in nutrient broth 3. Bacterial suspensions were diluted to an OD_{600nm} of about 0.2 immediately before use. Fifty microliters of inoculum was injected per square centimeter of Integra–AgNP and incubated at 37°C for 30 minutes to 24 hours in a humidified atmosphere. Controls consisted of using Integra without AgNP, 1 cm² pieces of Acticoat (Smith and Nephew, Hull, United Kingdom), which is a silver-coated wound dressing, and 1-cm² pieces of the Acticoat absorbant core (without the silver coating). At the appropriate times, the surviving bacteria were recovered by immersing the samples into neutralizing solution (1:10 dilution of inoculum) containing 0.4% sodium thioglycolate and 1% Tween 20. Samples were vortexed and serially diluted in phosphate-buffered saline PBS. 100 μL of each dilution was plated on Mueller Hinton II Agar (Becton Dickinson, Franklin Lakes, New Jersey), and colonies were grown at 37°C for 24 hours and then counted. Log reductions were calculated as log_{10} (CFU/mL)_{Integra} – log_{10} (CFU/mL)_{Acticoat}. Similarly, log reduction for Acticoat was calculated as log_{10} (CFU/mL)_{Acticoat} absorbant core – log_{10} (CFU/mL)_{Acticoat}.

Histology/Immunoperoxidase Staining
Integra constructs were washed two times in HBSS, embedded in Optimum Cutting Temperature compound (Tissue-Tek; Sakura-Finetek USA, Torrance, CA) medium and flash-frozen in liquid nitrogen. Twenty-micron-thick sections were stained with hematoxylin and eosin (H&E) for light microscopy, and 5 μm sections were processed using an immunoperoxidase stain against keratin AE1/AE3 by the Indiana University Health Immunohistochemical Lab.

Statistics
Linear regression analysis (Instat 3; Graphpad Software, La Jolla CA) was used to determine whether WST-1 absorbance of constructs containing ASC or keratinocytes increased with culture time to signify cell proliferation within the Integra matrix.

RESULTS AND DISCUSSION
Integra Supports ASC and Keratinocyte Proliferation
WST-1 absorbance readings are a measure of the amount of WST-1 reagent that is reduced by cell metabolism and therefore reflects the number of viable cells that are present in the Integra matrix. The metabolic activity of constructs containing human ASC or keratinocytes over increased over the time spent in culture (Fig. 1). The number of ASCs that were injected into the Integra was about two times the number of keratinocytes that were seeded on the surface. However, the finding that the absorbance readings were much higher for ASCs than for keratinocytes likely reflects a difference in plating efficiencies, which for freshly isolated keratinocytes on plastic is typically less than 5%. For fresh ASCs, the plating efficiency has been reported at about 22%.

Constructs containing both ASC and keratinocytes were examined histologically after 14 days, with the final 7 days of growth occurring at the air–liquid interface. The micrographs show that the ASCs were scattered throughout the volume of concentrated ASCs into the Integra versus pipetting a dilute suspension of keratinocytes onto its surface. Although our usual culture medium for ASCs is endothelial growth medium, the constructs containing both ASCs and keratinocytes were cultured in keratinocyte growth medium to support the proliferation of keratinocytes, which we have found to be the more difficult. The keratinocyte medium contained added serum and calcium to promote the differentiation of keratinocytes as previously reported.

The micrographs show the presence of a dense surface layer that stained strongly for keratin, but it is not clear if this represents differentiated or cornified cells or if it is an effect of dehydration due to the proximity of the water–air interface.

Integra Retains AgNP Following Immersion
Immersion of Integra into a 0.5% (w/v) AgNP suspension changed the Integra from colorless to a dark brown or black coloration (Fig. 3). Integra is made of type I collagen and chondroitin 6-sulphate in a 92:8 ratio. Ninety-eight percent of Integra by volume consists of fluid-filled pores that are 30 to 120 μm in diameter. When some of this large volume of fluid is wicked out of the matrix, its subsequent immersion into the AgNP suspension causes rapid imbibition of

FIGURE 1. Metabolic activity of Integra constructs containing adipose stem cells (○) or keratinocytes (●). Absorbance was calculated as (A_{630} – A_{450})_{construct} – (A_{630} – A_{450})_{blank}. The plotted values are the mean ± standard deviation of 2 to 4 replicate constructs. Slopes of regression lines are significantly different from 0 (●; p = 0.003; ○; p = 0.005).
FIGURE 2. (A) H&E staining of Integra cross sections shows the presence of adipose tissues (ASCs) and keratinocytes after 14 days. (B) Peroxidase staining with anti-cytokeratin AE1/AE3 shows the presence of keratinocytes. The ASCs were injected into the matrix and keratinocytes were pipetted onto the surface after 1 hour. The constructs were cultured at the air-liquid interface for the final 7 days. Bar = 200 μm.

FIGURE 3. Integra (A) was loaded with 200-nm-diameter AgNP (B) by simple immersion into a 0.5% (w/v) suspension for 1 minute. Representative scanning electron microscopy images of untreated Integra (C) and Integra–AgNP at (D) low magnification (x1,000) and (E) high magnification (80,000x).
the particles. SEM shows that the short exposure time was sufficient for the uptake and adherence of the particles to the collagen matrix (Figs. 3C–3E). Drobon et al studied AgNP in collagen solutions and found that AgNP bind to collagen mainly through strong electrostatic interactions with the carboxylate (COO−) and amine (NH2+) groups of the collagen.22 Moreover, particle-induced changes in collagen secondary structure occurred, from α-helix to the more compact β-sheet.22 In the Integra, the collagens are cross-linked by a glutaraldehyde treatment, which potentially prevents similar AgNP-induced changes to the collagen structure.

**Integra–AgNP Demonstrates Antibacterial Activity**

Figure 4 shows logarithmic reductions in the number of bacterial colonies that followed exposure of bacteria to Integra–AgNP. *S. aureus* and *P. aeruginosa* were injected into Integra that was pretreated with a 1% (w/v) suspension of 200-nm-diameter AgNP and incubated for 2 hours. Log reductions were less than 1 (0.10 and 0.44, respectively) and corresponded to decreases in colony-forming units (CFUs) of about 20% and 64%, respectively, when compared to Integra that did not contain AgNP (Fig. 4A). *K. pneumoniae* exposure to Integra–AgNP for 24 hours produced a log reduction of 0.33 or a decrease in CFUs of 54%. In comparison, a 30-minute exposure of *K. pneumoniae* to Acticoat wound dressing resulted in a log reduction of 3.1 (over a 1,000-fold decrease) when compared to a similar exposure to the Acticoat absorbant core that lacked the nanocrystalline silver sheet. Behm et al exposed *S. aureus*, *P. aeruginosa*, and *K. pneumoniae* to Acticoat for 30 minutes and measured reductions in CFUs of 3.7 to 4.7 logs,16 which is comparable to our measurements using Acticoat. Some Integra was pretreated with 75-nm-diameter AgNP, injected with *K. pneumoniae*, and incubated at 37°C for 24 hours. Figure 4B shows that there was an increasing trend in log reduction as the AgNP concentration increased between 0.1 and 1.0%. A reduction of 3 or more logs is considered “bactericidal.”16 Our Integra–AgNP constructs, on the other hand, may be defined more appropriately as “bacteriostatic.”

The antimicrobial activity of silver is due to its ionic form (Ag+) and has been attributed to several mechanisms, including the blocking of cellular respiration,23 binding to bacterial membranes to disrupt membrane function,24 and binding to and condensing bacterial DNA to inhibit cell replication.25 AgNP are composed of metallic silver, which is neutral, and the associated antimicrobial activity has been attributed to its release of ionic silver.26 However, some studies have implicated the particles themselves as the source of these toxic effects.27,28 The activity of AgNP is inversely proportional to their diameter as the surface to volume ratio of the particles increases with decreasing radius, thereby increasing the release of Ag+ ions. In our studies, the log reduction following exposure to 75-nm-diameter AgNP was about two times that of the 200-nm-diameter particles (0.66 vs. 0.33 CFU/mL; Fig. 4).

**Integra–AgNP Does Not Affect ASC Proliferation**

Integra samples were preloaded with 200-nm-diameter AgNP. Unattached particles that filled the pores of the matrix were washed out by vortexing the samples in Dulbecco’s phosphate-buffered saline. Samples were then injected with 10^7/cm^3 ASCs and cultured in EGM-2MV for 14 days. The metabolic activity of the constructs was measured using the WST-1 assay. The absorbance readings were not significantly different when the Integra was preloaded with 0, 0.25, 0.5, or 1% AgNP (Fig. 5), hence cell proliferation was not affected by the use of up to 1% (w/v) AgNP. The H&E staining of the constructs after 14 days shows that ASCs were attached to
the matrix and evenly distributed throughout for all concentrations of AgNP (Fig. 6). No AgNP or AgNP aggregates were present in the pores, and it demonstrates that the wash-out step was effective. However, the eosin staining of the collagen is slightly darker or appears to have a slight “dusty” appearance in samples that were treated with 1% AgNP (Fig. 6D). The wash-out step eliminated a significant source of toxicity (data not shown) and therefore is an important part of the sample preparation.

Cytotoxicity generally correlates with silver dissociation. Le Duc et al measured the effect of topical antiseptics on human skin substitutes in vitro and found that Acticoat decreased metabolic activity to about 60% of untreated control samples. However, decreases in metabolic activity due to the presence of AgNP did not equate with a loss of cell viability in vitro and did not inhibit wound healing, although aggregates of AgNP became localized in the cytoplasm of dermal fibroblasts of healed wounds. The implications of these reports and our findings are that higher concentrations of AgNP in Integra and increased antibacterial activity may be achievable before metabolic suppression, loss of cell viability, and impairment of wound healing are likely to occur.

CONCLUSIONS
Integra wound matrix was made bacteriostatic by the addition of AgNPs. The new antibacterial properties did not affect the proliferation of ASCs or keratinocytes that were cultured in the matrix. An expanded study of these properties at higher AgNP concentrations and their effects on oxidative stress, inflammation, vascularization, and epithelialization of wounds is expected to further enhance the performance of this matrix and increase its potential utility in current and future treatments of deep burn injuries.

FIGURE 5. Metabolic activity of Integra constructs that were pre-equilibrated in AgNP suspensions (200 nm diameter; 0, 0.25, 0.5, or 1% w/v), injected with $10^7$ ASC/cm$^3$ and cultured for 14 days. WST-1 absorbance values did not vary appreciably with exposure to the different AgNP concentrations.

FIGURE 6. H&E staining of cross sections of Integra that was preloaded with (A) 0%, (B) 0.25%, (C) 0.5%, or (D) 1.0% AgNP, injected with $10^7$ ASCs/cm$^3$ and cultured for 14 days. Bar = 200 μm.
ACKNOWLEDGMENT

This work was internally funded.

REFERENCES

Regenerative and Antibacterial Properties of Acellular Fish Skin Grafts and Human Amnion/Chorion Membrane: Implications for Tissue Preservation in Combat Casualty Care

Skuli Magnusson, BSc†; Baldur Tumi Baldursson, MD, PhD‡; Hilmar Kjartansson, MD‡; Ottar Rolfsson, PhD‡; Gudmundur Fertram Sigurjonsson, MEng*

ABSTRACT Background: Improvised explosive devices and new directed energy weapons are changing warfare injuries from penetrating wounds to large surface area thermal and blast injuries. Acellular fish skin is used for tissue repair and during manufacturing subjected to gentle processing compared to biologic materials derived from mammals. This is due to the absence of viral and prion disease transmission risk, preserving natural structure and composition of the fish skin graft. Objectives: The aim of this study was to assess properties of acellular fish skin relevant for severe battlefield injuries and to compare those properties with those of dehydrated human amnion/chorion membrane. Methods: We evaluated cell ingrowth capabilities of the biological materials with microscopy techniques. Bacterial barrier properties were tested with a 2-chamber model. Results: The microstructure of the acellular fish skin is highly porous, whereas the microstructure of dehydrated human amnion/chorion membrane is mostly nonporous. The fish skin grafts show superior ability to support 3-dimensional ingrowth of cells compared to dehydrated human amnion/chorion membrane (p < 0.0001) and the fish skin is a bacterial barrier for 24 to 48 hours. Conclusion: The unique biomechanical properties of the acellular fish skin graft make it ideal to be used as a conformal cover for severe trauma and burn wounds in the battlefield.

INTRODUCTION

A shift in the nature of injuries from a traditional battlefield pattern to fighting against insurgents who use improvised explosive devices has brought burn injuries to the forefront of care for injured servicemen. Furthermore, new types of directed energy weapons are being developed that are likely to call for improved therapies for thermal wounds. As the fighting patterns evolve, the need for new and improved therapies in combat casualty care (CCC) for wounds and tissue repair that can be used immediately on the battlefield increases. Nearly, 90% of combat-related deaths occur before arrival at combat field hospitals or transfer to an echelon V hospital capable of providing definitive care. The time it takes for wounded soldiers to be evacuated out of theater varies depending on location, evacuation from Iraq usually takes 48 to 72 hours.

Cadaver skin is the preferred burn wound cover for severe burns and is frequently used in the hospital setting. The use of cadaver skin is, however, impractical in battlefield and other austere environments. In order to respond to this unmet clinical need, attempts have been made to develop several new types of skin substitute in recent years. The main types of skin substitutes are either cellular or acellular. Cellular skin substitutes are not practical for use in a combat theater, as they either need special stabilization fluids with short shelf life or be frozen in liquid nitrogen. Acellular mammalian-derived products possess prolonged shelf life as compared to cellular skin substitutes but per Food and Drug Administration requirements have to undergo “viral inactivation.” Viral inactivation is performed with detergents that remove all soluble components from the tissue leaving behind an inert matrix of collagenous structure. The harsh viral inactivation processing removes lipids, glycans, elastins, hyaluronic acid, soluble collagen, and other important biological components from the tissue that potentially are beneficial to wound healing.

Acellular fish skin is remarkably similar to human skin, yet fundamentally different from mammalian-derived matrices, because of the preservation of structure, lipids, and other soluble components. Mammalian scaffolds require harsh chemical processing to reduce viral and prion transmission risk, such risk from the Atlantic cod (Gadus morhua) to humans is nonexistent. Fish skin grafts are subjected to gentle processing where structure and bioactive composition, including omega-3 polyunsaturated fatty acids is preserved. Studies have shown that omega-3 fatty acids possess antiviral7 and antibacterial8 properties and also act as regulators of inflammation.

Double-blind randomized clinical trials have shown that wound treatment with acellular fish skin allows for significantly faster wound closure than porcine small-intestinal-derived matrices. The acellular fish skin is currently being used in a regulatory approved and patented wound treatment product being marketed in the United States and in Europe under the brand name Kerecis Omega3 (Kerecis, Isafjordur, Iceland). After grafting, the fish skin is incorporated into the damaged area and infiltrated by autologous cells that convert...
the graft into functional, living tissue while the graft slowly breaks down. The fish skin graft has been used to treat a large number of wounds of various etiologies, both acute and chronic.11,12

Here, we set out to investigate properties of the acellular fish skin that are important for tissue repair and regeneration. Employing microscopy-based techniques, we compared the fish skin to dehydrated human amnion/chorion membrane (dHACM) allograft, with respect to structure and ability to sustain fibroblast ingrowth in culture. Furthermore, we investigated the ability of the acellular fish skin to resist bacterial invasion.

METHOD

Cell Culture

To assess cell ingrowth capability of the biologic materials, the scaffolds were seeded with cells. We conducted all cell culture experiments at 95% humidity and with 5% CO2. For all cell culture experiments, we used high glucose Dulbecco’s Modified Eagle’s medium (DMEM) (D6429; Sigma Aldrich, St. Louis, Missouri) with 10% fetal bovine serum (Cat. 10270-106; Thermo Fisher Scientific, Waltham, Massachusetts) and 1% Antibiotic-Antimycotic solution (Cat. 15240-062; Life Technologies, Carlsbad, California). Before seeding the biologic materials, we prehydrated them in DMEM for 24 hours and incubated in cell culture cabinets. The punches we made from the biologic materials were 6 mm in diameter, roughly the same size as the wells (0.32 cm²) of a 96 well plate. We cultured NIH 3T3 mouse embryonic fibroblast cell line (ATCC, CRL-1658). We split the cells when they reached 70 to 80% confluence and then seeded each biologic material with 16,000 cells in 50 μL volume of DMEM. After four hours, we added 150 μL of medium bringing the final volume to 200 μL in each well. Biologic materials that did not receive cells served as negative controls. After 24 hours, we transferred the biologic materials to 12-well plate and incubated in a cell culture hood for 12 days. Medium was changed every 3 days and wells checked for contamination. After 12 days, we harvested the samples and stained them after fixation.

Fixation

First we fixed the samples in 10% formalin for 48 hours. Next we placed the samples in a fixation machine (Tissue-Tek VIP 6; Sakura, Torrance, California) with the following program: 10% formalin: 2 hours, 70% alcohol: 2 hours, 90% alcohol: 2 hours, 100% alcohol: 2 hours, xylene: 2 hours. Lastly, we embedded the samples in paraffin for 2 hours. We cut 2 μm slices and put them on a slide in a 70°C heat cabinet for 30 minutes. Next, we embedded the slides in xylene for 15 minutes and washed with 100% alcohol for 5 minutes and water for 2 minutes.

Hematoxylin and Eosin Staining

We embedded slides with fixed samples with hematoxylin: 8 minutes, water: 2 minutes, 0.02 M HCL in 70% alcohol: 2 minutes, water: 4 minutes, eosin: 2 minutes, 100% alcohol: 14 minutes. Lastly, we glued cover slips on top of the slides.

Periodic Acid Schiff Staining

Samples are deparaffinized and hydrated in water then added to 0.5% periodic acid solution for 5 minutes and then Schiff’s reagent for 15 minutes. We used Mayer’s hematoxylin as a counter stain.

Masson Trichrome

Samples are fixed in 100% alcohol, 95% alcohol, 70% alcohol, Weigert’s iron hematoxylin working solution, and Bielbrich scarlet-acid in solution for 10 to 15 minutes. Sections are transferred to aniline blue solution and stained for 5 to 10 minutes. Lastly, the slides are rinsed briefly in distilled water and differentiated in 1% acetic acid solution.

Fluorescent Staining

Before staining, we used hydrophobic barrier pen to surround the fixed tissue on the slides. We diluted 6 μL of nuclear stain (NucBlue, Cat. R37605; Thermo Fisher Scientific) with 994 μL of 1x phosphate-buffered saline buffer with 1% Triton X-100 (Sigma Aldrich). We added 30 μL of the fluorescent staining solution to the slides and incubated at 37°C for 30 minutes. We washed the slides 3 times with 1x phosphate-buffered saline and lastly we added mounting medium (Fluormount, F4680; Sigma Aldrich) to the samples and glued cover slips on top.

Microscopy

We used a Field Emission Scanning Electron Microscope (Zeiss Supra 25, Oberkochen, Germany). The software program used to run the microscope is Smart SEM. The unit we used to gold coat the specimens is an Edwards S150B sputter coater. We coated the specimens for 2 minutes and afterward we mounted them on aluminum stubs with carbon tabs. We used light microscope (Leica DM-i6000B; Leica-Microsystems, Wetzlar, Germany) to observe hematoxylin and eosin (H&E)-stained samples and confocal microscope (FV1200; Olympus, Tokyo, Japan) to observe samples stained with fluorescent stains.

Porosity and Thickness

We counted the number of pores per 100 μm² of sample and measured the diameter of the pores using the software ImageJ (Wayne Rasband, National Institute of Health, Bethesda, Maryland) and scale bar on the image. We did not count pores with a diameter of 1 μm or less. ImageJ was also used to calculate the diameter of pores from the scanning electron microscope images.

Cell Ingrowth Quantification and Statistical Analysis

We quantified the cell ingrowth by counting fluorescently labeled cells within the biologic materials using the software.
We divided the images into columns with their width corresponding to 100 μm of sample and used the free hand tool in ImageJ to define the boundaries of the materials. Cells were considered ingrown when observed within the material. We only counted cells observed within the biologic materials but excluded the cells on top of the materials (Fig. 3). The mean values and standard deviations as well as all other statistical analysis were performed using the statistical application R.3.2.3 (Bell Laboratories, Murray Hill, New Jersey). All groups were tested for a normal distribution by the Shapiro–Wilk method and statistical methods were chosen according to those results and the group dynamics. Wilcoxon
rank sum test with continuity correction was used to analyze significant differences between groups when at least one of the groups did not contain normally distributed data.

**Bacterial Barrier**

We prepared an overnight culture of *Staphylococcus aureus* (ATCC 25923) and diluted it after thorough mixing for 60 seconds on a vortex mixer. From the overnight culture, we diluted 1.0 mL into 9.0 mL of broth repeatedly to make 10-fold dilutions from the maximum recovery diluent (MRD). From each dilution, we transferred 1.0 mL into separate petri plates and added 20 mL of melted Tryptone soya agar at 45°C. We incubated the plates at 35.0 ± 1.0°C for 48 hours. After the incubation, we counted colonies on the plates and calculated the number of colony-forming units in the initial suspension. We transferred 200 μL of precultured bacteria

![Image](image1)

**FIGURE 2.** Fibroblasts infiltrate and remodel the fish skin graft but do not interact with human amnion/chorion membrane. H&E staining of (A) fish skin graft and (B) dHACM. Cross sections are shown. Fibroblast cells infiltrate and remodel the fish skin graft following 12 days of culture. Cells infiltrate dHACM material to a much lesser extent and are primarily observed on the top layer of the dHACM material. Scale bars are 50 μm.

![Image](image2)

**FIGURE 3.** Cell ingrowth is clearly seen into the fish skin graft but to a much lesser extent to the human amnion/chorion membrane allograft. Fibroblasts were seeded onto fish skin graft and dHACM that were then imaged following 12 days of culture in the presence of a fluorescent antibody against nuclei (blue). Both the fish skin graft and the dHACM emit green autofluorescence at certain wavelengths. Cells are clearly observed to infiltrate the thicker (A1) fish skin graft, whereas fewer are observed in (B1) dHACM. The lower row (A2 and B2) shows cellular quantification in the respective grafts that are summarized in Figure 4. Scale bars are 100 μm.
Regenerative and Antibacterial Properties of Acellular Fish Skin Grafts

FIGURE 4. Significantly more cell ingrowth into acellular fish skin compared to human amnion/chorion membrane. Comparison of cellular infiltration into acellular fish skin graft compared to human amnion/chorion membrane allograft. There’s significant difference between cell ingrowth into the fish skin graft as compared to the dHACM material (n = 7) (p < 0.0001).

corresponding to 7,000 cfu into 1 of the 2 chambers unit. Pieces of membranes made of fish skin to be tested were cut to cover the insert with sterile scissors and placed between the insert units and a clamp was used to screw the 2 inserts tightly together. The cut was 1 mm outside the edge of the insert. We injected sterile tryptic soy broth (TSB) into 1 of 2 inserts and TSB with a bacterial strain into the other. Control samples containing only TSB were prepared. The chambers were incubated at 37.0 ± 1.0°C. Each sample was analyzed in triplicate as were the control samples. Recovery tests were done after 4, 24, and 48 hours. From the originally sterile TSB chambers, 1.0 mL was transferred into 9.0 mL of maximum recovery diluent and further 10-fold dilutions were made as described earlier. The plates were incubated at 35.0 ± 1.0°C for 48 hours. After the incubation, the colonies on the plates were counted and the number of colony-forming units in the initial suspension calculated. The same amount of sterile TSB was added into the unit to secure that the surface was always the same.

RESULTS AND DISCUSSION

Microstructure of acellular fish skin is highly porous and the microstructure of human amnion/chorion membrane is mostly nonporous. Initially, we wanted to compare the microstructure of the biologic materials and evaluate if the porosity was well suited for cell ingrowth. The fish skin graft is highly porous compared to human amnion/chorion membrane. The fish skin contains 16.7 holes (n = 6) per 100 μm on average, whereas the dHACM contains 1.7 holes (n = 6) per 100 μm on average (Fig. 1). The average diameter of the pores in the fish skin graft is 16.1 μm (n = 5) and 1.3 μm in the dHACM (n = 3). The thickness of the dHACM is 20.4 μm (n = 6) on average and the thickness of the fish skin graft is 450 μm (n = 6) on average (Fig. 1).

Acellular fish skin graft shows superior cell ingrowth when compared to human allograft. After evaluating the porosity of the biologic materials, we hypothesized that the fish skin grafts are well suited to support cell ingrowth since the pore size is within the range of typical cell size. Hematoxylin and eosin staining indicates that the fibroblasts do remodel the fish skin graft as they migrate and proliferate into the graft. The fibroblasts, however, form a layer on top of the dHACM allograft material and do not interact with the graft to the same extent (Fig. 2). The acellular fish skin grafts showed significantly (p < 0.0001) more 3-dimensional ingrowth of cells when compared to the dHACM (Figs. 3 and 4).

The acellular fish skin grafts can withstand bacterial invasion for up to 48–72 hours. In addition to regenerative capabilities, biologic materials used on battlefield wounds should, ideally possess antibacterial properties. Therefore, the fish skin graft was tested for its ability to resist bacterial invasion. The bacterial properties of the fish skin graft were tested by constraining the material between a 2-chamber unit and adding medium with bacteria on the other side and medium only in the second chamber. The fish skin graft can act as a bacterial barrier for up to 48–72 hours, n = 3 and when the fish skin graft supplemented with additional amount of omega-3, its bacterial barrier properties are augmented (Table I). Previously, it has been demonstrated that omega-3 PUFA have antibacterial properties against multiresistant bacteria and they might play a key role for the ability of the fish skin graft to act as a bacterial barrier.

Acellular fish skin has the potential to become a battlefield wound treatment of choice. The first few hours of CCC is vital for saving lives. An improvement in advanced field trauma wound treatment may play a large role in reducing morbidity, increased the viability of injured tissues. Burn injuries are often associated with significant complications such as localized or systemic infections, cosmetically unacceptable scar formation, and incomplete wound healing. Thermal injuries due to improvised explosives were the most serious threat to soldiers operating in Afghanistan and Iraq. Those burn patients face many complications, including supplementary injuries from multiple fragment wounds, high rate of infection, and long rehabilitation and hospitalization time. Cadaver skin is the preferred burn wound cover for severe burns and is frequently used in the hospital setting. The use of cadaver skin is however impractical in battlefield and other austere environments.

Overall, the results presented here show that acellular fish skin grafts act as a bacterial barrier and possess superior
The fish skin graft withstands bacterial invasion for up to 48–72 hours (n = 3)

<table>
<thead>
<tr>
<th>n</th>
<th>24 Hours</th>
<th>36 Hours</th>
<th>48 Hours</th>
<th>72 Hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>4.7</td>
<td>NA</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0</td>
<td>8.1</td>
<td>NA</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>8.2</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

This table summarizes the results from the 2-chamber assay with the fish skin graft, given in log(colony forming units/mL). The fish skin graft was able to act as a bacterial barrier for up to 48–72 hours. By increasing the omega-3 content of the fish skin graft the bacterial barrier properties are augmented.

There is an unmet gap and the potential of the fish skin graft exists for the development of better treatments, which will significantly improve functional and cosmetic outcomes in military service member suffering from severe burn or trauma wound victims resulting in reduced morbidity and mortality and better quality of life. Here, we have described the characteristics of the acellular fish skin graft that has the potential to fill this unmet gap.

ACKNOWLEDGMENTS

We want to thank the staff of the government research institute Matis for their contribution on the bacterial studies. This research was funded by the Office of Naval Research and Kerecis Limited, Eyrargata 2, 400 Isafjordur, Iceland.

REFERENCES

COL Shawn Kane, USA
James Karaszkiewicz, PhD
Jim Karesh, MD
Rahul Kashyap, MD
Hilmari Kjartansson, MD
Stephanie Kolakowsky-Hayner, PhD
Erik Kulstad, MD
Anthony LaPorta, MD
Srinivas Laxminarayan, MD
Frank Lebeda, PhD
Lt Col Denise Lennon, USAF
Amber Linde, PhD
Wei Liu, DSc
CJ Maldonado, PhD
COL Richard Malish, USA
Trudy Mallinson, OT
John Malone, MD
Rachel Markwald, PhD
Kacey Marra, PhD
Suzanne Martin, PsyD
Genny Maupin, MPH
Stephanie Maxfield-Panker, PhD, DPT
Geoffrey May, MD
Mary McCarthy, PhD
Joe McDonough, PhD
James McGhee, MD, MPH
Carolyn McGregor, PhD
LTC Dennis McGurk, USA
Francis McVeigh, MS
Thomas Merrill, PhD
Karen Mundy, PhD
Remington Nevin, MD
Omar Nunez Lopez, MD
LTC Jeremy Pamplin, USA
Paula Paradis, PT, MS, DPT, MBA, MHA
COL Robert Paris, USA
Charles Peterson, MD
CDR Travis Polk, USN
LTC Dan Porter, USA
Andrzej Przekwas, PhD
Christopher Râbago, PT, PhD
Susan Rasmussen, PhD
Vivek Raut, PhD
Jacques Reifman, PhD

MAJ Daniel Rhon, USA
COL Luis Rivero, USA
Farrukh Rizvi, PhD
MAJ Samantha Rodgers, USA
Lloyd Rose, PhD
Gregory Rule, PE
Steve Schauer, DO, RDMA, FAAEM
MAJ Matt Scherer, USA
Barri Schnall, MPT
COL Martin Schreiber, USA
Christopher Scully, PhD
Lt Col Thomas Shaak, USAF
Anita Sheoran, PhD
Lt Col Antoinette Shinn, USAF
Kanwarpal Singh, PhD
Teresita Sotomayor, PhD
Michael Steketee, PhD
Leonard Temme, PhD
Samuel Tisherman, MD
Ivo Torres Filho, MD, PhD
Yolandí van der Merwe, PhD
Eileen Villasante, PhD
LT Eric Vorm, USN
LCDR Peter Walker, USN
Heuy-Ching Wang, PhD
Kelly Watts, AuD
MAJ Marissa Weber, USA
LCDR Timothy Welsh, USN
Fletcher White, PhD
Mark Williams, PhD
MAJ Adam Willis, USAF
Yvette Woods, PhD
Nikki Wooten, PhD, LISW-CP
Don Workman, PhD
Gloria Workman, PhD, ABPP
CDR Abigail Yablonsky, USN
LTC Kathy Yancosek, USA
Shiming Yancosek, USA
Zheng Zhang, PhD