# Army Medicine Peer-Reviewed Publications

## November 2012

**INDEX**

## Executive Summary

November 2012 -- November featured heavy journal coverage on PTSD research with the details highlighting amygdala volume, perceived control of stressors, alcohol dependence, panic disorder, maternal illness and Asian American and Pacific Islander community susceptibility and response.

A considerable number of journals observed issues related to substance abuse with an especially relevant study on the abuse of energy drinks and the lack of sleep associated with the overconsumption of the beverages. Additionally, substance abuse also touched on performance enhancing drug abuse. Suicide research focused exclusively on the impact of the global economic downtown, with both studies indicating an increase of suicides during the research periods. An analysis on the potential effects of full EHR integration highlighted weaknesses in the proposed program, but offered solutions to these weaknesses. Miscellaneous coverage featured articles on geographic-specific illnesses, HIV infection in military personnel, aortic valve replacement surgery follow-up treatment and cancer incidence globally.

## Medical Journal Coverage

***Nutrition and Fitness***

### JAMA: Multivitamins in the Prevention of Cardiovascular Disease in Men

The study set out to determine whether long-term vitamin use among men would significantly decrease the risk of cardiovascular events. After a follow-up of approximately 11 years, the study saw 1,732 cardiovascular events with no significant effect of multivitamins on major cardiovascular events. Given October's Military Medicine article on the lack of awareness surrounding multivitamins, the supposed suggestion of multivitamins, especially as a preventive measure for cardiovascular health, should be reconsidered.

### PLOS Medicine: Leisure Time Physical Activity of Moderate to Vigorous Intensity and Mortality: A Large Pooled Cohort Analysis

Whether physical activity during times of leisure has an impact on premature mortality was the centerpiece of the study which measured the extension of life beyond 40 years. The study accounted for varied levels of activity based on age and BMI and found in

the pooled population that certain activities extend life expectancy. As the Army seeks to promote comprehensive health and nutrition, recognizing the positive role that leisure activity, especially in combat zones, can have on individual life expectancy is essential.

***Physiological Injuries and Surgery***

## Military Medicine: Incidence and Risk Factors for Acute Low Back Pain in Active Duty Infantry

Though copious studies have been performed on back pain, few have focused on lower back pain among infantrymen, a group within the military population that is tasked with carrying heavy loads over long distances. This study found that infantrymen had significantly lower instances of lower back pain when compared to their peers. One possible reason for this could be this population's tendency to not seek treatment for lower back pain or to normalize lower back pain as an inevitable result of their work. Another cause may be the high level of fitness and training that may prevent back pain. Additionally, fewer Marine infantrymen displayed lower back pain than infantrymen in the Army and junior enlisted had less lower back pain than senior enlisted. Risk for lower back pain increased with age and senior enlisted rank.

## International Orthopaedics: Long-term results of cementless hip arthroplasty with ceramic-on-ceramic articulation.

In a review of hip replacements, the study found positive results (including very good, good, and satisfactory) in 92 percent of cases and the 12-year prosthesis lifespan reaching nearly 84 percent. The adoption of the Mittelmeier prostheses is encouraged for Army use because of its positive reviews and long lifespan.

## Military Medicine: Changes in Hearing Thresholds as Measured by Decibels of Hearing Loss in British Army Air Corps Lynx and Apache Pilots

This study compared hearing loss of pilots flying primarily Lynx helicopters to those who exclusively fly an Apache. Results indicate that hearing loss is drastically less widespread than predicted, suggesting that the circumaural earmuffs issued to pilots are effective in limiting noise-associated hearing loss.

## Military Medicine: Maggot Debridement Therapy in Modern Army Medicine: Perceptions and Prevalence

An increasing resistance to antibiotics has created a need for Maggot Debridement Therapy (MDT) to remove decaying and infected flesh from open wounds. A survey of Army physicians found that 83 percent were familiar with MDT and, of those, 63 percent were aware that the FDA approves it but only 10 percent had used the treatment. The three most prominent reasons why physicians had not used the treatment was that there was no need, no access and they had no previous experience. Researchers indicated a need for more studies on MDT, as they noted the nature of warfare and antibiotic resistance points to an increasing need for debridement alternatives.

***PTSD***

## Archives of General Psychiatry: Amygdala Volume Changes in Posttraumatic Stress Disorder in a Large Case-Controlled Veterans Group

This study of brain physiology and PTSD found that those with PTSD had a smaller left amygdala, right amygdala and left hippocampus than those without PTSD, regardless of their traumatic experience. This is the first study to provide clear evidence that certain people are more susceptible to PTSD, and that the military can, for the first time, pre-screen service members to determine their ability to witness traumatic events without developing PTSD. This research received some coverage in the media market and is likely to receive additional coverage.

## Scientific American: It's not the stress that counts, it's whether you can control it

The study alleges that it isn't the type of stress that necessarily matters but it's the level of control one can exert on the stress. The control over certain stressors can impact future stresses by protecting against these future detriments, also known as "behavioral immunization." The study sought to determine whether the infralimbic or prelimbic cortex was responsible for the behavioral immunization phenomenon: Using electrophysiology, the scientists determined that the escapable stress response began in the prelimbic cortex. The study is directly applicable to human stress, a major concern to Army medicine, as future research can hone in on the prelimbic cortex to treat escapable stressors.

## Behavioral Modification: Cue-elicited affect and craving: advancement of the conceptualization of craving in co-occurring posttraumatic stress disorder and alcohol dependence.

Looking at the intersection of post-traumatic stress disorder and alcohol dependence, the study examines 1) the responses of the two conditions to stimuli for individuals seeking treatment, 2) the relationship between the elicited response and the craving for alcohol, and 3) the type and severity of craving as defined by the researchers. The most negative effect was seen in the trial combining trauma image and alcohol cues, and a positive affect was seen following alcohol cues supporting the respective hypotheses. Trauma cues, however, do not seem to increase positive affect after removing certain items related to anxious arousal. As the Army seeks to understand the motivation behind both traumatic stress and alcohol dependence, the study offers insight into the relationship between stimuli that affect the two.

## European Journal of Psychotraumatology: Fear of memories: the nature of panic in posttraumatic stress disorder.

Underlining the frequency of panic attacks in post-traumatic stress disorder sufferers, the study found sufferers of panic disorder (PD) more likely to report numbing or tingling than PTSD participants and they were more likely to report an increased fear of passing out. PTSD participants were reportedly more likely to relive a terrible event, injurious memories from the past, memories that drive them crazy, imprisonment in past situations, or a failure to block bad memories. The nuanced differences between PTSD and PD highlight the need for unique approaches by Army medicine to apply differing treatment measures.

## BMC Pregnancy Childbirth: A systematic review of the relationship between severe maternal morbidity and post-traumatic stress disorder.

A compilation of eleven relevant studies which met review criteria were analyzed to determine a relationship between the increased rate of maternal morbidity in high-income countries. The study seeks to close the gap of a lack of information of these instances of morbidity on postnatal psychological health, especially PTSD. Findings of a relationship between severe morbidity and PTSD or PTSD symptoms was inconsistent across the studies; however, pre-eclampsia was specifically targeted as a known risk factor for PTSD and its symptoms.

## Military Medicine: Mental Health of Asian American and Pacific Islander Military Veterans: Brief Review of an Understudied Group

Due to a lack of research on the Asian American and Pacific Islander veteran community, the study focused on the mental health of AAPI veterans. The review of prior studies found racism within the community who served during the Vietnam War, where these race-related issues pointed to more severe forms of PTSD symptoms. Anecdotally, AAPI veterans serve the longest and identify the most as being a veteran compared to other racial/ethnic groups. AAPI veterans are also less likely to experience homelessness In terms of physical health, AAPI veterans were, on average, healthier than their non-AAPI counterparts; however, they were less apt to take advantage of mental health services. As Army messaging seeks to remove the stigma of seeking out mental health services, the AAPI community should be a top priority.

***Sexual Health***

## Medscape: Military Sexual Trauma Complicates PTSD for Female Veterans

This interview with Ursula Kelly, PhD and assistant professor of Nursing at Atlanta Veterans Affairs Medical Center, discusses her research on the factors that may prevent female Soldiers who suffer from military sexual trauma from seeking treatment for resulting PTSD. Kelly notes that while men reject treatment because of the stigma associated with PTSD, women don't seek treatment because of the stigma associated with military sexual trauma. Kelly also said that female Soldiers will seek help when multiple sources, such as family members, friends and peers encourage them to do so and when they begin to realize that their behavior is not normal.

***Substance Abuse***

## BMC: A longitudinal assessment of alcohol intake and incident depression: the SUN project

Concerned with the effect of alcohol on depression in women, the study sought to determine the link between depression and alcohol abuse or alcohol dependence. Accounting for the variance of alcohol consumption within specific communities, the study separated the type of alcoholic beverage, quantity consumed, and the pattern of consumption. Finding that moderate consumption protects women against depression, there was no relationship found between higher consumption and the same protection. For Army purposes, the understanding of the unique relationship between alcohol intake and depression in females provides the nuance necessary to prescribe certain programs targeting women.

## Centers for Disease Control and Prevention: Energy Drink Consumption and Its Association with Sleep Problems Among U.S. Service Members on a Combat Deployment — Afghanistan, 2010

This study found that 45 percent of service members deployed in Afghanistan drank at least one energy drink per day with 14 percent drinking three or more per day. Those who consumed three or more drinks had a significantly higher percentage of sleep problems than their peers and were more likely to report having less than 4 hours of sleep per night. Those who consumed energy drinks were also more likely to fall asleep during briefings or while on guard duty and displayed a higher level of stress. Those who consumed the drinks in moderation, one to two per day, reported the same number of sleep issues as those who did not use caffeine. Researchers recommend that leaders educate service members on the potential effects of using too many energy drinks including withdrawal, caffeine intoxication or sleeplessness, noting that using too many energy drinks can lead to poor work performance and that the long term effects are unknown. This research received some coverage in medical trade publications.

## Annals of Internal Medicine: Behavioral Counseling After Screening for Alcohol Misuse in Primary Care: A Systematic Review and Meta-analysis for the U.S. Preventive Services Task Force

Alcohol misuse is the leading cause of preventable death in the United States, and researchers sought to determine the benefits and harms of behavioral health counseling and interventions on adults who misuse alcohol. Results indicate that multiple brief 10 to 15 minute interventions provided the best results with consumption decreasing by 3.6 drinks per week. Researchers found no evidence of harms. These findings indicate that brief but meaningful behavioral health interventions can help alleviate alcohol misuse in the Army and can be implemented in training barracks where alcohol misuse is widely reported on.

## British Journal of Sports Medicine: Performance enhancing drug abuse and cardiovascular risk in athletes: implications for the clinician

This literature review attempts to determine the health consequences illegal drugs have on cardiovascular health among athletes. Illegal drug use is linked to critical cardiovascular health, but more research should be undertaken to determine the actual effects. Healthcare providers should be able to determine the signs of drug use to help patients receive treatment and recovery. Though this review focuses on drug's effects on athletes, its findings can be applied to high-performance Soldiers.

***Suicide***

## Lancet: Increase in state suicide rates in the USA during economic recession

Comparing the rate of suicide from European countries to that of the US during the economic downturn, the study found that the rate of suicide in the US increased during the timeframe in question. The study found an additional 1580 suicides per year over and above the regular rate of suicide. Going further, the researchers looked into the relationship between unemployment and suicide where results indicated a 1 percent increase of unemployment caused a .79 percent increase of the European suicide rate and a .99 percent increase in the American. In terms of Army concern, stemming the rate of veteran unemployment can significantly affect the veteran suicide rate.

## Medscape: Sharp Increase in US Suicides Blamed on Financial Crisis

This article discusses the link between the 2007 world financial crisis and a worldwide increase in suicides. It does not refer to the interplay between financial worries, mental health and suicide among those in the military population, but research on this topic could add weight to Army messaging that the issues within the military mirror those of the civilian population.

***Traumatic Brain Injury***

## Medscape: Herbicide Exposure/TBI Combination Triples Parkinson's Risk

This article notes the enhancing capabilities that TBI and exposure to the widely used herbicide, paraquat, have on Parkinson's risk. This article brings up an interesting topic that may affect Soldiers, namely how the comingling of TBI and exposure to certain toxic chemicals may predispose Soldiers to certain diseases including Parkinson's. As more Soldiers who are diagnosed with TBI age, they may become sensitive to certain chemicals and, consequently, certain diseases.

## Military Medicine: The Influence of Sleep and Mood on Cognitive Functioning Among Veterans Being Evaluated for Mild Traumatic Brain Injury

This research supports previous studies indicating that poor sleep quality and quantity critically affects those with mental health issues. Poor sleep satisfaction among veterans with mild TBI was related to general distress, depression and anxiety and was

predictive of cognitive ability. Sleep satisfaction correlated with memory but not attention, indicating that those suffering from mild TBI who also have poor sleep satisfaction can pay attention but may have difficulty retaining information. The study controlled for the differences between sleep satisfaction and hours spent sleeping, noting that those suffering TBI symptoms may be more likely to register poor sleep despite sleeping for an adequate number of hours.

***Medical Protocol and Training***

## Military Medicine: Mobile Learning Module Improves Knowledge of Medical Shock for Forward Surgical Team Members

Mobile learning allows the Army to standardize education and treatment of medical shock. It also provides medics with an opportunity to refresh understanding of treatment and well as continue their education. This comparison of mobile and lecture-based education found that both groups performed the same despite the mobile learning lesson being 1/3 the length of lectures. Mobile education also enabled the user to pause and re-watch certain sections, resulting in a more efficient use of time. As budget cuts loom and medical demands continue to reach to the most remote areas of action, the use of mobile education can help reduce costs and enhance efficiency. Before mobile learning can be implemented, however, studies should be performed to assess curriculum.

## Military Medicine: Beyond Battlemind: Evaluation of a New Mental Health Training Program for Canadian Forces Personnel Participating in Third-Location Decompression

The analysis of Third-location Decompression training directed at easing the transition from a theater of war to civilian life documents a past method of administering TLDs and an updated version that is focused on mental health. The new program allows for a more flexible postdeployment transition and adjusts for critiques of the program. Program participants preferred the new program to the old version, in addition to a more favorable attitude towards mental health care following the program. The program did not, however, affect the stigma around receiving mental health care between the two programs. In developing postdeployment programs to reintegrate Soldiers, the content of the program should be evaluated to be sure to promote the work of mental health professionals and mental health care.

## Military Medicine: Multidisciplinary Response to the Escherichia coli O104 Outbreak in Europe

In light of a 2011 outbreak of E. coli in Europe, the study documents the outbreak response from U.S. Army Public Health Command Region-Europe. The description highlights the ad hoc advisory team, the ability to engage with German and European health professionals, a strong epicenter of communication, and tailored solutions related to food supply chains. Army messaging can utilize the documentation of an outbreak response such as this in the event of future outbreaks.

## Military Medicine: Medical Evacuation and Triage of Combat Casualties in Helmand Province, Afghanistan: October 2010–April 2011

In their study of helicopter medical evacuations in Helmand Province, researchers advise on three ways to enhance healthcare outcome. First, integrating pre-hospital care and the transfer of medical information will aid in restoring health during the "golden hour," when patients are most vulnerable. Second, mobile pre-hospital care units should be implemented as a midway point between the care that medics can provide during helicopter transport and comprehensive care at a dedicated medical facility. Finally, retrieving medics should be provided with protocol to select polytrauma victims to be immediately transferred to the highest care possible within the region.

## Military Medicine: Veteran Experiences Related to Participation in Shared Medical Appointments

This research sought to determine the effectiveness of shared medical appointments (SMA) in helping veterans stop smoking, select healthier eating options and increase physical activity. SMAs involve a team of healthcare providers meeting in a group session with a patient to discuss their health. Teams include social workers, physical therapists, physicians, psychologists etc. Veterans who participated in SMAs exhibited an increased knowledge of their complete health, which led to feelings of empowerment and control of their health. SMA programs that involved patients meeting with others who shared their medical concerns displayed improved health, lower blood pressure and weight loss among other benefits. Participants also cited convenience as a major benefit of SMAs.

## Military Medicine: Is the ASVAB ST Composite Score a Reliable Predictor of First-Attempt Graduation for the U.S. Army Operating Room Specialist Course?

Research indicates that a preliminary Armed Services Vocational Aptitude Battery (ASVAB) test score is a reliable predictor of service members' ability to pass the U.S. Army Operating Room Specialist Course. Specifically, those who scored 10 points or higher than their peers were five times more likely to pass the course. Additionally, women were 2.5 times more likely to pass the class than their male peers. Raising the ASVAB threshold will help determine which Soldiers are more likely to pass the course.

## Military Medicine: Stretch and Wrap Style Tourniquet Effectiveness With Minimal Training

Researchers set to determine if service members could apply a stretch, wrap and tuck (SWAT-T) tourniquet with only brief training. Fifteen military undergraduates watched a 19 second informational video three times, practiced twice and applied the tourniquet to a human volunteer on ten parts of the body. The students were able to use the tourniquet to stop arterial flow on limbs, but in some cases failed to wrap the tourniquet tight enough. Results indicate that in most instances, Soldiers can train easily to use SWAT-T, but more training is necessary for Soldiers in extreme environments where trauma can be expected.

***Sleep***

## Military Medicine: Maxillomandibular Advancement as Surgical Treatment for Obstructive Sleep Apnea in Active Duty Military Personnel: A Retrospective Cohort

Sleep apnea affects 3 percent of the military population, damaging service members' ability to sleep and work as well as increasing the likelihood that they will develop cardiovascular disease or be involved in traffic accidents. Research focused on determining the effectiveness of corrective jaw surgery in treating sleep apnea, and found that surgery is a viable treatment option with a significant success rate that will permanently treat sleep apnea and render service members deployable.

***Other***

## Military Medicine: Medical Costs of War in 2035: Long-Term Care Challenges for Veterans of Iraq and Afghanistan

This article delves into the medical costs to the United States of the wars in Iraq and Afghanistan noting that these conflicts differ in that more Soldiers are surviving war and IEDs cause expensive and complicated polytrauma injuries. By 2035, veterans of these conflicts will be middle aged with the health issues of previous veteran populations including back pain, joint issues, etc. but compounded by TBI, polytrauma and PTSD diagnoses. This article cites various efficiencies that can lower costs, and it suggests focusing on early detection for PTSD and TBI, early treatment and therapy for physiological injuries as well as more emphasis on preventative health including nutrition, exercise, weight control, substance abuse etc.

## Malaria Today: Is Dengue and Malaria Co-infection More Severe Than Single Infections?

Given that co-infection of dengue and malaria was first reported in 2005, the study seeks to describe the current rate and results of this phenomenon. The study discovered a greater risk of low blood platelet counts and anaemia in instances of co-infection of dengue and malaria. The recognition of the risk involved with co-infection of dengue and malaria in French Guiana, or other tropical environments, can be implemented into health policy conducted by the U.S. Army internationally.

## New England Journal of Medicine: Electronic Health Records and National Patient-Safety Goals

Turning to the variations of the implementation of EHR systems throughout hospitals, the researchers suggest that the variation includes risks to patient safety: device failures, miscommunication between EHR components, improper use, and coding discrepancies account for possible breakdowns of effective EHR usage. The study conducted by the VA preemptively suggests weaknesses in a fully integrated EHR system but offers adequate solutions for the potential problems.

## Archives of Neurology: Cholinergic Autonomic Dysfunction in Veterans With Gulf War Illness

The study objective was to confirm the previously held hypothesis that symptoms such as chronic diarrhea, dizziness, fatigue, and sexual dysfunction are as a result of cholinergic autonomic dysfunction in Gulf War veterans. The results indicated the above symptoms in addition to myriad others associated in the declared population as a consequence of cholinergic autonomic dysfunction.

## Military Medicine: Epidemiology of Contemporary Seroincident HIV Infection in the Navy and Marine Corps

Research considering the patterns, causes and effects of HIV on Navy and Marine populations determined that work factors are an important indication of HIV infection. A large percentage of infection occurred in the months after enlistment, indicating that medical personnel should focus on trainees and occupational schools. Deployment also increases the likelihood of contracting the virus, but enhanced early detection. The study noted that the Navy and Marines are unique populations whose research cannot always be accurately applied to other branches of the military. This is useful to MEDCOM as it helps determine the effectiveness of applying HIV research from other military populations to the Army.

## JAMA: Association of Warfarin Therapy Duration After Bioprosthetic Aortic Valve Replacement With Risk of Mortality, Thromboembolic Complications, and Bleeding

Warfarin therapy, traditionally used following aortic valve replacement (AVR) surgery, is prone to result in patient death if the treatment is discontinued within 3 months of surgery. The truncated treatment period can also result in thromboembolic complications and stroke. For populations requiring the AVR surgery, then, should expect to undergo the warfarin treatment for a minimum of 3 months and a suggested minimum of 6 months in order to mitigate the risk of cardiovascular death.

## Lancet: Global burden of cancer in 2008: a systematic analysis of disability-adjusted life-years in 12 world regions

The study identified country-specific designations of the share of cancer burden globally. Nearly 170 million healthy life-years were lost in 2008 to cancer, with China carrying 25 percent of the total. Divided by gender, Eastern Europe dominated men's cancer while Oceania and sub-Saharan African had the lion's share of women's cancer burden. Having a complete picture of the international incidence of cancer allows for research and policy to highlight possible environmental factors that can help to understand and treat cancers.

## Medical Journal Clips

***Nutrition and Fitness***

# Multivitamins in the Prevention of Cardiovascular Disease in Men

Journal of the American Medical Association
Howard D. Sesso, ScD, MPH; William G. Christen, ScD; Vadim Bubes, PhD; Joanne P. Smith, BA; Jean MacFadyen, BA; Miriam Schvartz, MD; JoAnn E. Manson, MD, DrPH; Robert J. Glynn, ScD; Julie E. Buring, ScD; J. Michael Gaziano, MD, MPH
7 Nov 2012

ABSTRACT

Context
Although multivitamins are used to prevent vitamin and mineral deficiency, there is a perception that multivitamins may prevent cardiovascular disease (CVD). Observational studies have shown inconsistent associations between regular multivitamin use and CVD, with no long-term clinical trials of multivitamin use.

Objective
To determine whether long-term multivitamin supplementation decreases the risk of major cardiovascular events among men.

Design, Setting, and Participants
The Physicians' Health Study II, a randomized, double-blind, placebo-controlled trial of a common daily multivitamin, began in 1997 with continued treatment and follow-up through June 1, 2011. A total of 14 641 male US physicians initially aged 50 years or older (mean, 64.3 [SD, 9.2] years), including 754 men with a history of CVD at randomization, were enrolled.

Intervention
Daily multivitamin or placebo.

Main Outcome Measures
Composite end point of major cardiovascular events, including nonfatal myocardial infarction (MI), nonfatal stroke, and CVD mortality. Secondary outcomes included MI and stroke individually.

Results
During a median follow-up of 11.2 (interquartile range, 10.7-13.3) years, there were 1732 confirmed major cardiovascular events. Compared with placebo, there was no significant effect of a daily multivitamin on major cardiovascular events (11.0 and 10.8 events

per 1000 person-years for multivitamin vs placebo, respectively; hazard ratio [HR], 1.01; 95% CI, 0.91-1.10; P = .91). Further, a daily multivitamin had no effect on total MI (3.9 and 4.2 events per 1000 person-years; HR, 0.93; 95% CI, 0.80-1.09; P = .39), total stroke (4.1 and 3.9 events per 1000 person-years; HR, 1.06; 95% CI, 0.91-1.23; P = .48), or CVD mortality (5.0 and 5.1 events per 1000 person-years; HR, 0.95; 95% CI, 0.83-1.09; P = .47). A daily multivitamin was also not significantly associated with total mortality (HR, 0.94; 95% CI, 0.88-1.02; P = .13). The effect of a daily multivitamin on major cardiovascular events did not differ between men with or without a baseline history of CVD (P = .62 for interaction).

Conclusion
Among this population of US male physicians, taking a daily multivitamin did not reduce major cardiovascular events, MI, stroke, and CVD mortality after more than a decade of treatment and follow-up.

Study Design

The PHS II was a randomized, double-blind, placebo-controlled, 2 × 2 × 2 × 2 factorial trial evaluating the balance of risks and benefits of a multivitamin (Centrum Silver or placebo daily [Pfizer; formerly Wyeth, American Home Products, and Lederle]), vitamin E (400-IU synthetic α-tocopherol or placebo on alternate days [BASF Corporation]), vitamin C (500-mg synthetic ascorbic acid or placebo daily [BASF Corporation]), and beta carotene (50-mg Lurotin or placebo on alternate days [BASF Corporation]) in the prevention of CVD, cancer, eye disease, and cognitive decline among 14 641 male physicians initially aged 50 years or older.32 The beta carotene component ended as scheduled in March 2003, and the vitamin E and C components ended as scheduled in 2007, with a lack of effect reported for CVD22 and cancer.33

As detailed previously,22 ,32 - 33 PHS II recruitment, enrollment, and randomization occurred in 2 phases (Figure 1). In phase 1, starting in July 1997, we invited 18 763 living participants from PHS I, a randomized trial of low-dose aspirin34 and beta carotene17 among 22 071 male physicians, to participate in PHS II. Men were ineligible if they reported a history of cirrhosis or active liver disease, were taking anticoagulants, or reported a serious illness that might preclude participation. Men also must have been willing to forgo current use of multivitamins or individual supplements containing more than 100% of the recommended dietary allowance of vitamin E, vitamin C, beta carotene, or vitamin A. Men with a history of myocardial infarction (MI), stroke, or cancer remained eligible. We randomized 7641 (41%) willing and eligible PHS I participants into PHS II.

Phase 2 of the PHS II began in July 1999 with invitational letters and baseline questionnaires sent to 254 597 additional US male physicians 50 years or older identified from a list from the American Medical Association that excluded PHS I participants. By July 2001, 42 165 men (16.6%) had completed the baseline PHS II questionnaire, of whom 11 128 (26.4%) were willing and eligible to participate based on the same eligibility criteria as PHS I participants. Of 11 128 physicians who entered a run-in phase, 7000 (63%) were adherent with their pills and were randomized into PHS II.

A total of 14 641 men were randomized into PHS II in blocks of 16 and stratified by age, prior diagnosis of cancer, prior diagnosis of CVD, and, for 7641 PHS I participants, their original beta carotene treatment assignment. There were 754 (5.1%) men with a history of MI or stroke before randomization.

All participants provided written informed consent, and the institutional review board at Brigham and Women's Hospital approved the research protocol.

Study Treatment, Follow-up, and Adherence

Every 6 months for the first year, then annually thereafter, PHS II participants were sent monthly calendar packs containing a multivitamin or placebo. Annual mailed questionnaires asked about adherence, adverse events, end points, and risk factors. Blinded treatment and follow-up continued through June 1, 2011, the scheduled end of the PHS II multivitamin component. Data analyses include follow-up and validation of reported end points through August 2012. Morbidity and mortality follow-up in PHS II were high—98.2% and 99.9%, respectively. In addition, morbidity and mortality follow-up as a percentage of person-time each exceeded 99.9%. Adherence with the multivitamin component was defined from participant self-report, which has been shown to be highly reliable in physicians,35 as taking at least two-thirds of the pills.

Confirmation of CVD End Points

For the multivitamin component, a primary end point was major cardiovascular events (including nonfatal MI, nonfatal stroke, and CVD mortality). Prespecified secondary end points included in this report include total MI and total stroke. Other end points considered in these analyses included fatal and nonfatal MI and stroke, cardiovascular death, ischemic and hemorrhagic stroke, and total mortality.

For each of the above self-reported end points, we requested permission from the participant to examine all relevant medical records. On receipt of consent, medical records were requested and reviewed by an end points committee of physicians blinded to randomized treatment assignment. We were unable to obtain adequate medical records for less than 5% of the end points.

The diagnosis of MI was confirmed by evidence of symptoms in the presence of either diagnostic elevations of cardiac enzyme levels or diagnostic changes on electrocardiograms. For fatal events, the diagnosis of MI was also accepted based on autopsy findings.34 We confirmed diagnoses of stroke defined as a typical neurologic deficit of sudden or rapid onset and vascular origin, lasting more than 24 hours. Stroke was classified according to National Survey of Stroke criteria into ischemic, hemorrhagic, and unknown subtype,36 with high interobserver agreement.37

Participant deaths were usually reported by family members or postal authorities. Following a report of a participant death, we obtained death certificates, autopsy reports, or both. Total mortality was confirmed by the end points committee or by death

certificate. Mortality attributable to CVD was additionally documented by convincing evidence of a cardiovascular mechanism from all available sources. For men with unknown vital status, we used web and National Death Index searches to identify deaths.

Only confirmed end points of MI, stroke, and CVD death were included in this analysis. We also collected data on participant self-reports of congestive heart failure, angina pectoris, and revascularization (including coronary artery bypass graft surgery and percutaneous coronary intervention) for inclusion in our analyses.

Statistical Analyses

All primary analyses were based on the intention-to-treat principle, in which all 14 641 randomized PHS II participants were classified according to their randomized multivitamin treatment assignment and underwent follow-up until the occurrence of major cardiovascular events, death, loss to follow-up, or the end of the multivitamin component of PHS II on June 1, 2011.

We performed all analyses using SAS version 9.2 (SAS Institute Inc) and S-Plus (Insightful Corp), with statistical significance set at $P < .05$ using 2-sided tests. The PHS II was estimated to have 80% power to detect a 12% reduction in the primary end point of major cardiovascular events.

We initially compared baseline characteristics by multivitamin treatment assignment to ensure that randomization equally distributed baseline characteristics by active vs placebo groups. As done in previous PHS II trial analyses,22 ,33 Cox proportional hazards models estimated hazard ratios (HRs) and 95% CIs, comparing event rates in the multivitamin and placebo groups. For each prespecified end point, we stratified on the presence of CVD at randomization and adjusted for PHS II study design variables, including age (in years), PHS cohort (original PHS I participant, new PHS II participant), and randomized vitamin E, vitamin C, and beta carotene assignments. For analyses of total major cardiovascular events, all new events were included, regardless of whether the participant had a baseline history of CVD. Analyses of individual cardiovascular end points did not censor men on occurrence of another cardiovascular end point. For analyses of total and cardiovascular mortality, we included all 14 641 PHS II participants; for total mortality, we additionally stratified by history of cancer at randomization.
We tested the proportional hazards assumptions by including an interaction term for multivitamin treatment with the logarithm of time; this assumption was not violated for major cardiovascular events, total MI, and total stroke ($P > .05$ for each). Cumulative incidence curves compared the overall effect of the multivitamin component on major cardiovascular events, total MI, and total stroke using a crude log-rank test. We investigated the effect of adherence to the multivitamin intervention on our primary results using sensitivity analyses with censoring and stratification.

We then conducted additional exploratory analyses on the effect of the multivitamin intervention on major cardiovascular events, total MI, and total stroke after excluding the first 2 or 5 years of follow-up to explore a possible early or late benefit associated with long-term multivitamin use. We also conducted subgroup analyses stratified by major risk factors, parental history of MI at ages younger than 60 years, and selected coronary biomarkers and dietary factors available in a subgroup of PHS II participants. We evaluated the

effect of a daily multivitamin within the prespecified subgroups of 754 men with and 13 887 men without a baseline history of CVD. Effect modification was assessed using interaction terms between subgroup indicators and randomized multivitamin treatment assignment.

RESULTS.

We randomized a total of 14 641 men into PHS II; the mean age of participants was 64.3 (SD, 9.2) years. Factors measured at baseline were similar between the multivitamin and placebo groups (Table 1). Among coronary risk factors, there was a low proportion (3.6%) of current smokers and a relatively high proportion (59.9%) of men who exercised 1 time/wk or more, which was countered with 42.0% of men reporting a history of hypertension, 35.4% a history of high cholesterol levels, and 6.2% a history of diabetes. Baseline aspirin use was high (77.4%) in this population of physicians, in part reflective of their previous participation and results of the PHS I trial assessing aspirin use and CVD.34 There were 754 men (5.1%) with a baseline history of CVD and 1312 (9.0%) with a baseline history of cancer.

Median follow-up of PHS II participants was 11.2 years (interquartile range, 10.7-13.3 years; maximum, 13.8 years), totaling 164 320 person-years. Adherence was 76.8% in the multivitamin group and 77.1% in the placebo group at 4 years (P = .71); 72.3% in the multivitamin group and 70.7% in the placebo group at 8 years (P = .15); and 67.5% in the multivitamin group and 67.1% in the placebo group at the end of follow-up (P = .70). There were small differences between the multivitamin and placebo groups when comparing the avoidance of individual nontrial multivitamin use (<30 days/y) at 4 years of follow-up (86.7% and 85.4%, respectively; P = .03) and 8 years of follow-up (78.5% and 75.8%, P = .01) but not by the end of multivitamin follow-up (81.0% and 80.3%; P = .35). During multivitamin treatment, we confirmed that 1732 men had major cardiovascular events, including 652 cases (first events) of MI and 643 cases of stroke (527 ischemic stroke, 94 hemorrhagic stroke), and 829 had cardiovascular death, with some men experiencing multiple events. A total of 2757 men (18.8%) died during follow-up.

Multivitamin Use and Major Cardiovascular Events

The rates of major cardiovascular events were 11.0 per 1000 person-years in the multivitamin group and 10.8 per 1000 person-years in the placebo group.

In secondary analyses, there were fewer MI deaths among multivitamin users (HR, 0.61; 95% CI, 0.38-0.995; P = .048). Among stroke subtypes, a daily multivitamin had no effect on either ischemic stroke (HR, 1.10; 95% CI, 0.92-1.30; P = .29) or hemorrhagic stroke (HR, 1.08; 95% CI, 0.72-1.63; P = .69). We found no significant effect of a daily multivitamin on rates of congestive heart failure (HR, 0.95; 95% CI, 0.83-1.09; P = .47), angina (HR, 1.00; 95% CI, 0.91-1.09; P = .96), and coronary revascularization (HR, 1.03; 95% CI, 0.94-1.13; P = .50). Taking a daily multivitamin was not significantly associated with CVD mortality (5.0 and 5.1 events per 1000 person-years for multivitamin and placebo, respectively; HR, 0.95; 95% CI, 0.83-1.09; P = .47). There were fewer total deaths among multivitamin users (HR, 0.94; 95% CI, 0.88-1.02; P = .13), but this was not statistically significant.

In secondary analyses, exclusion of the first 2 or 5 years of follow-up did not alter the results for major cardiovascular events, total MI, or total stroke. Analyses adjusting for adherence either during follow-up or averaged over the whole trial, or adjusting for drop-ins, did not materially change the effect of multivitamin use on risk of major cardiovascular events.

Modifiers of the Effect Between Multivitamin Use and Major Cardiovascular Events

In subgroup analyses, we examined whether baseline clinical, lifestyle, familial, biochemical, and dietary risk factors for CVD, along with the other randomized PHS II interventions, modified the effect of a daily multivitamin on major cardiovascular events (eTable 1). There was a suggestion of a differential effect across age groups (P = .041 for interaction), with possible differences among men aged 50 to 59 years (HR, 1.27; 95% CI, 0.99-1.63; P = .06) and men 70 years or older (HR, 0.91; 95% CI, 0.81-1.03; P = .14). We found no other evidence of effect modification by baseline risk factors on major cardiovascular events (P > .05 for interaction for all). There also were no multiplicative or subadditive interactions of the multivitamin component with randomized vitamin C, vitamin E, or beta carotene treatment in PHS II (P > .05 for interaction for all).

We found no significant interaction by baseline CVD history status (P = .62 for interaction) for primary (HR, 1.02; 95% CI, 0.92-1.13) vs secondary (HR, 0.96; 95% CI, 0.75-1.22) prevention (Table 3). The cumulative incidence curves did not differ for primary (crude log-rank P = .71) or secondary (crude log-rank P = .94) prevention during up to 14 years of treatment and follow-up (Figure 3). The apparent lower rate of MI death among multivitamin users persisted (HR, 0.56; 95% CI, 0.33-0.95; P = .03), whereas power was limited, with only 9 cases of MI death among those with baseline CVD (P = .31 for interaction). The effect of a daily multivitamin on total MI, total stroke, and other cardiovascular end points did not differ between men with and without baseline CVD (P > .05 for interaction for all).There was a similar lack of significant benefit for the secondary end points of total MI (3.9 and 4.2 events per 1000 person-years for multivitamin and placebo, respectively; HR, 0.93; 95% CI, 0.80-1.09; P = .39) and total stroke (4.1 and 3.9 events per 1000 person-years; HR, 1.06; 95% CI, 0.91-1.23; P = .48) compared with men taking placebo. This lack of effect is illustrated in the corresponding cumulative incidence curves (crude log-rank P > .05 for both) (Figure 2).

Potential Adverse Effects of Daily Multivitamin Use

Besides the primary and secondary end points, we assessed several potential adverse effects of daily multivitamin use and found no significant effects on gastrointestinal tract symptoms (peptic ulcer, constipation, diarrhea, gastritis, and nausea), fatigue, drowsiness, skin discoloration, and migraine (P > .05 for all). Participants taking the multivitamin vs placebo were more likely to have skin rashes (2125 in the multivitamin group and 2002 in the placebo group; HR, 1.07; 95% CI, 1.01-1.14; P = .03). In addition, findings were inconsistent for effects of daily multivitamin use on minor bleeding, suggesting the role of chance. There was a reduction in hematuria (1194 men in the multivitamin group and 1292 in the placebo group; HR, 0.91; 95% CI, 0.84-0.98; P = .02), an increase in epistaxis (1579 in the multivitamin group and 1451 in the placebo group; HR, 1.10; 95% CI, 1.02-1.18; P = .01), and no effect on easy bruising or other bleeding (2786 in the multivitamin group and 2806 in the placebo group; HR, 0.99; 95% CI, 0.94-1.05; P = .77).

COMMENT.

The PHS II represents to our knowledge the only large-scale, randomized, double-blind, placebo-controlled trial testing the long-term effects of a commonly available multivitamin on the prevention of chronic disease. We found that after more than a decade of daily multivitamin use among middle-aged and older men, daily multivitamin use did not reduce the primary end point of major cardiovascular events. Multivitamin use also did not reduce the risk of total MI; total, ischemic, or hemorrhagic stroke; cardiovascular death; or other cardiovascular end points, including congestive heart failure, angina, or coronary revascularization. The reduction observed in fatal MI (P = .048) may have been attributable to chance. These findings on CVD and the decision to take a multivitamin should be considered in the context of initial nutritional status and other outcomes to be considered in this trial.

The lack of an effect of a daily multivitamin on CVD appears consistent with what is known to date. Basic research indicates several mechanisms by which specific micronutrients contained in multivitamins may prevent CVD38 - 39 through modifications in platelet activity,40 reductions in thrombotic potential,41 and modifications in vascular reactivity.42 The consistent observation that people consuming greater amounts of fruits and vegetables tend to have lower rates of coronary heart disease43 and stroke44 supports the idea that combinations of vitamins at moderate doses may offer protection against CVD.

Observational data examining multivitamin use and CVD are sparse and inconsistent. Among 1 063 023 US adults from the Cancer Prevention Study II, men without CVD taking a multivitamin had an age-adjusted relative risk of death from ischemic heart disease of 0.91 (P ≤ .001), attenuated on multivariate adjustment10 ; similar results were noted for women. In 80 082 Nurses' Health Study participants, multivitamin use was associated with a significant reduction in coronary heart disease incidence (relative risk, 0.76; 95% CI, 0.65-0.90) after 14 years,9 a result further confirmed with additional follow-up.12 In a Swedish population-based case-control study in adults aged 45 to 70 years, the multivariate odds ratio of MI comparing regular users vs nonusers of multivitamins was 0.79 (95% CI, 0.63-0.98) among 2053 men and 0.66 (95% CI, 0.48-0.91) among 928 women.13

In contrast, in the PHS I enrollment cohort of 83 936 initially healthy male physicians, there was no association between baseline multivitamin use and either CVD or coronary heart disease mortality.11 Among 161 808 Women's Health Initiative participants, of whom 41.5% took a multivitamin, there was no association between multivitamin use and the risk of CVD, MI, or stroke after a median of 8 years of follow-up.14 Last, there was also no association between multivitamin use and cardiovascular mortality in 182 099 men and women from the Multiethnic Cohort Study after a mean follow-up of 11 years.16

Several large trials of single agents or combinations of vitamins and minerals, generally at doses well above recommended dietary allowances and the multivitamin dose used in PHS II, have demonstrated no effect on CVD.17 - 18 ,20 ,45 Primary prevention trials that have examined smaller combinations of vitamins and minerals, including the Linxian Chinese Cancer Prevention Trial26 and the Supplementation en Vitamines et Minereaux Antioxydants (SU.VI.MAX) trial28 as well as secondary prevention trials such as the Heart Protection Study,27 found no effect on CVD. Other randomized trials have tested combinations of B vitamins with folic acid at

high doses, particularly in the secondary prevention of CVD, but have found no protective effect.46 Moreover, the Women's Health Initiative calcium and vitamin D trial, testing vitamin D3 (400 IU/d) plus calcium (1000 mg/d), found no effect on CVD.47

Baseline nutritional status among our physician participants remains a critical consideration in the interpretation of our findings. PHS II participants likely represent, on average, a well-nourished population who already have adequate or optimum intake levels of nutrients, for which supplementation may offer no additional benefit.48 However, the requirement for PHS II participants to avoid personal use of multivitamin supplements also lowered their in-trial intake of essential vitamins and minerals. Additional studies are needed to understand how the range of baseline nutritional status among PHS II participants and other populations may modify the effect of a daily multivitamin on cardiovascular end points. Further, several behavioral (eg, exercise, weight loss) and pharmacological (eg, lipid-lowering therapies) interventions are available to effectively lower CVD risk. This may make it difficult for vitamin supplements such as a multivitamin to meaningfully contribute toward risk reduction.

Several unique strengths of this trial include more than a decade of treatment and follow-up, high statistical power for our primary end point of major cardiovascular events, consistently good adherence in taking a daily multivitamin, and the inclusion of physician participants providing high-quality reporting of health information. We are unaware of any other long-term clinical trials that have tested use of a multivitamin in the prevention of CVD and other chronic diseases, highlighting the importance of trials like PHS II to test the efficacy of supplements and assess potential causality across a range of clinically relevant outcomes. In addition, we selected a commonly used multivitamin formulation when we initiated PHS II to increase the generalizability of our findings.

This trial also has important potential limitations to be considered. We relied on a specific, constant multivitamin formulation ( eTable 2), which is one of many multivitamin formulations. There was an observed reduction in total cancer found for the PHS II multivitamin,31 suggesting that the formulation used may be adequate for cancer but not for CVD. This highlights the need to understand how essential vitamins and minerals may differentially interact and influence cardiovascular and cancer mechanisms, even at usual levels of vitamin and mineral intake. Although PHS II included more than a decade of treatment, an even longer duration of multivitamin use may be required to derive any cardiovascular benefits. Existing epidemiologic data can provide insight on this concept, while PHS II remains the only trial of its kind for which extended follow-up of CVD end points can provide important longer-term mechanistic perspectives.

The PHS II also may have limited generalizability, because our study population was confined to middle-aged and older, predominantly white, male physicians. Despite some multivitamin nonadherence and drop-ins during PHS II, adjusted analyses adjusted for adherence and drop-ins reiterated a lack of effect of multivitamin use on major cardiovascular events. As with any trial, chance may be important when multiple hypotheses are tested; thus, cautious interpretation of secondary analyses is warranted. In addition, long-term multivitamin use may be more effective when initiated earlier in life to counter the initiation and progression of atherosclerosis that often begins at an earlier age.

After a mean of 11.2 years of treatment and follow-up in 14 641 men, daily multivitamin supplementation in this trial did not reduce the risk of major cardiovascular events. These data do not support multivitamin use to prevent CVD, demonstrating the importance of long-term clinical trials of commonly used nutritional supplements. Whether to take a daily multivitamin requires consideration of an individual's nutritional status, because the aim of supplementation is to prevent vitamin and mineral deficiency, plus consideration of other potential effects, including a modest reduction in cancer31 and other important outcomes in PHS II that will be reported separately.

AUTHOR INFORMATION.

Corresponding Author: Howard D. Sesso, ScD, MPH, Brigham and Women's Hospital, 900 Commonwealth Ave E, Third Floor, Boston, MA 02215 (hsesso@hsph.harvard.edu).

Author Contributions: Drs Sesso and Gaziano had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Study concept and design: Sesso, Christen, Bubes, Manson, Glynn, Buring, Gaziano.

Acquisition of data: Sesso, Bubes, Smith, MacFadyen, Schvartz, Manson, Buring, Gaziano.

Analysis and interpretation of data: Sesso, Christen, Bubes, Manson, Glynn, Buring, Gaziano.

Drafting of the manuscript: Sesso, Gaziano.

Critical revision of the manuscript for important intellectual content: Christen, Bubes, Smith, MacFadyen, Schvartz, Manson, Glynn, Buring, Gaziano.

Statistical analysis: Sesso, Bubes, Glynn, Gaziano.

Obtained funding: Sesso, Buring, Gaziano.

Administrative, technical, or material support: Sesso, Bubes, Smith, MacFadyen, Schvartz, Manson, Gaziano.

Study supervision: Sesso, Bubes, MacFadyen, Schvartz, Gaziano.

Conflict of Interest Disclosures: All authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Dr Sesso reported receiving investigator-initiated research funding from the National Institutes of Health (NIH), the Tomato

Products Wellness Council, and Cambridge Theranostics Ltd. Dr Christen reported receiving research funding support from the NIH, Harvard University (Clinical Nutrition Research Center), and DSM Nutritional Products Inc (formerly Roche Vitamins). Dr Manson reported receiving investigator-initiated research funding from the NIH; assistance with study pills and packaging from BASF and Cognis Corporations for the Women's Antioxidant and Folic Acid Cardiovascular Study and from Pronova BioPharma and Pharmavite for the VIT amin D and OmegA -3 TriaL ; and funding from the nonprofit Aurora Foundation. Dr Glynn reported receiving investigator-initiated research funding from the NIH, Bristol-Meyers Squibb, AstraZeneca, and Novartis, and signing a consulting agreement with Merck to give an invited talk. Dr Buring reported receiving investigator-initiated research funding from the NIH and assistance with study pills and packaging from Natural Source Vitamin E Association and Bayer Healthcare for the Women's Health Study. Dr Gaziano reported receiving investigator-initiated research funding from the NIH, the Veterans Administration, and BASF Corporation to assist in the establishment of this trial cohort; assistance with study agents and packaging from BASF Corporation and Pfizer (formerly Wyeth, American Home Products, and Lederle); and assistance with study packaging provided by DSM Nutritional Products Inc (formerly Roche Vitamins). No other authors reported disclosures.

Funding/Support: This study was supported by grants CA 097193, CA 34944, CA 40360, HL 26490, and HL 34595 from the NIH (Bethesda, Maryland) and an investigator-initiated grant from BASF Corporation (Florham Park, New Jersey). Study agents and packaging were provided by BASF Corporation and Pfizer (formerly Wyeth, American Home Products, and Lederle) (New York, New York), and study packaging was provided by DSM Nutritional Products Inc (formerly Roche Vitamins) (Parsippany, New Jersey).

Role of the Sponsor: NIH, BASF, Pfizer, and DSM Nutritional Products Inc had no role in the design and conduct of the study; the collection, management, analysis, and interpretation of the data; or the preparation, review, or approval of the manuscript.

Data and Safety Monitoring Board: Voting members over the course of the Physicians' Health Study (PHS) II trial included Lawrence Cohen, Rory Collins, Theodore Colton, I. Craig Henderson, Andrea LaCroix, Ross Prentice, and Nanette Wenger (chair); ex-officio members included Mary Francis Cotch, Jeffrey Cutler, Frederick Ferris, Jerome Fleg, Peter Greenwald, Natalie Kurinij, Howard Parnes, Marjorie Perloff, Eleanor Schron, and Alan Zonderman.

Disclaimer: Dr Gaziano, a contributing editor for JAMA, was not involved in the editorial review of or decision to publish this article.

Additional Contributions:We are deeply indebted to the 14 641 physician participants for their long-standing dedication and conscientious collaboration. We also acknowledge the long-term contributions of Charles Hennekens, MD, DrPH, of Florida Atlantic University to the PHS and the exemplary contributions of the staff of the PHS at Brigham and Women's Hospital, under the leadership of Joanne Smith: Charlene Belanger, Eileen Bowes, Kenneth Breen, Mary Breen, Mary G. Breen, Jose Carrion, Shamikhah Curry, Colleen Evans, Ivan Fitchorov, Natalya Gomelskaya, Cindy Guo, Delia Guo, Jasmah Hanna, Beth Holman, Andrea Hrbek, Gregory Kotler, Tony Laurinaitis, Hannah Mandel, Chandra McCarthy, Geneva McNair, Annie Murray, Leslie Power, Philomena Quinn, Harriet Samuelson, Fred Schwerin, Andromache Sheehey, Sara Tower, Martin Van Denburgh, Diana Walrond, Phyllis Johnson Wojciechowski, and Angela Zhang. Last, we are deeply grateful for the efforts of the PHS Endpoints Committee,

including Samuel Goldhaber, Carlos Kase, Meir Stampfer, and James Taylor, over the course of PHS II. Each named individual was compensated for his or her contribution as part of the grant support.

Online-Only Content: The Author Video Interview is available here.

BACK TO TOP

## Leisure Time Physical Activity of Moderate to Vigorous Intensity and Mortality: A Large Pooled Cohort Analysis

PLOS Medicine

Steven C. Moore1, Alpa V. Patel2, Charles E. Matthews, Amy Berrington de Gonzalez, Yikyung Park, Hormuzd A. Katki, Martha S. Linet1, Elisabete Weiderpass, Kala Visvanathan, Kathy J. Helzlsouer, Michael Thun, Susan M. Gapstur, Patricia Hartge, I-Min Lee

6 Nov 2012

Abstract

Background

Leisure time physical activity reduces the risk of premature mortality, but the years of life expectancy gained at different levels remains unclear. Our objective was to determine the years of life gained after age 40 associated with various levels of physical activity, both overall and according to body mass index (BMI) groups, in a large pooled analysis.

Methods and Findings

We examined the association of leisure time physical activity with mortality during follow-up in pooled data from six prospective cohort studies in the National Cancer Institute Cohort Consortium, comprising 654,827 individuals, 21–90 y of age. Physical activity was categorized by metabolic equivalent hours per week (MET-h/wk). Life expectancies and years of life gained/lost were calculated using direct adjusted survival curves (for participants 40+ years of age), with 95% confidence intervals (CIs) derived by bootstrap. The study includes a median 10 y of follow-up and 82,465 deaths. A physical activity level of 0.1–3.74 MET-h/wk, equivalent to brisk walking for up to 75 min/wk, was associated with a gain of 1.8 (95% CI: 1.6–2.0) y in life expectancy relative to no leisure time activity (0 MET-h/wk). Higher levels of physical activity were associated with greater gains in life expectancy, with a gain of 4.5 (95% CI: 4.3–4.7) y at the highest level (22.5+ MET-h/wk, equivalent to brisk walking for 450+ min/wk). Substantial gains were also observed in each BMI group. In joint analyses, being active (7.5+ MET-h/wk) and normal weight (BMI 18.5–24.9) was associated with a gain of

7.2 (95% CI: 6.5–7.9) y of life compared to being inactive (0 MET-h/wk) and obese (BMI 35.0+). A limitation was that physical activity and BMI were ascertained by self report.

Conclusions

More leisure time physical activity was associated with longer life expectancy across a range of activity levels and BMI groups.

Editor's Summary

Background

Regular physical activity is essential for human health. It helps to maintain a healthy body weight and prevents or delays heart disease, type 2 diabetes, and some cancers. It also makes people feel better and increases life expectancy. The World Health Organization (WHO) currently recommends that adults do at least 150 minutes of moderate- to vigorous-intensity physical activity every week. Moderate-intensity physical activities (for example, brisk walking and gardening) require a moderate amount of effort and noticeably increase the heart rate; vigorous-intensity physical activities (for example, running or fast swimming) require a large amount of effort and cause rapid breathing and a substantial heart rate increase. Worryingly, people in both developed and developing countries are becoming increasingly physically inactive. People are sitting at desks all day instead of doing manual labor; they are driving to work in cars instead of walking or cycling; and they are participating in fewer leisure time physical activities.

Why Was This Study Done?

Although various studies suggest that physical activity increases life expectancy, few have quantified the years of life gained at distinct levels of physical activity. Moreover, the difference in life expectancy between active, overweight individuals and inactive, normal weight individuals has not been quantified. Thus, it is hard to develop a simple public health message to maximize the population benefits of physical activity. In this pooled prospective cohort analysis, the researchers determine the association between levels of leisure time physical activities, such as recreational walking, and years of life gained after age 40, both overall and within body mass index (BMI) groups. A pooled prospective cohort analysis analyzes the combined data from multiple studies that have followed groups of people to investigate associations between baseline characteristics and outcomes such as death. BMI is a ratio of weight to height, calculated by dividing a person's weight by their height squared; normal weight is defined as a BMI of 18.5–24.9 kg/m2, obesity (excessive body fat) is defined as a BMI of more than 30 kg/m2.

What Did the Researchers Do and Find?

The researchers pooled self-reported data on leisure time physical activities and BMIs from nearly 650,000 individuals over the age of 40 years enrolled in one Swedish and five US prospective cohort studies, most of which were investigating associations between lifestyle factors and disease risk. They used these and other data to calculate the gain in life expectancy associated with specific levels of physical activity. A physical activity level equivalent to brisk walking for up to 75 minutes per week was associated with a gain of 1.8 years in life expectancy relative to no leisure time activity. Being active—having a physical activity level at or above the WHO-recommended minimum of 150 minutes of brisk walking per week—was associated with an overall gain of life expectancy of 3.4–4.5 years. Gains in life expectancy were seen also for black individuals and former smokers, groups for whom relatively few data had been previously available. The physical activity and life expectancy association was also evident at all BMI levels. Being active and normal weight was associated with a gain of 7.2 years of life compared to being inactive and class II+ obese (having a BMI of more than 35.0 kg/m2). However, being inactive but normal weight was associated with 3.1 fewer years of life compared to being active but class I obese (having a BMI of 30–34.9 kg/m2).

What Do These Findings Mean?

These findings suggest that participation in leisure time physical activity, even below the recommended level, is associated with a reduced risk of mortality compared to participation in no leisure time physical activity. This result may help convince currently inactive people that a modest physical activity program may have health benefits, even if it does not result in weight loss. The findings also suggest that physical activity at recommended levels or higher may increase longevity further, and that a lack of leisure time physical activity may markedly reduce life expectancy when combined with obesity. Although the accuracy and generalizability of these findings may be limited by certain aspects of the study's design (for example, some study participants may have overestimated their leisure time physical activity), these findings reinforce the public health message that both a physically active lifestyle and a normal body weight are important for increasing longevity.

BACK TO TOP

***Physiological Injuries and Surgery***

## Incidence and Risk Factors for Acute Low Back Pain in Active Duty Infantry

Military Medicine
CPT Justin Ernat, MC USA; CPT Jeffrey Knox, MC USA; LTC Joseph Orchowski, MC USA; LTC Brett Owens, MC USA

ABSTRACT

Although much research has been performed on occupational risk factors for low back pain, little has been published on low back pain among infantrymen. This purpose of this study is to evaluate the incidence of acute low back pain amongst active duty infantrymen as compared to a matched control population. The Defense Medical Epidemiology Database was searched and incidence rates were calculated and compared between infantry and noninfantry soldiers. Data was stratified and controlled for age, race, marital status, rank, and branch of service using the Poisson multivariate regression analysis. Significantly lower rates of acute low back pain were discovered in active duty infantrymen when compared to matched controls (32.9 versus 49.5 cases per 1,000 person-years). Additionally, significantly lower rates were identified in the Marines versus the Army, and among junior enlisted compared to senior enlisted service members.

## INTRODUCTION

Low back pain is a significant cause of disability in the working population and carries a considerable economic impact with an estimated annual productivity loss of $28 billion with up to 149 million lost work days in the United States. Because of the significant economic burden of this disease, numerous studies have focused on occupational factors involved in its development. Cited contributing occupations include, but are not limited to, truck, taxi, and bus drivers; operators of heavy machinery; construction workers; pilots; as well as numerous others. Aside from the actual physical demands, other factors to include job satisfaction, monotonous tasks, work relations, demands, stress, and perceived ability to work all influence a worker's risk for low back pain.

The active duty military population infantryman represents a high-demand occupational group that is regularly involved in physically challenging activities that regularly place high loads to their spinal column. Although such activities would be expected to place these individuals at increased risk of low back pain events, the incidence of low back pain in this population has not been well described. The purpose of our study is to assess the incidence and demographic risk factors of acute low back pain resulting in a visit to a health care provider in active duty infantrymen when compared to a matched control population.

## METHODS AND MATERIALS

To evaluate our study population of active duty infantrymen, a retrospective database analysis was performed using the Defense Medical Epidemiology Database (DMED). The DMED compiles diagnoses made by physicians and physician extenders for every patient encounter in a military treatment facility according to International Classification of Diseases, Ninth Revision (ICD-9) coding information. In addition, the DMED compiles information including the total number of service members on active duty each year, as well as their demographic and military specific information. It can track members as they move throughout the world and can also track service members' visits to nonmilitary outpatient health care facilities that provide care for active duty soldiers. It has been used in past studies to provide information on various musculoskeletal conditions, including low back pain.

To determine the total number of infantry patients with low back pain, we queried the ambulatory DMED system for the years 1998–2006 using the ICD-9 CM code 724.20. This code was used for ambulatory patients' initial visits for “lumbago” (low back pain) so that there was no repeat coding. This code can be used for the study with confidence because it is an easy diagnosis for health care providers to make without concern for variability in diagnostic criteria or method of diagnosis. Patients' results were then stratified by race, rank, age, branch of service, and marital status. Race was categorized as white, black, and other according to self-report. Junior (E1–E4) and senior (E5–E9) enlisted infantry members were included in the analysis. Officers were excluded from the analysis because of the variability in daily activities of senior infantry officers compared to enlisted infantrymen. Age was categorized as <20, 20 to 24, 25 to 29, 30 to 34, 35 to 39, and 40+ years. Marital status was classified as single, married, or “other” as determined by a self-report. This data was then compared to a control population consisting of those with all occupational codes other than infantry and matched for the above demographic factors.

Incidence rates (IRs) and IR ratios (IRRs) were calculated and compared using the Poisson multivariate regression analysis. All IRRs were calculated using the control group as the referent category. This study received institutional review board approval from Tripler Army Medical Center. No external funding was received for this study.

## RESULTS

A total 791,526 person-years of data were analyzed with 26,044 cases of low back pain yielding an overall incidence of 32.9 per 1,000 person-years among all infantrymen. A total of 4,125,727 person-years of data were analyzed with 204,404 cases of low back pain with an overall incidence of 49.5 per 1,000 person-years in the control group. After adjusting for age, race, rank, branch of service, and marital status infantrymen were found to have a significantly lower rate of low back pain resulting in a visit to a health care provider with an adjusted IRR of 0.69 (95% CI: 0.68–0.70) ($p < 0.0001$). The control group was used as the referent category .

Older age was found to be a significant risk factor in this population with the highest unadjusted IRs in the 40+ year-old age group with an incidence of 52.8 per 1,000 person-years. The 20- to 24-year-old age group was found to have the lowest rate of low back pain with an unadjusted IR of 27.2. The <20-year-old and 25- to 29-year-old age groups had intermediate rates of 34.8 and 34.8 per 1,000 person-years, respectively. IRRs using the control group as the referent category of the ages groups were as follows: 0.63 for <20, 0.58 for 20 to 24, 0.72 for 25 to 29, 0.83 for 30 to 34, 0.88 for 35 to 39, and 0.91 for 40+ year-olds.

The relative risk of low back pain between infantry and noninfantry soldiers was found to be significantly affected by age. In all age groups, the control group was used as the referent category and IRR was calculated in this manner. The overall adjusted IRR was 0.69, this ranged from 0.61 in the <20 age group to 0.91 in the 40+ age.

Race was found to be a significant factor in the incidence of low back pain in the infantrymen population with the lowest incidence among the “other” category (30.3 per 1,000 person-years). White infantrymen had the next highest incidence with an unadjusted IR of 32.2 per 1,000 person-years and blacks had the highest rate of 42.2 per 1,000 person-years. IRRs were also calculated for each of the individual race categories using the control group as the referent category. The unadjusted IRR are 0.67 for whites, 0.77 for blacks, and 0.65 for those classified as “other.”

Married infantrymen had a higher unadjusted IR of low back pain when compared to single infantrymen, 37.8 versus 29.3 per 1,000 person-years, respectively. The “other” category had the highest rate with 42.8 per 1,000 person-years. The unadjusted IRR were 0.61 for singles, 0.76 for married individuals, and 0.67 for those classified as “other.” These were calculated using matched controls as a referent category.

Rank was also a significant factor with senior enlisted infantrymen demonstrating higher rates of low back pain compared to junior enlisted. The unadjusted IR for low back pain among junior enlisted infantrymen was 31.8 per 1,000 person-years. The unadjusted rate for senior infantry enlisted infantrymen was 35.2 per 1,000 person-years. Unadjusted IRR for junior and senior enlistees were 0.59 and 0.80, respectively, using matched controls as the referent category.

Low back pain rates were significantly different between the branches of service with Army infantrymen demonstrating a much higher rate when compared to Marine infantrymen (41.4 per 1,000 person-years versus 15.7 per 1,000 person-years). When compared to matched controls, unadjusted IRR for the Army and Marines were 0.70 and 0.59, respectively.

## DISCUSSION

Our study analyzed 4,917,253 person-years of data with 230,448 episodes of acute low back pain resulting in a health care encounter. Our study demonstrated a significantly lower rate in active duty infantrymen compared to a matched control population. These results are lower than many prior series on both active duty military and civilian populations which report IRs of 40.5 per 1,000 person-years in active duty military and between 24.2 and 44.7 in civilian populations. Additionally, this difference was found to be significantly age-dependent with incidence approaching that of the control group as a patient increases in age.

Our study of infantry service members identified multiple risk factors contributing to an increased incidence in low back pain to include increasing age, senior enlisted rank, members of the Army, and marital status other than single. These factors are well described in a previous study by the same authors.

The infantry is a particularly interesting group to study with regards to musculoskeletal injuries. The loads carried by soldiers have been documented as far back as the Byzantines where weaponry and equipment could weigh as much as 80 kg. Modern infantrymen, in turn, are exposed to very high load carriage along with long and strenuous road marches in addition to multiple other very physically demanding tasks that have been associated with multiple injuries. Knapik et al demonstrated that 24% of light infantry soldiers carrying 46 kg over a 20-km course developed at least one injury, 23% of such were back injuries. Other injury-specific research associated with road marching and load carrying includes correlations with gait abnormalities, cardiovascular functioning, shoulder issues, foot blistering, effects on posture, and other overuse injuries.

Despite these risk factors, we identified a significantly lower rate of low back pain resulting in a health care encounter in the infantry population. The reason for this is unclear but likely multifactorial. One likely possibility is a decreased willingness to seek treatment for this population. The expectations of meeting certain professional and physical demands as well as the idea of succumbing to injury are psychosocial factors that this subgroup may face preventing them from going to see a provider. Carragee et al demonstrated that many soldiers feel that back pain is a normal part of their occupation and therefore tend to ignore or minimize this condition. Another potential factor is the training and fitness level required of the infantry soldier, which may provide a protective effect against the significant loads and forces placed on their lower backs. This remains unclear and represents avenues of further potential research in this population.

A previous study by the authors had demonstrated similar results with regards to branch of service in that Marines had lower rates of acute low back pain when compared to the Army, Navy, or Air Force. The overall IRs for the active duty Army and Marines in that study were 57.0 and 25.3, in comparison to the active duty infantry Army and Marines in the current study which were 41.4 and 15.7, respectively. The lower IRs in the current study could be as a result of some of the physical and mental factors described above that are associated with infantry service in comparison to other active duty soldiers. The reason behind this difference between services

is likely multifactorial but may be related to differences in training regimens which may provide an additional protective effect. Alternatively, differences in care-seeking behavior may likely contribute to some degree.

BACK TO TOP

## Long-term results of cementless hip arthroplasty with ceramic-on-ceramic articulation.

International Orthopaedics
Synder M, Drobniewski M, Sibiński M.
November 2012

Abstract

PURPOSE

The goal of the study was to evaluate long-term results of hip arthroplasty in patients with ceramic-on-ceramic articulation.

METHODS

The follow-up involved 220 primary total hip arthroplasty procedures (188 patients, 101 women and 87 men) after implantation of the Mittelmeier cementless hip endoprosthesis. The mean age of patients at surgery was 44.5 years and the mean follow-up was 19.6 years, with a minimum of 12.3 years. Dysplastic, idiopathic and post-traumatic coxarthrosis were the most frequent forms of degenerative hip changes. The Merle d'Aubigné and Postel classification, as modified by Charnley, was used for clinical evaluation.

RESULTS

Very good results were obtained in 39.5 % of the patients, good results in 43.6 %, satisfactory results in 9.1 % and poor results in 7.8 %. Twelve-year survival for the whole prosthesis was 86.36 %, for the acetabulum 89.99 % and for the stem 91.36 %.

CONCLUSIONS

Long-term results of hip arthroplasty using the Mittelmeier prosthesis are fairly encouraging with their low incidence of loosened prosthesis components after surgery.

BACK TO TOP

## Changes in Hearing Thresholds as Measured by Decibels of Hearing Loss in British Army Air Corps Lynx and Apache Pilots

Military Medicine
LTC Gregory T. Lang, MC USA; Lt Col Michael J. Harrigan, RAMC
Nov 2012

ABSTRACT

Objective: Helicopter pilots are exposed to noise at work and are at risk of developing hearing loss in excess of that which naturally results from aging. We investigated whether Lynx pilots demonstrated changes to hearing thresholds that differed from Apache pilots. Methods: Survey responses were combined with audiometric data from a retrospective cohort of 59 Lynx and 87 Apache pilots. Subjects' audiograms were analyzed for air conduction thresholds with age correction performed in accordance with ISO 7029. Annual changes in low frequencies (0.5–2 kHz) and high frequencies (3–6 kHz) were calculated. Subjects were categorized for time in service and flying hours. Results: Hearing was better than predicted at nearly all frequencies in both ears for Lynx and Apache pilots. There were no differences in hearing between groups of pilots. Significant differences in hearing threshold changes existed for pilots with 20 or more years of service compared to those in other categories. Discussion: The results suggest that the circumaural earmuffs currently incorporated into the flying helmet mitigate the risk of noise-induced hearing loss in these pilots.

INTRODUCTION

Military rotary wing pilots are required to perform their job in an extremely noisy environment in which the principal form of hearing protection remains the use of personal protective equipment (PPE). Noise attenuating PPE is typically worn in the form of

circumaural earmuffs integrated into the aircrew helmet system and offers the simplest, least expensive, and most operationally effective method of providing hearing protection for military helicopter aircrew.

The high noise levels to which aircrew personnel are subjected can lead to a number of problems. Interference with speech and nonspeech communications, such as auditory warning alerts, can have both safety of flight and operational implications. Over time, the exposure to aircraft noise has the potential to cause permanent hearing damage with resultant long-term occupational and social repercussions. Several studies have corroborated an association between hearing loss and an occupation in military aviation.

The noise dose received by the crewmember is a combination of both ambient aircraft noise transmitted through the helmet and the communications signal that is delivered directly to the ears from within the helmet. Mean noise levels for turboshaft powered, single rotor airframes approach 100 dB(A) and are most intense at lower frequencies. Associations with cumulative flying hours, years of flying service, type of aircraft, and years of military service have been implicated as risk factors for the development of noise-induced hearing loss (NIHL) in aircrew personnel.

Although it seems axiomatic that aircraft noise exposure would be the major factor associated with hearing loss in military aircrew, the reported results have not always supported this assumption. Sutherland and Gasaway reported that noise-exposed U.S. Air Force personnel had better hearing than the general U.S. population. Ribak et al concluded that hearing loss in Israeli pilots was the result of presbyacusis and not a factor of aircraft noise or the accumulation of flying hours. Kuronen et al found that Finnish military pilots had better hearing than predicted using an international standardization for estimation of occupationally related NIHL. Barney and Bohnker demonstrated that U.S. Marine Corps aviation personnel were no more likely to have elevated hearing thresholds than other Marines.

One possible explanation for these unexpected findings is the improved level of noise attenuation offered by the modern military flight helmet. Flying helmets originated as a means to protect against the weather and were originally made of canvas or leather. In contrast, modern flight helmets are constructed from advanced technology materials that provide crash protection for the head, act as a mount for visual enhancement devices, and provide built-in communications and noise attenuation systems. Importantly, the performance of the helmet as means of providing hearing protection depends not only on the design and material characteristics of the shell and the circumaural pads, but also on the capability to afford a customized fit and the proper wear by the user.

Currently, the British Army Air Corps (AAC) employs two helicopters in an operational role. Pilots of the Lynx wear the Mk4 flying helmet, whereas pilots of the Apache wear the Integrated Helmet and Display Sighting System (IHADSS). These helmets differ in their noise attenuation characteristics. Additionally, differences in helmet fit and wear may result in noise leakage around the earmuffs and may be a risk factor for hearing loss.

Based on technical data analyzed by another agency, the Mk4 was reported to have a noise reduction rating (NRR) of 22 dB, whereas the NRR for the IHADSS was reported to be 19 dB. Though the difference may not seem like much, it is important to note that current U.K. Control of Noise at Work Regulations stipulate that an increase of 3 dB in recorded noise equates to a doubling of the sound pressure level and a requirement to halve the permissible exposure time.Therefore, the purpose of this retrospective cohort study was to assess changes in hearing thresholds that could not be accounted for by age-related hearing loss and to determine if significant differences existed between British Lynx and Apache pilots. The results may be useful for providing guidance regarding hearing conservation measures and use of noise attenuating PPE in the military aviation environment.

## METHODS

All British AAC pilots are administered screening pure-tone air-conduction audiograms as part of the aircrew selection process and subsequent annual flying duty medical examinations. After obtaining approval to conduct the study through the Ministry of Defense Research Ethics Committee, the audiograms from all available medical records for AAC pilots currently flying either the Lynx or Apache helicopter were reviewed. Of the 150 records that were initially identified, three were excluded for a documented history of otologic pathology and one was excluded for incomplete data. The remaining 146 records consisted of 59 Lynx and 87 Apache pilots. Both groups of helicopter pilots had previous exposure to fixed wing aircraft and other rotary wing types as part of their military flying training, though in most cases, this was less than 200 hours all flown with the Mk4 series helmet.

Existing audiogram data for frequencies ranging from 500 Hz to 6,000 Hz collected as part of the routine aircrew medical examinations were reviewed. Values at 8 kHz were not consistently recorded in the medical record and were therefore not analyzed. For all pilots, the most recent audiogram and the audiogram dated closest to the start of flight training were selected for data analysis. For Apache pilots only, an additional audiogram closest to the date of conversion to type was selected. All audiograms were performed at U.K. military medical centers in a sound-insulated audiometry booth. However, specifics of the testing equipment used are known only for the most recent audiograms following the standardization of audiometers in 2008 to the Amplivox CA850 Series 4, which uses a modified Hughson–Westlake technique for automatic pure-tone threshold detection (Amplivox, Oxford, England). Any abnormalities documented in the medical records with respect to the ear, nose, and throat examination were recorded for review in order to maintain a study population of otologically normal persons in accordance with ISO 7029 validity criteria.

Demographic data from a self-reported survey collected as part of this same study were categorized for years of service (<10, 10 to 19, and 20 or more) and flying hours (<1000, 1000 to 1999, and “2000 or more”) to facilitate analysis. Approximately 200 surveys were distributed. Exact figures of how many pilots received the survey cannot be ascertained as some were sent electronically, and

others never made it to eligible subjects who were deployed in support of ongoing operational requirements. A total of 101 surveys were returned of which 34 Lynx and 47 Apache responses could be matched to the 146 subjects with audiogram data previously described.

All audiogram data were age-corrected according to the formulas detailed in ISO 7029—statistical distribution of hearing thresholds as a function of age. This standard has been used previously to control for the effects of age on hearing threshold shifts in order to assess for other independent variables. Using the subject's age at the time of the audiogram, the median thresholds for 0.5, 1, 2, 3, 4, and 6 kHz were calculated according to ISO 7029. These figures were then subtracted from the measured hearing threshold of the same frequency as recorded on the subject's audiogram. This produced unique age-corrected audiograms for each subject at various points in time.

The age-corrected thresholds for each audiogram were then summed to create a low frequency (LF) value (0.5 + 1 + 2 kHz) and a high frequency (HF) value (3 + 4 + 6 kHz) according to the medical grading standards used by the AAC. By subtracting earlier LF and HF values from more recent values and dividing the result by the interval in years between the two measurements, we were able to calculate age-corrected LF and HF threshold changes per year. Flying career threshold shifts were calculated as the difference between the earliest and the most recent age-corrected audiograms. For Apache pilots only, the age-corrected audiogram closest to, but not more than 12 months from, the date of Apache conversion training was used to calculate “after conversion” threshold shifts. These figures were used to evaluate for a possible association between the hearing threshold shifts and the type of aircraft flown. Other factors such as the number of years of military service and total flying time were also analyzed for association with hearing threshold shifts.

Statistical analysis was performed using Predictive Analytics Software. The Student's two-tailed *t*-test for paired samples was used for statistical analyses of threshold changes between audiograms of the same subject. Statistical analyses of threshold changes between the two groups of Lynx and Apache pilots were performed using Student's two-tailed *t*-test for unpaired samples with unequal variance. Analysis of variance was used for comparison of hearing threshold shifts with the categorical variables of “years of service” and “flying hours.” Power analysis was performed using Power Analysis and Sample Size 2000. The high variance of hearing levels in healthy individuals resulting in large standard deviations (SD) has been previously identified. Therefore, a difference of 10 dB or more at any individual frequency in age-corrected hearing thresholds between the two groups of Lynx and Apache pilots was chosen to be clinically significant. In order to have a power of 90% at a significance level of 0.05, a group size of 22 subjects was necessary for SD = 10, whereas a group size of 48 subjects would be necessary for SD = 15.

RESULTS

The audiograms from 59 Lynx and 87 Apache pilots were analyzed. The mean (range, SD) of pilot age, years in service, and total military flight time in hours. The mean interval between the earliest and most recent audiograms was 12.0 (3–31, 8.2) years for Lynx pilots and 10.9 (2–32, 6.6) years for Apache pilots. .

For Lynx pilots, the hearing threshold for the left ear at 6 kHz was the only value that was significantly different than the ISO predicted values ($p = 0.01$). For Apache pilots, threshold values at 2 kHz and 6 kHz for the right ear were significantly different than the ISO predicted values ($p = 0.01$ and $p = 0.03$, respectively); however, the value at 2 kHz was less than the ISO predicted value. The only hearing threshold value that was significantly different between Lynx and Apache pilots was the left ear at 6 kHz, the Apache pilots' threshold being less than that of the Lynx pilots'.

For the group of all pilots ($n = 146$), threshold values of LF and HF were compared between the initial and most recent audiograms. Values of LF were significantly different for both the right and left ear ($p < 0.01$) with the mean values for the right ear being 11.85 dB initially and 3.29 dB recently, and the means for the left ear being 11.44 dB initially and 5.89 dB recently. The values of HF were not significantly different. These findings remained consistent when pilots were grouped by aircraft type, and there were no differences between the two groups.

Age-corrected hearing threshold values demonstrated no significant differences between groups of Lynx and Apache pilots in either ear at all frequencies for both the initial and the most recent audiograms. Within groups analysis revealed that there were statistically significant differences in hearing threshold values at all frequencies for both Lynx ($p < 0.01$ for all 6 frequencies) and Apache pilots ($p < 0.02$ for all 6 frequencies) when comparing the initial to the most recent age-corrected audiograms; however, the direction of the change was exactly opposite of that which was expected. For all frequencies in either ear in both groups of pilots, the means of the most recent age-corrected thresholds were significantly less than the means of the initial age-corrected thresholds ($p < 0.05$).

Again, there were no significant differences between groups of pilots. Note that the age-corrected LF and HF threshold values for the most recent audiograms of both groups of pilots are “less” (i.e., indicates better hearing) than the age-corrected threshold values derived from the initial audiograms. For Apache pilots, there were highly significant differences in mean age-corrected LF and HF values for both ears when comparing the most recent audiogram to the one obtained closest to the date of Apache conversion ($p < 0.01$ for all values) demonstrating an improvement in LF and HF hearing thresholds.

Analysis of variance showed consistently significant differences for age-corrected LF and HF threshold shift per year in both ears for the group of all subjects when “years of service” was selected as an independent categorical variable. Post hoc Games–Howell

testing for unequal variances confirmed statistical significance with the means and 95% confidence intervals. Note that all values of the shift per annum are negative indicating an overall “improvement” in age-corrected hearing thresholds. Post hoc significance was not maintained for the independent variable of “flying hours.”

## DISCUSSION

Hearing thresholds between 0.5 and 4 kHz in British AAC Lynx and Apache pilots are not significantly different than the median age-specific value as predicted by ISO 7029 after a mean time in service of 17.9 and 15.8 years, respectively. The only exception is for Apache pilots' right ear at 2 kHz where their hearing threshold is better than the predicted value. At 6 kHz, the Lynx pilots' right ear and the Apache pilots' left ear demonstrated no significant difference compared to their respective ISO predicted values; whereas, threshold values in Lynx pilots' left ear and Apache pilots' right ear were significantly worse than the predicted values. However, the difference at 6 kHz was only 5.21 dB for Lynx pilots and 3.03 dB for Apache pilots. This difference, although statistically significant, is likely not clinically relevant as it has been shown that hearing thresholds can vary by as much as 10 dB between two appropriately calibrated audiometers.

Unexpectedly, this study revealed that the age-corrected threshold values derived from the most recent audiograms for the six pure-tone frequencies tested between 0.5 and 6 kHz, as well as the summations that yielded values for LF and HF were lower than the age-corrected threshold values from the initial audiograms. It might be argued that the initial audiograms were obtained using audiometric equipment that differed from current testing apparatus making it inappropriate to compare the two sets of data. This is a valid concern, and we suspect that the initial audiograms recorded thresholds that were erroneously high. With the employment of standardized audiometric equipment and sound-insulated testing booths, the threshold values recorded on the most recent audiograms provided a more sensitive measure of hearing at thresholds below 10 dB. Thus, age-corrected threshold values would erroneously show a decrease over time.

Whether our suspicion that the initial audiogram values are spuriously high is correct or not, based on the most recent audiograms of Lynx and Apache pilots it is possible to conclude that the AAC hearing conservation program in combination with the current hearing protection afforded by the Mk4 and IHADSS flying helmets is appropriately preventing hearing loss beyond that which can be attributed to the natural aging process for frequencies between 0.5 and 4 kHz. At 6 kHz, the slight increase in hearing threshold is so minimal as to render it clinically insignificant. These findings are consistent with several other studies that have found aircrew personnel to have no increased risk for NIHL. Wagstaff and Arva proposed that civilian rotary wing pilots' use of circumaural headsets provided a higher level of noise attenuation and resulted in these pilots having hearing that was no worse than civilian

airline pilots and air traffic controllers. Such headsets are very similar to the ones incorporated into military flying helmets like the Mk4 and IHADSS. Looking at machinists and wood workers with high levels of occupational noise exposure, Johansson and Arlinger concluded that workers' increased knowledge about the risk of NIHL and their increased tendency to wear hearing protection contributed to the improvement of hearing thresholds over the course of 2 decades.

The AAC has operated a robust hearing conservation program supported by legislated requirements for over 30 years. Current directives stipulate an Exposure Limit Value of 87 dB(A) for continuous noise measured at the ear averaged over an 8-hour working day. The recent approval for use of the Communications Ear Plug (CEP) by Communications & Ear Protection for both flight helmets is an opportunity to provide higher levels of noise attenuation for aircrew personnel. Their use by Apache pilots began within 1 year of data collection, so it is very unlikely that the CEP had any effect on the results of this study, but future studies should evaluate their capability to impart additional hearing protection. Pilots are required to have their helmets checked for proper fit and wear every 15 weeks, and as previously mentioned, they receive an annual screening audiogram that is then reviewed during their physical examination by a physician trained in aviation medicine. These encounters reinforce the role of PPE in reducing the risk of NIHL, and it is believed that they have changed behavior not only in the workplace, but also in off-duty environments as well.

Risk factors associated with changes in aircrew hearing are not always consistent across the many published studies. In this study, we found that pilots with 20 or more years of military service had annual LF and HF age-corrected threshold changes that were significantly different from pilots with less than 10 years of service, and in most cases these changes were significantly different from pilots with 10 to 19 years of service. Until recently, it was fairly typical for younger pilots to have fewer hours than older pilots who had more time in military service. Hence, potential risk factors such as age, years in military service, and total flying hours were directly proportional. However, operational requirements in the last decade have changed this pattern such that now many junior pilots are logging flying time that only senior aircrew in years gone by would have accumulated. This may provide a partial explanation for why we found no association with total flying time.

During the progress of this study, the RAF Centre of Aviation Medicine tested the attenuation characteristics of several flying helmets and reported a NRR of 25 dB and 27 dB for the Mk4 and IHADSS, respectively. When worn with the CEP, the NRR reached 36 dB and 39 dB, respectively. These figures indicate that both flying helmets are providing greater hearing protection than previously anticipated. From our findings, it is reasonable to conclude that flying military helicopters and the concomitant exposure to aircraft noise are not necessarily associated with hearing loss provided proper PPE is utilized. Reports that found only age-related hearing loss in military aircrew and one that found hearing to be better than predicted according to ISO 1999 lend support for this conclusion. Furthermore, there is no indication that any AAC pilot has been medically discharged from duty in the last 30 years as a result of a

hearing loss. Since 1999, there have been four pilots who have been medically downgraded for poor hearing; however, all four remain on active flying status.

This study had several limitations. As with many previous publications, occupational and recreational noise exposure data are not available for the personnel involved. The audiograms taken before 2008 were recorded using several different models of audiometers under various testing conditions. It is suspected that the newer automatic audiometers are more accurate particularly at lower thresholds erroneously leading to an observed “improvement” in hearing over time. Unfortunately, such changes in technology and policy cannot be controlled in the nonlaboratory environment and present a challenge to retrospective studies of medical data. Another limitation of this study is related to the data collected from the survey. The distribution of surveys was facilitated by electronic transmission; however, it makes calculation of a return rate near impossible. Furthermore, survey data are not available for all subjects, and those who did respond may have a bias in their answers. However, post hoc power analysis indicates that threshold differences as small as 3.8 dB were detectable.

The findings of this study should not abrogate the need for continued health surveillance, risk communication, or research and development of improved PPE. Future studies should evaluate nonrated aircrew personnel as well as the long-term effects of the very recent introduction of CEP to the AAC flying helmets. Similar results would be expected to be found in U.S. military helicopter pilots, who have been using CEP for over a decade.

In summary, the hearing of 146 British AAC pilots followed for a mean time of greater than 10 years is no worse than the hearing of a corresponding standardized reference population of equivalent age over the 0.5 to 4 kHz frequency range, with only minor clinically insignificant variations observed at 6 kHz. Hearing thresholds did not differ between the groups of Lynx or Apache pilots indicating that the Mk4 and IHADSS flying helmets provide similar levels of noise attenuation in the work place. After correcting for age using ISO 7029, we found that the annual changes in LF and HF hearing thresholds were associated with the number of years in military service with AAC pilots having 20 or more years of service differing from pilots with fewer years of military service. We found no such associations for total flying hours or type of aircraft flown. We conclude that the existence of a comprehensive hearing conservation program in combination with the noise attenuation afforded by the Mk4 and IHADSS flying helmets are the principal factors that mitigate the risk of NIHL as a result of an occupation in military aviation. Military helicopter pilots using equivalent (or better) noise attenuating PPE would be expected to show similar results.

BACK TO TOP

## Maggot Debridement Therapy in Modern Army Medicine: Perceptions and Prevalence

Military Medicine
Rae A. Heitkamp, BS; George W. Peck, PhD; CPT Benjamin C. Kirkup, MS USA
November 2012

## ABSTRACT

Maggot debridement therapy (MDT), despite its long history and safety profile, finds limited use in the military health care system. Although new methods are continually being investigated to debride wounds more quickly and effectively, MDT remains largely a therapy of last resort. We evaluated the frequency of MDT in the Army sector of the MHS and the decision-making process surrounding its use. A 22 question survey of Army physicians was prepared and distributed through select Medical Corps Consultants in specialties likely to practice debridement. 83% of respondents were familiar with MDT, and of those familiar, 63% were aware of FDA approval for the product and 10% had used the product themselves. The three most frequently cited reasons for not using the therapy were no need (52%), no access (23%), and insufficient experience (19%). Informing the 37% of physicians who are not aware of FDA approval is an obvious target for program improvement. However, as many do not find a need for MDT, targeted improvements to MDT access and education for those physicians who encounter indications for MDT would permit them to apply MDT where there is an unmet need.

## INTRODUCTION

Antibiotic resistant infections are a major concern in military medicine. Both Wounded Warriors returning to the United States and Veterans with chronic health issues such as diabetes or amputation are susceptible to multidrug-resistant bacteria. Because of the failure of antibiotics to treat these resistant infections, the biomedical research community is increasingly turning its attention to other antimicrobial therapies.

One alternative to antibiotics for infected wounds is maggot debridement therapy (MDT), also known as biodebridement, biosurgery, and larvae therapy. MDT is FDA approved, is also approved in Europe and the United Kingdom, and was used widely before antibiotics were discovered. It is experiencing a renaissance in the antibiotic resistance era. MDT consists of the application of sterile fly maggots to a wound, construction of an enclosure around the treatment, and removal and replacement of maggots every 48 to 72 hours. MDT appears to heal wounds by at least three mechanisms: debridement of necrotic tissue, disinfection of the wound bed, and stimulation of healing.

In spite of its low cost and high treatment efficacy, current knowledge indicates that MDT is rarely considered by physicians until alternative treatments have failed. If the therapy were used earlier in the treatment cycle, perhaps patients could be spared amputation revisions and days in the hospital. The aversion to this therapy is not attributable to any unreasonable risks to the wellbeing of patients or health professionals; it seems to be, rather, an image problem of the maggot itself, a superficiality that is incongruent with the aims of medicine. The reasons why MDT is not used more widely in medicine require further elucidation.

While not the first choice for wound debridement of many clinicians, MDT has a long record of efficacy and favorable opinion in clinical practice. An early survey of surgeons published in 1935 gave MDT high approval ratings. A more recent survey published in 2007 focused on veterinary use and reported favorable MDT effects. Military medical facilities and organizations offer a promising environment for the use of MDT because of the nature of the patient population and the implicit organizational disinterest in purely cosmetic concerns during therapy. In order to assess the experience of military medical organizations with MDT, we surveyed a sample of Army practitioners that we considered likely to encounter indications for this therapy. As the active prescribers of wound therapy within the Army medical establishment, these practitioners represent the operational attitudes and effective posture of the organizations in which they serve. The purposes of the survey were to gauge the current practice of MDT within the Army and to assess strategies to promote the use of MDT in military medicine.

## METHODS

### Mode of Distribution

We developed an anonymous survey through Survey Monkey (http://www.surveymonkey.com/) and distributed it via e-mail to Medical Corps Consultants (MCC) in the U.S. Army. We requested that MCCs distribute the survey to physicians working under their authority. We collected responses from May 4, 2011 to July 5, 2011.

### Population and Sample

We contacted MCCs in the following departments for survey distribution: Anesthesiology, Cardiothoracic Surgery, Clinical Investigation, Emergency Medicine, General Surgery, Infectious Disease, Internal Medicine, Neurosurgery, Orthopedic Surgery, Otolaryngology/Head and Neck Surgery, Pain Management, Plastic Surgery, Transplant Surgery, Trauma, Urological Surgery, and Vascular Surgery. The survey link did not preclude responses from physicians in other specialties or from outside the Army, but because of the low number of responses from the other branches ($n$ = 3), non-Army physicians' responses were removed from the study population.

Survey Contents

The survey had a branching structure and included both multiple choice and free-response. Respondents would answer at most 22 questions, depending on the branches along which the survey directed them. Questions covered demographic information, experience with MDT, exposure to MDT, and opinions about the use of MDT in the hospital. After completion, the survey directed respondents to the BioTherapeutics Education and Research Foundation homepage (www.bterfoundation.org), a physician resource for MDT information and guidance.

Data Analysis

Before analysis, we set aside data for respondents who failed to answer more than three of the demographics questions ($n = 2$). We conducted tests of independence (G-statistic with Williams correction) in Microsoft Excel for 2 × 2 contingency tables. For tables found to be significant ($p < 0.05$), we used logistic regression in *R* (version 2.14.0; The R Foundation for Statistical Computing, http://www.r-project.org/) to determine odds ratios with 95% confidence intervals.

We processed free-response data from question 19 ("Why have you never attempted to use this therapy?") using published methods as follows: We recruited volunteers unassociated with the survey development or distribution. Individual volunteers grouped all physicians' responses to question 19 ($n = 125$) into as many distinct categories as they thought were appropriate. Then they regrouped the responses into the consensus categories selected from the first round of categories submitted. We allowed volunteers to select more than one category for the physician responses because many physicians listed multiple items in their responses. We considered a response reportable by category if more than half of the seven volunteers classified it in that category.

RESULTS

In total, 180 Army physicians answered the MDT survey. Because of the branching of the survey and the option of not responding to individual questions, the sample size of responses varies among the questions included in the survey. We collected responses from medical treatment facilities in each of the four regions of MEDCOM and from overseas installations in Korea, Afghanistan, and Iraq. We contacted 16 MCCs and received responses from physicians in 7 specialties.

The prevalence of MDT in Army medicine was high in our sample. The majority of survey respondents were familiar with MDT before beginning the survey: 82.7% (of 179) reported that they had heard of or used the therapy in the past. A total of 10.4% (of 144) responded that they had practiced MDT themselves in the past, and 24.5% (of 143) was aware of colleagues who had used the

therapy. Accuracy of knowledge about MDT was also high in the sample. Most respondents were aware that MDT is FDA approved (63.2% of 144); however, when presented with statistics on the rates of usage of MDT worldwide near the end of the survey, 56.1% (of 171) was surprised at the prevalence of the therapy worldwide, expecting the rates to be lower.

We ran eight tests of independence to assess differences in groups of respondents relative to MDT awareness and use. Using two-way tables, we tested the impact of specialty and demographics on MDT practice. Of the tables analyzed, the only significant group difference was in the perceived value of an MDT network of practitioners between younger and older physicians: Physicians less than or equal to 40 years old were 3.22 times (95% CI, 1.348–7.692; G = 6.810, $df = 1$, $p < 0.05$) more likely than doctors older than 40 to have a favorable perception of the utility of more MDT network resources.

Seven independent observers ("volunteers") grouped data for the question "Why have you never attempted to use this therapy?" which received a substantial number of physician responses, into seven distinct categories among the free-response data. For question 19, there were 147 instances (seven categories × seven volunteers; $n = 123$ physician responses evaluated) where more than half the volunteers ($n \geq 4$) were in agreement about the categorization of a given physician response.

An additional set of multiple-choice questions evaluated how organizational behavior would impact the use of MDT within military medical organizations. A total of 85.2% (of 169) indicated that having access to a network of MDT practitioners would be a good resource, and one-third (48 respondents) of that group said that a network of practitioners would increase their own use of the therapy. Additionally, 71.9% ($n = 23$) of physicians who knew others who had successfully applied MDT said they were more likely to use MDT because of their colleagues' success.

## DISCUSSION

We are encouraged to find that MDT is already used considerably in Army wound care, and that even where rate of use is low, the community is typically familiar with the therapy. The explosion of antibiotic resistance in pathogenic microbes has created a motivation for the use of other therapies in wound treatment. The military has a particular need for research in this field because of the nature of modern combat wounds and the care of veterans. The development of novel therapeutics takes years, thus the re-evaluation, optimization, and repurposing of already-approved drugs and devices for wound treatment is an attractive alternative to new development.

MDT is an effective treatment that is already FDA approved as a medical device. There is, however, very little comparable data on its use in the medical community. Most of the studies on MDT in the clinic use slightly different methodologies or estimates of success,

making direct comparison of results challenging. However, as far back as 1935 physicians have been cognizant of the general value of MDT practice. Robinson conducted a survey of 947 North American surgeons known to have employed MDT. Of the 635 responding surgeons, 91.2% expressed a favorable opinion to MDT; only 4.4% expressed an unfavorable view. In reply to the question: “Do you have any specific objection to this method?” 72% of the respondents answered “No.” Of the respondents objecting, the most common complaints raised by surveyed practitioners were the cost of the maggots, the time and effort required to construct the maggot dressings, and discomfort to patients. In 2007, Dr. Ronald Sherman and others surveyed 23 wound care therapists, 30 nonwound care clinicians, and 66 nonclinicians. They reported that 72% said they would use MDT in the future, 6% said they used MDT previously, and 22% said they would never use it. Most wound care therapists who had not used MDT before cited the following reasons: they thought their colleagues and/or facilities would not allow it, poor reimbursement, or they thought their patients would not want it. Most nonwound care clinicians stated that they do not use MDT because they either did not care for eligible patients or because patients would not agree to MDT. However, among nonclinicians (e.g., potential patients), only 14% said they would refuse MDT; 86% said they would accept or consider MDT if recommended by their doctor. Sherman et al concluded that many therapists are not making use of MDT because of real or perceived barriers by their colleagues or facilities, whereas other therapists are misinformed about costs, reimbursement, availability, or patient acceptance. They also emphasized that better education is essential if wound care therapists are to make use of MDT (personal communication).

The first step to identifying the apparent public relations problem this therapy experiences is to assess organizational attitudes. Our survey is, we believe, the only modern survey of MDT practice in the military. The results indicate that some of the previously reported characterizations of MDT resistance (in particular, the “yuck” factor) are not particularly prevalent among Army physicians and thus do not restrict the effective application of cosmetically unappealing therapies in military medical organizations. According to our survey, physicians are not using MDT because they have not been trained to use it and do not know how to access it. This is a far more attractive problem to approach than the psychological aversion to maggots in wounds. It points to some obvious organizational initiatives (e.g., training), which could improve military medical effectiveness and efficiency.

A substantial contingent of the respondent pool noted that they believe patients will be resistant to this mode of treatment. However, the results of a small survey published in 2005 indicate that resistance to this therapy lies more on the practitioner side than with the patient. Similar asymmetries in physician and patient perception in the broader context of medical care were later characterized by Lesho et al Having addressed the organizational attitude about MDT, in future studies we would like to survey wounded patients to identify their perceptions of MDT. If the military health care organizations and military patients are unknowingly in agreement about the attractiveness of MDT as a therapy to treat nonhealing wounds, the remaining obstacles to expanding the use of MDT in medicine are training and access.

BACK TO TOP

***PTSD***

## Amygdala Volume Changes in Posttraumatic Stress Disorder in a Large Case-Controlled Veterans Group

Archives of General Psychiatry
Rajendra A. Morey, MD, MS; Andrea L. Gold, MS; Kevin S. LaBar, PhD; Shannon K. Beall, BS; Vanessa M. Brown, BA; Courtney C. Haswell, MS; Jessica D. Nasser, BA; H. Ryan Wagner, PhD; Gregory McCarthy, PhD
November 2012

ABSTRACT

Context  Smaller hippocampal volumes are well established in posttraumatic stress disorder (PTSD), but the relatively few studies of amygdala volume in PTSD have produced equivocal results.

Objective  To assess a large cohort of recent military veterans with PTSD and trauma-exposed control subjects, with sufficient power to perform a definitive assessment of the effect of PTSD on volumetric changes in the amygdala and hippocampus and of the contribution of illness duration, trauma load, and depressive symptoms.

Design  Case-controlled design with structural magnetic resonance imaging and clinical diagnostic assessments. We controlled statistically for the important potential confounds of alcohol use, depression, and medication use.

Setting  Durham Veterans Affairs Medical Center, which is located in proximity to major military bases.

Patients  Ambulatory patients (n = 200) recruited from a registry of military service members and veterans serving after September 11, 2001, including a group with current PTSD (n = 99) and a trauma-exposed comparison group without PTSD (n = 101).

Main Outcome Measure  Amygdala and hippocampal volumes computed from automated segmentation of high-resolution structural 3-T magnetic resonance imaging.

Results  Smaller volume was demonstrated in the PTSD group compared with the non-PTSD group for the left amygdala (*P* = .002), right amygdala (*P* = .01), and left hippocampus (*P* = .02) but not for the right hippocampus (*P* = .25). Amygdala volumes were not associated with PTSD chronicity, trauma load, or severity of depressive symptoms.

Conclusions  These results provide clear evidence of an association between a smaller amygdala volume and PTSD. The lack of correlation between trauma load or illness chronicity and amygdala volume suggests that a smaller amygdala represents a vulnerability to developing PTSD or the lack of a dose-response relationship with amygdala volume. Our results may trigger a renewed impetus for investigating structural differences in the amygdala, its genetic determinants, its environmental modulators, and the possibility that it reflects an intrinsic vulnerability to PTSD.

BACK TO TOP

## It's not the stress that counts, it's whether you can control it

Scientific American
6 Nov 2012

Stress is generally not a good thing. Most of us who live stressful lives (which, I suppose, would be all of us), are well aware of this. We try to reduce our stress, or even stress about how stressed we are. Traumatic stress increases the risk for all sorts of psychiatric disorders, including major depressive disorder, anxiety, and post traumatic stress disorder. But not all stresses are created equal, even the traumatic ones. And it turns out that it's not the stress itself that is important…it's whether or not you have any control over it.

Varela et al. "Control over Stress, But Not Stress Per Se Increases Prefrontal Cortical Pyramidal Neuron Excitability" Journal of Neuroscience, 2012.

A stress that you can control is a very different one from a stress that you can't. I usually think of a stress you cannot control as something like the illness of a family member, as compared to a stress you can control, say, the stress involved in training for and running a marathon (which is definitely a physical stressor). These are both stresses, but they aren't alike. While the stress that you cannot control is often a very traumatic experience, and can predispose people to psychiatric disorders, a controllable stress is

actually a good event. Not only does it blunt the impact of the stressor itself, it can be protective against the detriments of future uncontrolled stresses. Scientists call this “behavioral immunization” against future stress.

Behavioral immunization involves the recruitment of very specific brain regions, especially the medial prefrontal cortex of the brain.

After exposure to a controllable stress, there is increased activity in the medial prefrontal cortex, and it is thought that the increase in activity is important for the development of behavioral immunization. If you stop this increased activity from taking place during controllable stress, you can prevent behavioral immunization.

But the medial prefrontal cortex is not a single area. Instead, it is divided into the infralimbic cortex and the prelimbic cortex. Each of these areas plays a different role in emotionally processing and expression. But which one is involved in behavioral immunization?

To look at this, the authors of this study exposed rats to either escapable or inescapable stress. The rats are put in a chamber where they will receive a tail shock at random intervals. Half of them have no choice, but the other half have a little wheel. When they turn the wheel, the shock will stop, allowing them to escape the stress. They also had another groups of control rats that never got the shock at all.

After exposure to escapable or inescapable stress, the authors used electrophysiology to see how the neurons in the prelimbic cortex responded. Electrophysiology is a very widely used technique in neuroscience, and is a fascinating way to peer inside the electrical life of a cell. You take a cell (this study worked in slices, but you can also do it in the whole animals), and very, very carefully impale it with a small glass electrode. The electrode is hollow with a little hole in the tip, and by suctioning very carefully, you can merge your electrode in with the membrane of your cell. From there, you have access to all the electrical impulses that run across the cell membrane. You can then take a stimulating electrode, stimulate somewhere else, and watch the cell respond. How much and in what way the cell responds tells you not only what kind of cell it is, but how responsive it is following something like inescapable stress.

The authors found that rats exposed to escapable stress had the makings of behavioral immunization beginning in the prelimbic cortex. When rats were allowed to escape the stress, after the stress the neurons in the prelimbic cortex showed increased electrical activity. The neurons were more excitable. This is a state that has been linked to the development of behavioral immunization. In contrast, the animals that could not escape the stress showed no changes when compared to control, nonstressed rats. This would mean that the rats exposed to escapable stress would end up with more responsive neurons in the prelimbic cortex, which might help them adapt to further stressors.

The medial prefrontal cortex in general has a lot of potential in stress research. Just exciting this area is enough to create resilience to stress in animals. It appears that this effect may be particularly important in the prelimbic area. And this could have potential for work in human stress. Not all stresses are necessarily entirely bad. And work in this area could give us important information on

ways to help people who are, or have been, exposed to trauma. If we could produce a behavioral immunization beforehand, for example, we may be able to help people who are likely to go into combat, or if we can target this area to promote stress resilience, we could help those who have already been exposed to traumatic stress.

BACK TO TOP

## Cue-elicited affect and craving: advancement of the conceptualization of craving in co-occurring posttraumatic stress disorder and alcohol dependence.

Behavioral Modification
November 2012

Abstract

Posttraumatic stress disorder (PTSD) commonly co-occurs with alcohol dependence (AD) and negatively affects treatment outcomes. Trauma-related negative affect enhances substance craving in laboratory cue-reactivity studies of AD individuals, but the role of positive affect has not been established. In this study, 108 AD treatment-seeking adults with current PTSD and AD were presented with four counterbalanced trials consisting of an audio cue (personalized trauma or neutral script) followed by a beverage cue (alcohol or water). Results revealed alcohol cues increased positive and negative affect, and positive affective responses explained significant incremental variance in self-reported craving and salivation, but only when cues were accompanied by neutral not trauma imagery. Ambivalent (high negative and positive) responses were associated with strongest craving. Findings advance the conceptualization of craving in individuals with PTSD-AD and highlight the importance of independently assessing positive and negative affective responses to cues in individuals with co-occurring PTSD-AD.

Keywords: PTSD, alcohol abuse, craving, approach-avoidance, cue reactivity, salivation

Researchers, clinicians, and policy makers have been increasingly attentive to the issue of posttraumatic stress disorder (PTSD) and substance use disorder (SUD) comorbidity. Between 20% and 40% of individuals seeking treatment for alcohol and drug use disorders meet current diagnostic criteria for PTSD (e.g., Dansky, Roitzsch, Brady, & Saladin, 1997; Reynolds et al., 2005). Lifetime prevalence rates of PTSD in this population range from 30% to 52% (Back et al., 2000; Reynolds et al., 2005). Alcohol and drug users with PTSD present a more complicated clinical picture and experience poorer treatment outcomes, such that individuals with PTSD improve less in SUD treatment programs, relapse more quickly on discharge, and experience additional difficulties compared with individuals with SUD alone (e.g., poorer occupational functioning, poorer physical health and greater likelihood of overdose; Brown, Stout, & Mueller, 1999; Norman, Tate, Anderson, & Brown, 2007; Ouimette, Finney, & Moos, 1999; Read, Brown, & Kahler, 2004).

PTSD is characterized by experiences of heightened negative affect, as well as constant efforts to manage and/or avoid such states (e.g., Beckham et al., 2000). At least two theoretical models suggest that negative affect plays a pivotal role in the maintenance of PTSD-SUD comorbidity (see Stewart & Conrod, 2003, for review). The self-medication hypothesis proposes that individuals with PTSD drink to relieve distress evoked by traumatic memories (Khantzian, 1997). The mutual maintenance hypothesis, a bidirectional model, suggests that substance use may also interfere with natural resolution of trauma-related distress, and that PTSD symptoms may negatively affect substance use symptoms (Brown, Stout, & Gannon-Rowley, 1998). Consistent with these models, alcohol-dependent individuals report increased alcohol craving in response to stress cues and negative mood induction paradigms (e.g., Fox, Bergquist, Hong, & Sinha, 2007; Rubonis et al., 1994). Similarly, individuals with comorbid PTSD and alcohol dependence (AD) evidence increased self-reported alcohol craving and salivary responding in response to trauma cues, even when no alcohol cue is present (e.g., Coffey et al., 2010), with these responses being dampened after distress elicited by traumatic memories is reduced through repeated exposure to the traumatic memory (Coffey, Stasiewicz, Hughes, & Brimo, 2006). Conversely, exposure to alcohol and drug cues has also been shown to increase negative affect in cuereactivity paradigms. For example, among recently abstinent alcohol-dependent individuals, alcohol cue-elicited craving is related to increases in subjective and physiological anxiety and fear (Fox et al., 2007). Individuals with AD also tend to report increased guilt on exposure to alcohol-related cues (Cooney, Gillespie, Baker, & Kaplan, 1987). Thus, trauma-related negative affect induces craving among SUD individuals, and exposure to drug and alcohol cues in turn exacerbates negative affect.

Although cue-reactivity paradigms have predominantly focused on negative affect, positive affect may also be important to consider. Researchers are moving toward a multidimensional conceptualization of craving that includes alcohol approach and avoidance motivations (Anton, 1999; Breiner, Stritzke, & Lang, 1999; McEvoy, Stritzke, French, Lang, & Ketterman, 2004; Stritzke, Breiner, Curtin, & Lang, 2004; Stritzke, McEvoy, Wheat, Dyer, & French, 2007). From this perspective, anticipated positive (e.g., pleasure, stress relief) and negative (e.g., hangover, interpersonal difficulties) consequences simultaneously attract and deter an individual from drinking (Breiner et al., 1999). Cognitive models of cravings suggest that individuals in treatment for SUD are particularly vulnerable to what are termed abstinence promotion cravings, which are characterized by ambivalence about use (Tiffany, 1990). Essentially, when a SUD individual is confronted by a cue (e.g., beer in the grocery store) he or she experiences conflicting approach (e.g., positive outcome expectancies) and avoidance (e.g., desire to remain sober) impulses that are subjectively experienced as craving (Stritzke et al., 2007; Tiffany, 1990). The ambivalence model (e.g., Breiner et al., 1999) predicts urges to drink to be strongest when positive responses to alcohol are high and negative responses are low (i.e., pure approach motives). Cravings are theoretically weaker when negative responses are high and positive responses are low (i.e., pure avoidance) and when positive and negative responses are low (i.e., indifference). Ambivalence, characterized by strong positive and negative responses, is theorized to be a middle ground between these poles. These models emphasize the importance of examining multidimensional, orthogonal responses to drug and alcohol cues. Despite this, the role of positive affect in craving has not been well established using cue-reactivity paradigms. Fox and colleagues (2007) found that alcohol-dependent individuals reacted to alcohol cue-elicited craving with decreases in positive affect. However, related literature suggests that positive affect, as well as positive outcome expectancies, correspond with increased levels of craving (Johnson & Fromme, 1994; Litt, Cooney, & Morse, 2000). Consistent with this, positive affective stimuli increase alcohol craving in non-treatment-seeking alcoholics (Mason, Light, Escher, & Drobes, 2008). Furthermore,

increases in self-reported positive affect following alcohol cue exposure also have been demonstrated among regular, nonabstinent social drinkers (Kambouropoulos & Rock, 2010; Kambouropoulos & Staiger, 2004). The results of these few studies notwithstanding, no studies to our knowledge have investigated the impact of alcohol cue presentation on positive affect among PTSD-SUD individuals. Given the complexity of these individuals' drinking patterns and motivations, research examining the effects of alcohol cues on positive affect among SUD individuals with PTSD may help advance our conceptualization of craving and decision making about drinking in this population. Examining affective responses in the context of alcohol and trauma cues will be particularly important in this respect.

The relationship between affective ambivalence and craving also requires further study. Ambivalence models of cravings focus on conflicting approach and avoidance inclinations. Affective responses to cues are not directly addressed by these models, but they may influence and/or reflect these motivational dispositions. Indeed, affective ambivalence has been examined in the context of several facets of desire and desire-related behavior (e.g., chocolate consumption, sexual arousal; Hormes & Rozin, 2011; Peterson & Janssen, 2007). Simultaneous positive and negative affective responses to cues have received inadequate attention in alcohol craving research, however, and no studies have examined how these affective responses relate to the ambivalence model. Extension of this model to affect suggests that positive and negative affective responses to cues should be associated with variations in alcohol craving.

Present Study.

Ultimately, understanding affective responses to alcohol and trauma cues will be critical to advancing our conceptualization of co-occurring PTSD and AD. Thus, the first aim of the current study was to examine affective responses to alcohol and trauma cues in treatment-seeking individuals with concurrent PTSD and AD. Toward this aim, alcohol-dependent substance abuse treatment seekers with PTSD were presented with alcohol or water in vivo cues, combined with personalized trauma and neutral imagery cues. Alcohol craving, along with positive and negative affect, was measured in response to each cue combination. It was hypothesized that (a) the combined trauma script-alcohol cue would correspond with the highest levels of alcohol craving and negative affect compared with the other cue combinations and (b) alcohol cues would elicit higher levels of positive affect compared with the neutral (i.e., water) cue combinations.

The second aim was to examine the relationship between cue-elicited affect and alcohol craving. PTSD-related negative affect is known to be a strong predictor of SUD craving. However, recent models of cravings highlight the importance of examining multidimensional, orthogonal responses to cues. As such, craving may be better understood as a function of positive and negative affect. Accordingly, we hypothesized that increases in positive affect would be related to increases in craving, above and beyond variance accounted for by increases in negative affect.

The final aim of the current study was to examine cue-elicited alcohol craving as a function of response types defined by the ambivalence model of craving. This model implies that urges to drink occur in a multidimensional space, with stronger urges

occurring in the context of either pure approach (i.e., high positive, low negative) or ambivalent (i.e., high positive, high negative) responses. Extending this model to affective responses, we examined how patterns in affective responsivity related to cue-elicited craving. Participants' affective reactions to cues were classified as "positive" (high positive/low negative), "negative" (low positive/high negative), "ambivalent" (high positive/high negative), or "indifferent" (low positive/low negative). It was hypothesized that cue-elicited cravings would be strongest among "positive" and "ambivalent" affective responders.

## Method

### Participants

Participants were 108 adults (50 women) meeting Diagnostic and Statistical Manual of Mental Disorders (4th ed.; DSM-IV; American Psychiatric Association, 1994) criteria for current AD and PTSD who were recruited from a community residential chemical dependency treatment facility as a part of a larger, ongoing study. Participants were required to report at least one heavy drinking day in the past 60 days. To eliminate acute effects of intoxication or withdrawal, participants were also required to be alcohol and drug free for at least 4 days immediately preceding the cue-reactivity assessment. Exclusion criteria included meeting diagnostic criteria for a psychotic disorder, currently experiencing a manic episode, or current use of benzodiazapines or other medications that could interfere with cue-elicited craving or salivation (e.g., naltrexone, tricyclic antidepressants, anticholinergic medication; see Table 1 for demographics).

### Measures
### Screening

The PTSD Checklist (PCL; Weathers, Litz, Herman, Huska, & Keane, 1993), a reliable, valid, and widely used PTSD screening tool (e.g., Blanchard, Jones-Alexander, Buckley, & Forneris, 1996), and the Alcohol Use Disorders Identification Test (AUDIT; Saunders, Aasland, Babor, de la Fuente, & Grant, 1993), a 10-item self-report screening measure for alcohol problems that was developed by the World Health Organization, were used to screen prospective participants. Individuals whose scores were ≥44 on the PCL (Blanchard et al., 1996) and ≥8 on the AUDIT (Babor, de la Fuente, Saunders, & Grant, 1992) were scheduled for a comprehensive assessment.

### Diagnosis

The Computerized Diagnostic Interview Schedule (C-DIS; Robins et al., 2000) was used to diagnose SUDs. The C-DIS is a computerized version of the Diagnostic Interview Schedule (DIS), which is a structured clinical interview for the assessment of DSM-IV Axis I disorders. It has demonstrated sound psychometric properties in the diagnosis of substance abuse and dependence (e.g., Vandiver & Sher, 1991).

The National Women's Study (NWS) PTSD Module (Resnick, 1996), as adapted by Dansky, Bryne, and Brady (1999), was used to establish PTSD Criterion A events. The NWS-PTSD module is a structured interview based on the DIS. This module has been frequently used with populations consisting of men and women (e.g., Coffey et al., 2006; Saladin et al., 2003) and has demonstrated good reliability and concurrent validity (Resnick, Kilpatrick, Dansky, Saunders, & Best, 1993).

The Clinician-Administered PTSD Scale (CAPS; Blake et al., 1995) is a structured clinical interview that was used in the current study to diagnose current (past month) PTSD symptoms for all Criterion A events combined. The CAPS is considered the "gold standard" for PTSD assessment (Weathers, Keane, & Davidson, 2001).

The Mini-International Neuropsychiatric Interview (MINI; Sheehan et al., 1998), a commonly used structured diagnostic interview, was used to assess diagnostic criteria for common mood, eating, and anxiety disorders, with the exception of PTSD and SUDs.

Self-report

The Impact of Event Scale-Revised (IES-R; Weiss & Marmar, 1997) is a widely used measure of the three DSM-IV PTSD symptom clusters (intrusion, avoidance, and arousal). The IES-R has demonstrated sound reliability and validity (e.g., Creamer, Bell, & Failla, 2003) and was used in the current study to assess PTSD symptoms specific to the target trauma used in the individualized trauma script.

The Alcohol Dependence Scale (ADS; Skinner & Horn, 1984) measures alcohol-related problems and symptoms, including withdrawal symptoms, tolerance, impaired control over drinking, and alcohol-seeking behavior. It displays solid psychometric properties (e.g., Drake, McHugo, & Biesanz, 1995).

Cue reactivity

The Positive and Negative Affect Scale (PANAS; Watson, Clark, & Tellegen, 1988) is a 20-item self-report scale that measures positive and negative affect. Each item is rated from 1 = very slightly or not at all to 5 = extremely. Items are summed to provide totals ranging from 10 to 50 for positive and negative affect. The PANAS has demonstrated sound psychometric properties (Crawford & Henry, 2004; Watson et al., 1988) and was used in the current study to measure affective responses to trauma and alcohol cue combinations.

Self-reported cue-elicited alcohol craving was assessed with three Likert-type scale items (Kozlowski, Pillitteri, Sweeney, Whitfield, & Graham, 1996). Participants rated three statements: (a) "I crave a drink right now," (b) "I have a desire for a drink right now," and (c) "I want a drink right now," on a 0 to 10 Likert-type scale following each cue-combination trial. Items demonstrated high internal consistency ($\alpha$ = .97 following NN [neutral imagery cue followed by a neutral cue] condition) and were averaged to produce a single craving score.

Salivary flow was used as a physiological measure of cue-elicited craving. This method is described fully by Monti and colleagues (1987). In brief, three preweighed dental cotton rolls were inserted under the tongue and between the inner check and lower gum on each side of the mouth and weighed again following the presentation of each cue combination. Magnitude of salivation was calculated by subtracting the pretrial weight from the posttrial weight.

Procedure
Assessment session

Prospective participants with scores above the cutoffs on the AUDIT and PCL were invited to complete an assessment session. At the assessment session, participants completed an Institutional Review Board (IRB)-approved, documented informed consent procedure. Participants who consented to participate then completed the C-DIS, NWS-PTSD module, CAPS, MINI, IES-R, and ADS to fully assess inclusion and exclusion criteria. Most participants had experienced multiple traumatic events. Therefore, following the recommendations of Weathers and Keane (1999), it was not required that their PTSD symptoms be associated with a single traumatic event on the CAPS. Instead, participants could relate each PTSD symptom to one or more traumatic events that fulfilled Criterion A for PTSD.

If eligible, participants were next asked to describe their worst traumatic event using multiple sensory dimensions (i.e., physical sensations, thoughts, emotions, olfactory sensations, visual details, and avoided activities and events), which provided information for a 60-s audiotaped narrative to be used in their laboratory session. They were then scheduled for a laboratory session within the next week and asked to maintain abstinence from alcohol and illicit drugs for 4 days prior to the laboratory session.

Laboratory session

To control for diurnal variations that may affect cue reactivity, laboratory sessions were scheduled between 1:00 p.m. and 3:00 p.m. On arrival for the laboratory session, participants were asked to remove any gum and were given a urine drug screen (UDS; Instant Technologies, Inc., Norfolk, Virginia) to assess recent drug use. Expired air samples were analyzed (Alco-sensor IV, Intoximeters, Inc., St. Louis, Missouri) to assess current alcohol intoxication. Participants who reported or tested positive for the metabolites of cocaine, opiates, benzodiazapines, amphetamines, methamphetamine, oxycodone, propoxyphene, barbiturates, and 3,4-methylenedioxymethamphetamine (MDMA) were rescheduled. However, participants who tested positive for Tetrahydrocannabinol (THC) and reported marijuana use in the past 30 days, but not the past 4 days, were allowed to participate in the laboratory session due to the long half-life of THC. Participants were then seated in a sound attenuated experimental room and provided detailed instructions of the laboratory procedure.

This laboratory procedure has been described in detail elsewhere (Coffey et al., 2010). Briefly, participants were presented four counterbalanced imagery in vivo cue combinations (i.e., trauma imagery cue followed by an alcohol cue [TA], trauma imagery cue

followed by a neutral cue [TN], neutral imagery cue followed by an alcohol cue [NA], and NN). The trauma imagery cue was a 60-s narrative description of the participant's subjectively rated worst traumatic event, the alcohol cue was their preferred alcoholic beverage, and the neutral cues were a 60-s narrative about changing a light bulb and a bottle of water. Prior to the presentation of these cue combinations, participants were led through a practice trial using an NN combination. Dental cotton rolls were inserted into participants' mouths and they were instructed to close their eyes and listen to the narrative. Following the imagery presentation, the experimenter placed the beverage cue in front of the participant, and they were instructed to open their eyes and continue imagining the scene described in the narrative as vividly as possible for 2 min. Immediately following the 3-min imaginal in vivo cue exposure, the dental cotton rolls were removed and participants were instructed to complete self-report measures (i.e., craving and PANAS). Following the practice trial, this procedure was repeated for the four cue combinations (i.e., TA, TN, NA, NN). Participants provided a final craving rating at the end of the session to assure their safety on dismissal; a clinical psychologist was available to assist any distressed participants.

Analyses

Aim 1: Responses to alcohol and trauma cues

Repeated-measures ANOVA was used to examine differences among the four cue-combination trials on positive and negative affect (i.e., PANAS) and craving (i.e., self-report and salivary response). Statistically significant omnibus Fs were investigated with paired-samples t tests.

Aim 2: Cue-elicited affect and craving

To examine the relationship between cue-elicited alcohol craving and affect, change scores were first calculated to isolate the specific effects caused by the addition of alcohol cues. These change scores were calculated by subtracting affect or craving ratings administered after the neutral trials from the same measure administered after the alcohol trials. For example, two positive affect change scores were calculated for each person, one as the difference between NA and NN trials (NA-NN; representing affective responses to alcohol cues outside a trauma context) and another as the difference between TA and TN trials (TA-TN; representing affective responses to adding alcohol cues to trauma cues). Because the raw PANAS scores range from 10 to 50, these change scores could potentially range from +40 (indicating extreme increase in affect following cue presentation) to −40 (indicating extreme decrease in affect). Two negative affect, self-report craving and salivation change scores were calculated for each person in the same fashion. Four hierarchical multiple regression analyses were used to examine the incremental utility of adding positive affect (Step 2) to negative affect (Step 1) in the prediction of cue-elicited cravings (self-report and salivation).

Aim 3: Ambivalence model response types and cue-elicited craving

To examine how cue-elicited affect maps on to the ambivalence model of cravings, participants were classified into one of four responder types based on their positive and negative affect scores. First, individuals demonstrating an increase in positive affect from neutral to alcohol cue presentation trials (i.e., NA-NN) were designated “high positive”; individuals demonstrating either a decrease or no change in positive affect were classed as “low positive.” Participants were labeled as either “high negative” or “low negative” in the same manner. These labels were used to classify participants into one of the four ambivalence model responder types: “positive” (high positive/low negative), “negative” (low positive/high negative), “ambivalent” (high positive/high negative), and “indifferent” (low positive/low negative).

Repeated-measures ANOVA was used to examine differences between the four responder types and change in craving (self-report and salivary response) associated with adding alcohol cues to neutral cues (i.e., NA-NN). Statistically significant omnibus Fs were investigated with independent-samples t tests. Because group sizes were quite unequal for these analyses, conservative Scheffé tests were applied to limit Type 1 error. To examine whether results differed in the context of trauma cues, categorization and ANOVAs were repeated using affective responses to alcohol in trauma trials (i.e., TA-TN).

Results.
Means, standard deviations, and between-trial effect sizes for negative affect, positive affect, and craving are presented in Table 2; see Figure 1 for graphic presentation. No gender, education, employment, or ethnicity differences were found on PANAS scores or alcohol craving during the cue-reactivity paradigm.

Table 2

Raw Means (SDs) and Between-Trial Effect Sizes (d) of Participants’ (N = 108) Ratings and Salivary Responding Following Each Cue-Combination Trial

Figure 1

Cue-elicited affect and self-report craving means

Aim 1: Responses to Alcohol and Trauma Cues

Negative affect

A significant main effect was found for trial, $F(3, 321) = 172.71$, $p < .001$, $\eta p2 = .62$. Post hoc tests revealed the TA trial elicited significantly more negative affect than the TN ($p = .004$), NA ($p < .001$), and NN trials ($p < .001$). The TN trial elicited significantly greater negative affect than the NA ($p < .001$) and NN trials ($p < .001$). Last, the NA trial elicited significantly greater negative affect than the NN trial ($p < .001$).

Positive affect

A significant main effect was found for trial, $F(3, 321) = 17.30$, $p < .001$, $\eta p2 = .14$. Post hoc analysis of the trial types revealed that the TA and NA trials elicited higher levels of positive affect than the TN and NN trials, all ps < .001. The TA and NA trials elicited equivalent levels of positive affect ($p = .80$), whereas the TN trials elicited significantly greater positive affect than the NN trial ($p = .01$).

To examine whether the TN increased positive affect over NN as a function oof PANAS positive affect scale items potentially tapping anxious arousal ("alert," "attentive," "active"), analyses were repeated after removing these three items. Consistent with this explanation, a significant main effect was found for trial, $F(3, 321) = 19.05$, $p < .001$, $\eta p2 = .15$. Post hoc analysis of the trial types revealed that the NA trials elicited higher levels of positive affect than all other trials, including TA, ps < .01, ds > .26. The TA trial elicited higher levels of positive affect than TN and NN, ps < .01, ds > .41. The TN and NN trials were no longer significantly different, $p = .47$, $d = .09$.

Self-reported craving

Results from the self-report craving measure were similar to the PANAS findings. First, a significant main effect was found for trial, $F(3, 321) = 91.71$, $p < .001$, $\eta p2 = .46$. Post hoc analysis of the trial types revealed that the TA trial produced higher craving than the TN, the NA, and the NN trials (all ps < .001). The TN trial produced higher craving than the NN trial ($p < .001$) but did not differ statistically from the NA trial ($p = .12$). The NA trial produced higher craving ratings than the NN trial ($p < .001$).

Salivation

Analyses of salivation showed congruent effects. A significant main effect was found for trial, $F(3, 318) = 19.31$, $p < .001$, $\eta p2 = .15$. Post hoc analysis of the trial types revealed that the TA trial produced greater salivation than the TN, NA, and the NN trials (all ps < .001). The TN trial produced greater salivary response than the NN trial ($p < .001$) but differed only marginally from the NA trial ($p = .07$). The NA trial produced more salivation than the NN trial ($p < .001$).

Aim 2: Cue-Elicited Affect and Craving

Correlation and multiple regression analyses were conducted to examine the relationship between affective responses to cues and alcohol cue-elicited craving. Increases in self-reported craving associated with adding alcohol cues to neutral cues (NA-NN) correlated with an increase in negative ($r = .34$, $p < .001$) and positive affect ($r = .23$, $p = .02$); increases in craving associated with adding alcohol cues to trauma cues (TA-TN) was also associated with an increase in negative ($r = .25$, $p = .01$) and positive affect ($r = .19$, $p = .049$). Increases in salivation associated with the addition of alcohol cues to neutral cues correlated with increases in

positive affective responses (r = .20, p = .04) but was unrelated to change in negative affect (r = .06, p = .51). Increases in salivation associated with the addition of alcohol cues to trauma cues were unrelated to affective changes of either valence (rs < .09, ps > .45).

Correlations between subjective craving and salivary responsivity to alcohol cues were not significant, either in the context of neutral cues (NA-NN; r = .15, p = .11) or trauma cues (TA-TN; r = .10, p = .31). Within modality, there was only a marginally significant relationship between subjective craving reactivity to alcohol cues in the context of neutral versus trauma cues, r = .18, p = .07. Salivary responses to alcohol were more consistent, with NA-NN and TA-TN reactivity correlating at r = .45, p < .001.

Self-reported craving

The full model accounted for 15% of variance in craving score change from the NN to NA condition, F(2, 105) = 9.25, p < .001. Change in negative affect from NN to NA accounted for 12% of variance at Step 1 of the model, β = .34, p < .001. Change in positive affect from NN to NA accounted for an additional 3% of variance at Step 2 of the model, β = .19, p = .04.

With regard to difference in craving scores between the TA and TN conditions, the full model accounted for 8% of variance in craving score change from the TN to TA condition, F(2, 105) = 4.80, p = .01. Change in negative affect from TN to TA accounted for 6% of variance at Step 1 of the model, β = .25, p < .001. Change in positive affect from TN to TA did not explain any additional variance at Step 2 of the model, β = .15, p = .11.

Salivation

When examining salivary responses to the addition of alcohol cues to neutral cues (i.e., NA-NN), the full model accounted for 4% of variance in salivation change from the NN to NA condition but did not reach statistical significance, F(2, 104) = 2.19, p = .12. Change in negative affect from NN to NA was unrelated to change in salivation at Step 1 of the model, β = .06, p = .51. However, change in positive affect explained significant incremental variance at Step 2, r2 change = .04, β = .19, p = .05.

With regard to difference in salivation scores between the TA and TN conditions, the full model accounted for 0.6% of variance in salivation score change from TN to TA and was not statistically significant, F(2, 105) = 0.292, p = .75. Change in negative and positive affect did not account for variation in salivary responding at either step of the model, βs < .08, ps >.45.

Aim 3: Ambivalence Model Response Types and Cue-Elicited Craving

Means, standard deviations, and between-group effect sizes comparing cueelicited cravings among individuals classified as “positive,” “negative,” “ambivalent,” or “indifferent” affective responders are presented in Table 3.

Table 3

Affective Response Types: Means (SDs) and Between-Group Effect Sizes (d) of Cue-Elicited Change in Craving and Salivary Responding

Self-reported craving

For changes in self-reported craving associated with adding alcohol to neutral cues (NA-NN), a significant main effect was found for affective responder classification, $F(3, 104) = 7.69$, $p < .001$, $\eta p2 = .18$. Post hoc Scheffé tests revealed that "ambivalent" responders reported significantly greater cue-elicited increases in craving than individuals classified as "positive" ($p = .004$) or "indifferent" ($p = .01$) and marginally greater craving than those classified as "negative" responders ($p = .07$). Individuals classified as "positive," "negative," or "indifferent" did not report significantly different cue-elicited changes in craving, $ps > .58$.

When adding alcohol to trauma cues (TA-TN), there were again significant differences between affective responder classifications, $F(3, 104) = 3.08$, $p = .03$, $\eta p2 = .08$. Post hoc Scheffé tests revealed that "ambivalent" responders reported significantly greater cue-elicited increases in craving than individuals classified as "indifferent" ($p = .03$). There were no other statistically significant group differences, $ps > .42$.

Salivation

There were no statistically significant differences among the four responder groups when adding alcohol to neutral cues (NA-NN), $F(3, 104) = 1.11$, $p = .35$, $\eta p2 = .03$, or when adding alcohol to trauma cues (TA-TN), $F(3, 104) = 0.38$, $p = .77$, $\eta p2 = .01$.

Discussion.

Responses to Alcohol and Trauma Cues

The first goal of the present study was to investigate affective responses to alcohol and trauma cues among individuals with PTSD and AD. First, it was hypothesized that the combined trauma image-alcohol cue (TA trials) would increase negative affect and craving more than any other cue condition. Consistent with this hypothesis, trauma cues (TA and TN) elicited the strongest negative affect, with the greatest distress occurring following presentation of trauma and alcohol cues together (TA trials). Together, this evidence indicates a potential additive effect whereby trauma and alcohol cues trigger stronger negative affect in individuals with PTSD-AD than either cue alone. This builds on previous work demonstrating a relationship between cue-elicited craving and increases in guilt, anxiety, and fear in alcohol-dependent individuals (Cooney et al., 1987; Fox et al., 2007). Consistent with the cue-reactivity literature (see Carter & Tiffany, 1999), trauma and alcohol cues also produced moderate to large increases in self-reported craving and salivation. Craving was highest when trauma and alcohol cues were presented together, relative to when either type of cue was presented alongside a neutral counterpart (i.e., TA > TN and NA). This finding is in line with previous work indicating that negative affective states trigger urges to drink (Cannon et al., 1992; Fox et al., 2007; Rubonis et al., 1994). Taken together, the

present results provide support for a model whereby negative affect, exposure to drug/alcohol and trauma cues, and craving mutually encourage one another. This process is consistent with bidirectional conceptualizations of concurrent PTSD-SUD (Brown et al., 1998).

Second, it was hypothesized that alcohol cues would elicit increases in positive affect. This hypothesis was also supported, given that the trauma script-alcohol and neutral script-alcohol conditions (TA and NA) were associated with the greatest levels of positive affect. This effect was most obvious when trials involving alcohol were compared with neutral-neutral cue conditions (i.e., TA and NA > NN; medium effect sizes), but small effects were also observed between the trauma script-alcohol and trauma script-neutral cues (i.e., TA > TN). These findings replicate and extend past research demonstrating cue-elicited positive affect among normative social drinkers (Kambouropoulos & Rock, 2010; Kambouropoulos & Staiger, 2004) within a treatment-seeking PTSD-AD sample. The finding that alcohol cues increase positive affect even in the context of trauma imagery is novel and underscores the reinforcing properties of alcohol-seeking behavior. This may be an important mechanism underlying the maintenance of alcohol use among highly distressed populations.

Trauma cues also elicited more positive affect than neutral script cues (i.e., TN > NN), but this effect disappeared once three PANAS-PA items overlapping with anxious arousal ("alert," "attentive," "active") were removed. Thus, trauma cues do not appear to appreciably increase positive affect. Future work examining affective responses to trauma may wish to account for this measurement effect or supplement with alternative assessment modalities (e.g., facial electromyography responses).

Cue-Elicited Affect and Craving

The second aim of the current study was to examine the relationship between cue-elicited affect and craving. Previous research, as well as available data, indicates that increases in negative affect are associated with increased craving. However, given the paucity of research on positive affect in this population, as well as models highlighting the importance of appetitive responses in craving (e.g., Breiner et al., 1999; Stritzke et al., 2007; Tiffany, 1990), the role of positive affect in craving, above and beyond that of negative affect, was of particular interest. As hypothesized, when examining the addition of alcohol to neutral cues (i.e., NA-NN responses), positive affect explained a significant proportion of variance in self-reported craving and salivation responses, over and above that already accounted for by negative affect. Together, these results underscore the importance of positive and negative affective responses in understanding cue-elicited alcohol craving. In fact, this is the first evidence to our knowledge that cue-elicited positive affect explains incremental variance in alcohol cravings among individuals with comorbid PTSD and AD. This work complements multidimensional models of cravings as well as literature indicating relationships among positive affect, outcome expectancies, and craving in alcohol users without comorbid PTSD (Johnson & Fromme, 1994; Litt et al., 2000).

When examining alcohol cue-elicited craving in the context of trauma cues (i.e., TA-TN responses), negative affect was positively correlated with craving, a finding consistently demonstrated in previous work (Coffey et al., 2002; Coffey et al., 2006; Saladin et al., 2003). However, positive affect did not explain a significant proportion of variance in self-reported craving over and above that

accounted for by negative affect in the presence of trauma cues. Similarly, neither positive nor negative affective responses were associated with cue-elicited salivation in this context. The discrepant findings between the NA-NN and TA-TN findings are intriguing and underscore the importance of examining cravings from a multidimensional perspective. Specifically, the relative importance of positive or negative affect in SUD may vary depending on contextual factors. When triggered by traumatic memories, distress may be the strongest determinant of alcohol use among individuals with PTSD and AD. However, in a neutral environment, feeling distressed or happy may trigger drinking. Treatments for individuals with PTSD-AD may need to address these different pathways to use.

### Ambivalence Model Response Types and Cue-Elicited Craving

Our final study goal was to examine the relationship between cue-elicited affect and cravings as an extension of the ambivalence model of cravings. Partially consistent with the hypotheses, participants classified as “ambivalent” affective responders (i.e., demonstrating increases in positive and negative affect in response to cues) reported significantly greater increases in self-reported craving than individuals classified as “positive” (increases in positive but not negative affect) or “indifferent” (no increases in either affect valence) and marginally greater increases in craving than “negative” responders (increases in negative but not positive affect). There were no group differences in salivary responses.

These results are generally consistent with models of cravings that point to the role of conflicting positive and negative responses in the subjective experience of craving (Stritzke et al., 2007; Tiffany, 1990). It is possible, however, that positive and negative affect in response to alcohol and trauma cues do not reflect pure approach and avoidance motivations, given that individuals classified as “positive” did not evidence relatively stronger cravings. Affective ambivalence may instead relate to craving and subsequent motivation to drink as a function of the greater affective intensity, more generally, as present in the “ambivalent” class of responders. Alternatively, as initial euphoric responses to drug cues have been shown to give way to increasing levels of irritability and frustration over time (O’Brien, Ehrman, & Ternes, 1986; see Stasiewicz & Maisto, 1993, for discussion), craving could also be higher among “ambivalent” responders as a function of frustration. Although this is one of the first studies to examine affective ambivalence in substance use cravings, it builds on some preliminary work examining the role of affective ambivalence in desire for food and sex (e.g., Hormes & Rozin, 2011; Peterson & Janssen, 2007).

### Limitations

Several limitations of the current work are worth noting. First, no diagnostic control group was used. As such, although results shed light on affective responses to drug and alcohol cues, it cannot be concluded that these effects are unique to individuals with PTSD and AD. Individuals with comorbid SUD and depression, for example, may display similar response patterns. Comparison with nonclinical and alternative clinical groups would clarify the specificity of results. Second, the sample comprises individuals from a residential SUD treatment facility. People who are seeking treatment may experience different affective responses to cues and may arguably be less ambivalent about their use than substance users in the general population. In addition, although participants

completed the laboratory session less than a week after being admitted into the treatment program, they may already have learned new coping strategies that influenced their affective responses to cues. Thus, results may not generalize to nontreatment seekers. Third, although the present study examines positive and negative affect as an extension of the ambivalence model of cravings, several constructs that are important to these models (e.g., expectations about the effects of use, motivation, cognition, and actual drinking behavior; Breiner et al., 1999) were not assessed. Future laboratory work examining affective ambivalence would benefit from the inclusion of explicit self-report and behavioral measures of approach and avoidance motivation, such as the Approach and Avoidance of Alcohol Questionnaire (McEvoy et al., 2004). Finally, self-reported craving and affective ratings were collected within the same assessment period of the laboratory study, and therefore, it is unclear whether affect promoted craving, whether the reverse was true, or whether bidirectional influences existed. Future work using diverse study designs could help determine the interrelations among these variables.

Conclusions and Future Directions.

Understanding affective responses to alcohol and trauma cues will be an important step toward advancing conceptual models of concurrent PTSD-AD, particularly as it relates to craving. The present study demonstrates that exposure to alcohol cues among treatment-seeking individuals with comorbid PTSD and AD increases positive and negative affect in addition to self-reported craving and salivation. Negative affective responses to alcohol cues correlated significantly with self-reported craving in the context of neutral and trauma cues. A novel finding was that positive affective responses explained significant incremental variance in self-reported craving and salivation when cues were accompanied by neutral, but not trauma imagery. Consistent with this, participants classified as “ambivalent” affective responders reported the greatest cue-elicited increases in cravings, particularly in the context of neutral imagery cues. Overall, findings highlight the importance of independently assessing positive and negative affective responses to cues in understanding the complexity of cravings in individuals with comorbid PTSD and AD. Results are consistent with multidimensional conceptualizations of cravings and encourage greater consideration of the rewarding elements of drinking and the importance of resolving ambivalence in the treatment of SUDs (Breiner et al., 1999; Miller & Rollnick, 1991). Increasing awareness of affective responses to cues and developing targeted strategies to cope with cravings in different emotional contexts may be a valuable area for further research and intervention.

Acknowledgments.

The authors wish to thank M. Trost Friedler, Jackie Lampley, and the staff and patients of Harbor House Recovery Center for their cooperation on this study. They also wish to thank Robert C. Schlauch for his helpful comments on an earlier draft of this manuscript.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported, in part, by National Institute on Alcohol Abuse and Alcoholism Grant R01AA016816 (PI: Coffey).

BACK TO TOP

## Fear of memories: the nature of panic in posttraumatic stress disorder.

European Journal of Psychotraumatology
Joscelyne A, McLean S, Drobny J, Bryant RA.
November 2012

Abstract

BACKGROUND

Although there is increasing evidence that panic attacks are common in posttraumatic stress disorder (PTSD), little is known if posttraumatic panic is comparable to panic attacks observed in panic disorder (PD).

OBJECTIVE

This study examined the cognitive responses to panic attacks in participants with PD and PTSD.

METHOD

Participants with PD (n=22) and PTSD (n=18) were assessed on the Anxiety Disorder Interview Schedule for DSM-IV and subsequently administered the Agoraphobic Cognitions Questionnaire and a measure of fears related to trauma memories.

RESULTS

Although participants did not differ in terms of catastrophic appraisals about somatic sensations, PTSD participants were more likely to experience fears about trauma memories and being harmed by trauma again during their panic attacks than PD participants.

CONCLUSIONS

These findings suggest that although PTSD participants fear somatic outcomes during panic attacks, their panic attacks are distinguished by a marked fear of trauma memories.

Abstract

There is increasing evidence that panic attacks are common in people with posttraumatic stress disorder (PTSD). An analysis of the US National Comorbidity Survey found that 35% of people with PTSD had panic attacks in the past year, and this was linked to greater disability and comorbidity (Cougle, Feldner, Keough, Hawkins & Fitch, 2010). Many trauma survivors report experiencing panic attacks during the traumatic event; 90% of rape victims (Resnick, Falsetti, Kilpatrick & Foy, 1994) and 53% of motor vehicle and non-sexual assault survivors (Bryant & Panasetis, 2001) report at least four panic symptoms during the trauma. People with acute stress disorder (ASD) are more likely to experience panic attacks at the time of the trauma than those without ASD (Bryant & Panasetis, 2001). Furthermore, people with ASD also report more posttraumatic panic when compared to trauma survivors without ASD (Nixon & Bryant, 2003). Falsetti and Resnick (1997) found that 69% of treatment-seeking trauma survivors had experienced at least one panic attack in the 2 weeks prior to presenting for treatment. The importance of panic in the trajectory of PTSD responses is also highlighted by findings that initial dissociation mediates the relationship between peritraumatic panic and subsequent PTSD (Bryant et al., 2011), and peritraumatic panic predicts long-term mental health outcomes (Boscarino & Adams, 2009). Despite the relevance of panic attacks in PTSD, little is known about the nature of posttraumatic panic.

The intersection between PTSD and panic disorder (PD) is highlighted in recent years by fear circuitry models, which posit common etiologies and mechanisms across fear-based anxiety conditions, including PTSD and PD (Andrews, 2009). It is proposed that fear circuitry disorders share fear-conditioning processes at their point of origin such that otherwise benign stimuli are paired with an aversive experience; subsequent exposure to the conditioned stimuli signals threat and results in anxiety (Lanius, Frewen, Vermetten & Yehuda, 2010; Milad, Rauch, Pitman & Quirk, 2006). In the context of PTSD, this would involve reminders of the threat, whereas in the context of PD, it would require reminders of physical fears, such as choking, having a heart attack, or dying. This accords with models that posit that the arousal and panic experienced at the time of a traumatic experience become part of the conditioned stimuli, and thereafter somatic cues can trigger re-experiencing symptoms (Hinton, Hofmann, Pitman, Pollack & Barlow, 2008). Consistent with animal and human fear-conditioning research (Rauch & Drevets, 2009), fear circuitry disorders are characterized by excessive amygdala reactivity and impaired regulation of that response by the medial prefrontal cortex (Shin & Liberzon, 2010).

The major cognitive model of PD postulates that people catastrophically misinterpret somatic sensations to the extent that they fear that benign sensations are perceived as signals of impending death or severe illness (Clark, 1986, 1996). For example, sensations such as mild chest pain and dizziness may be viewed as being indicative of an impending heart attack. Supporting this model is the evidence that PD patients are more likely to interpret situations containing ambiguous internal stimuli as threatening (Clark et al., 1988; McNally & Foa, 1987; for a review, see McNally, 1994). Models of posttraumatic panic posit that panic that occurs at the time of trauma contributes to strong fear conditioning, and the somatic cues associated with the panic become associated with many other cues related to the traumatic experience (Falsetti, Resnick, Dansky, Lydiard & Kilpatrick, 1995; Jones & Barlow, 1990). These models propose that subsequent internal (e.g., emotions, physiological arousal, and cognitions) and external (e.g., places, objects, and smells) triggers elicit subsequent panic attacks, which in turn trigger trauma-related associations. This proposal is supported by evidence that 84% of a sample of trauma patients experiencing panic attacks reported that trauma reminders cued their panic attacks (Falsetti & Resnick, 1997).

Cognitive models are capable of explaining both PTSD and PD. The emphasis on cognitive responses in models of PD converges with cognitive PTSD models, which also propose that traumatic experiences can lead to catastrophic interpretations about the experience, the potential of future harm, and how one manages the effects of the traumatic experience (Ehlers & Clark, 2000). This is supported by evidence that maladaptive appraisals after trauma are predictive of subsequent PTSD (Dunmore, Clark & Ehlers, 1999; Ehlers, Mayou & Bryant, 1998; Warda & Bryant, 1998). Fear network models posit that mental representations of the feared content are highly connected and readily activated by cues that are related to the feared event; when activated, these representations can involve catastrophic appraisals about the feared event, thereby exacerbating the fear (Foa & Kozak, 1986). These representations can apply to traumatic or somatic representations, thereby being able to explain the cognitive responses of both PTSD and PD. It has been suggested that PTSD is characterized by a more widely activated fear network than other anxiety disorders as a result of the severity of the threat (Foa, Steketee & Rothbaum, 1989). Consistent with this proposal, trauma survivors display catastrophic appraisals about traumatic, somatic, and social events (Smith & Bryant, 2000).

An outstanding issue concerns the cognitive responses to posttraumatic panic attacks. Although PD models posit that the major mechanism underpinning the disorder is fear of aversive consequences of somatic events, it is possible that different fears underpin panic attacks in PTSD. Specifically, PTSD models emphasize that panic attacks cue conditioned responses that developed at the time of the traumatic experience and these attacks should accordingly trigger trauma memories. This hypothesis has indirect support from evidence indicating that inducing arousal in trauma survivors elicits trauma memories, as well as more flashback phenomena, in trauma victims with PTSD or ASD (Bremmer et al., 1997; Nixon & Bryant, 2005). On the basis of this hypothesis, we predicted that whereas panic attacks in the context of PD would be predominantly associated with fear of aversive outcomes from somatic perceptions, we expected that panic attacks in the context of PTSD would be associated with fear of trauma memories.

## Method

### Participants

The PD sample comprised 22 consecutively assessed participants (9 male and 13 female) of mean age 39.55 years (SD=13.08), who were seeking treatment at the Anxiety Treatment and Research Unit at Cumberland Hospital, Sydney; 21 participants met criteria for PD with agoraphobia and 1 had PD without agoraphobia. The PTSD sample comprised 18 consecutively assessed participants (8 male and 10 female) of mean age 42.22 years (SD=10.24), who were seeking treatment at the Traumatic Stress Clinic at Westmead Hospital, Sydney. Participants presented after motor vehicle accidents (n=11) or non-sexual assault (n=7). Inclusion criteria were (1) met criteria for PD or PTSD, (2) proficiency in English, (3) aged between 16 and 65 years, and (4) no diagnosis of organic mental disorder or psychosis. Patients in the PTSD group met DSM-IV diagnostic criteria for PTSD, did not meet criteria for PD, and had experienced at least one panic attack following their trauma. Patients in the PD group met DSM-IV criteria for either PD, with or without agoraphobia, but failed to meet criteria for a PTSD diagnosis. In terms of comorbidity of the PTSD participants, six participants had major depressive disorder and four had substance abuse. In terms of comorbidity of the PD participants, six

participants were diagnosed with generalized anxiety disorder, five with major depressive disorder, two with obsessive–compulsive disorder, and one with social phobia. Four PD participants also reported past trauma, with two experiencing childhood abuse and two involved in a motor vehicle accident, but none reported PTSD symptoms related to these events.

## Measures

Diagnosis of PD and PTSD was ascertained using the Anxiety Disorder Interview Schedule for DSM-IV (ADIS-IV; Brown, Di Nardo & Barlow, 1994), which is a clinician-administered structured diagnostic interview following DSM-IV criteria. This schedule was also used to determine the presence of any comorbid anxiety disorders. The test–retest reliability of the ADIS-R (the predecessor of the ADIS-IV) ranges from 0.57 to 0.82 (di Nardo, Moras, Barlow, Rapee & Brown, 1993). PTSD severity was measured using the Clinician-Administered PTSD Scale (CAPS; Blake et al., 1995). The CAPS is a structured clinical interview that indexes the 17 symptoms described by the DSM-IV PTSD criteria.

Participants in both groups were also administered two questionnaires to index their panic-related cognitions. The Agoraphobic Cognitions Questionnaire (ACQ; Chambless, Caputo, Bright & Gallagher, 1984) consists of 14 cognitive statements that represent cognitive misappraisals of somatic symptoms. To assess trauma-related panic cognitions, a 12-item Traumatic Panic Cognitions Scale (TPCS) was developed. Items for this measure were based on proposals from experienced PTSD clinicians concerning the common fears that PTSD patients report during their panic attacks. This questionnaire comprised items that index the extent to which respondents may worry about traumatic memories or traumatic events occurring during a panic attack, such as, “Memories of the past are hurting me”, “I am reliving a terrible event”, and “I will never escape my memories”. The TPCS showed strong internal consistency, with a Cronbach's-α of 0.94. Both these questionnaires specifically indexed cognitions that occur during a panic attack and each utilized a five-point Likert scale (1=never, 2=hardly ever, 3=sometimes, 4=often, 5=always).

## Procedure

Following informed written consent, participants were administered the ADIS-IV by clinical psychologists. Two weeks after the clinical interview, participants were administered the ACQ and TPCS in a random order of presentation.

# Results

## Participant characteristics

An independent sample t-test indicated no difference between the PTSD and PD groups in terms of their age ($t(38)=-0.708$, ns). There was also no difference in time since the onset of panic attacks between the PTSD (M=68.50 months, SD=49.56) and PD (M=60.41 months, SD=56.56) groups ($t(38)=0.33$, ns).

Somatic reactions

Participants in the PTSD group (M=10.05, SD=2.95) and the PD group (M=9.00, SD=1.14) did not differ in terms of their total number of ADIS-IV assessed somatic panic symptoms (t(38)=1.42, ns). Table 1 presents the percentage of participants in both groups who reported each somatic symptom. Chi-square analyses that adopted a Bonferroni-adjusted α of p<.005 indicated that PD participants were more significantly likely to report numbing/tingling than PTSD participants ($\chi 2(N=40)=7.78$, p<.005).

Cognitive reactions

The mean rating of each ACQ item is presented in Table 2. The PD (M=36.34, SD=11.83) and PTSD (M=32.50, SD=8.38) groups did not differ in terms of total ACQ scores, t(38)=1.25, ns. Multiple comparisons that adopted a Bonferroni-adjusted α of p<.005 indicated that PD participants reported a fear of passing out [t(38)=3.14, p<.005] and a fear of going blind [t(38)=3.20, p<.005] more often than PTSD participants.

Table 3 presents the mean rating of TPCS items. The PTSD participants (M=40.72, SD=10.06) scored higher than PD participants (M=26.78, SD=14.49) on TPCS items (t(38)=3.45, p<.001). PTSD participants were significantly more likely than PD participants to report that during a panic attack they relived a terrible event [t(38)=−2.94, p<.005], that memories of their past were hurting them [t(38)==−3.91, p<.001], that something was going to hurt them again [t(38)=−3.55, p<.001], that their memories were driving them crazy [t(38)=−3.11, p<.005], that they feel trapped in the past [t(38)=−3.06, p<.005], and that they cannot block out terrible memories [t(38)=−3.03, p<.005].

Discussion

PD and PTSD participants reported comparable somatic symptoms during their panic attacks. This finding accords with previous evidence that PTSD and PD participants report comparable somatic experiences during panic attacks (Falsetti & Resnick, 1997). Despite this similarity in somatic response, PTSD participants reported marked fear of trauma memories more than PD participants. PTSD participants tended to report that during their panic attacks, they experience traumatic intrusions, fears that they would be harmed again, and an inability to control trauma memories. This pattern suggests that the cognitive profile of panic attacks in PTSD may be qualitatively different from the panic attacks typically observed in PD. Whereas PD is characterized by fear of aversive consequences of somatic events, PTSD participants predominantly fear the consequences of trauma memories. This observation is consistent with models of posttraumatic panic that posit that panic reactions at the time of trauma are conditioned with the traumatic event, and that subsequent panic reactions will trigger associated memories (Falsetti et al., 1995; Jones & Barlow, 1990).

Contrary to expectation, PTSD participants reported comparable catastrophic interpretations about somatic sensations as PD participants. This pattern suggests that panic attacks in PTSD participants also involve maladaptive cognitive responses about somatic events. Previous work has demonstrated that people with posttraumatic stress exaggerate the probability of aversive events occurring and the aversive consequences of these events (Warda & Bryant, 1998). Importantly, this pattern extends to exaggerating

aversive outcomes from somatic events, as well as physical harm (Smith & Bryant, 2000). This pattern is consistent with evidence that people with ASD have higher anxiety sensitivity scores than those without ASD (Bryant & Panasetis, 2001), and that anxiety sensitivity scores are strongly predictive of posttraumatic panic (Nixon & Bryant, 2003). Taken together, these findings suggest that the fear conditioning that occurs during the traumatic event encompasses the aversive outcomes of somatic events, and in this sense the panic attacks in PTSD participants share the fear of somatic catastrophes that are observed in PD panic attacks. The panic attacks in PTSD participants, however, are additionally characterized by the fear of trauma memories.

These findings may have implications for clinical management of posttraumatic panic. PD is traditionally treated using interoceptive exposure that allows the patient to learn that the somatic sensations do not result in the feared outcome (Barlow & Craske, 1988). More recent treatment protocols have been developed to specifically treat comorbid PTSD and PD, which have commenced with interoceptive exposure prior to treating traumatic stress by trauma-focused exposure (Falsetti, Resnick, Davis & Gallagher, 2001). This approach presumes that the panic should be managed via panic management strategies prior to addressing the trauma symptoms. In contrast to this approach, the current findings suggest that effective management of panic in PTSD may occur after imaginal exposure that involves prolonged exposure to the trauma memories, and may not directly require interoceptive exposure. It is worth noting that whereas prolonged exposure is efficacious in treating PTSD symptoms (Harvey, Bryant & Tarrier, 2003), treatment studies to date have not indexed the influence of trauma-focused exposure therapy on posttraumatic panic attacks. Interestingly, recent treatment studies of PD have found that panic-focused cognitive behavior therapy is beneficial in reducing comorbid anxiety conditions; however, these studies have not included PTSD patients (Craske et al., 2007; Tsao, Mystkowski, Zucker & Craske, 2005). We note that whereas our PTSD participants suffered panic attacks, they did not meet the criteria for PD; patients with comorbid PTSD and PD may require interventions that do involve interoceptive exposure because of the significant ongoing fear of somatic outcomes. The extent to which prolonged exposure of trauma memories will reduce posttraumatic panic attacks, and PD, remains to be tested by randomized controlled trials.

This study's conclusions are qualified by the small sample size; future studies that employ larger samples may provide a more robust test of differences in cognitions, and also allow analysis of cognitions about panic following different types of trauma. We note that we relied on a measure of posttraumatic panic cognitions that has not been validated. Development of a measure of these cognitions that is subsequently validated and used in samples of PTSD and PD patients would strengthen the current findings. These limitations notwithstanding, the current findings suggest that posttraumatic panic needs to be understood as a different experience from panic attacks in PD. The fear of trauma memories experienced during posttraumatic panic suggests that these memories may need to be the focus of therapy intervention. The extent to which prolonged exposure of trauma memories resolves posttraumatic panic remains to be tested.

Acknowledgements

This research was funded by an NHMRC Program grant.

Conflicts of Interest

No authors are reporting a conflict on interest. This research was not funded by an Australian National and Health Medical Research Council grant.

BACK TO TOP

## A systematic review of the relationship between severe maternal morbidity and post-traumatic stress disorder.

BMC Pregancy Childbirth
Furuta M, Sandall J, Bick D.
10 Nov 2012

Abstract

ABSTRACT

BACKGROUND

The incidence of severe maternal morbidity is increasing in high-income countries as a consequence, in part, of increased obstetric intervention and increasingly complex medical needs of women who become pregnant. Access to emergency obstetric care means that for the majority of women in these countries, an experience of severe maternal morbidity is unlikely to result in loss of life. However, little is known about the subsequent impact on postnatal psychological health resulting in an evidence gap to support provision of appropriate care for these women. There has recently been increasing recognition that childbirth can be a cause of post-traumatic stress disorder (PTSD). The combination of experiencing a life-threatening complication and its management may culminate in psychological trauma. This systematic review examined the association between women's experience of severe maternal morbidity during labour, at the time of giving birth or within the first week following birth, and PTSD and its symptoms.

METHODS

Relevant literature was identified through multiple databases, including MEDLINE, PsycINFO, EMBASE, CINAHL, British Nursing Index, Web of Science, Cochrane library and the British Library, using predetermined search strategies. The search terms included "post-traumatic stress disorder", "PTSD", "stress disorders, post-traumatic", "maternal morbidity", "pregnancy complications" "puerperal disorders", "obstetric labo(u)r complication", "postpartum h(a)emorrhage", "eclampsia". Studies identified were categorised according to pre-defined inclusion and exclusion criteria. The quality of included studies was assessed using the relevant CASP appraisal tools.

RESULTS

Eleven primary studies met review criteria. Evidence of a relationship between severe maternal morbidity and PTSD/PTSD symptoms was inconsistent and findings varied between studies. Nevertheless, there is some evidence that severe pre-eclampsia is a risk factor for PTSD and its symptoms, an association possibly mediated by other factors such as fetal/neonatal condition.

CONCLUSIONS

Despite the absence of robust evidence regarding the relationship between severe maternal morbidity and PTSD/PTSD symptoms, it is crucially important that clinicians and policy makers are aware of a potential higher risk of PTSD among women who experience severe morbidity. Further studies are now needed to confirm this risk as well as to understand underlying mechanisms in order to minimise the longer term psychiatric impact of severe maternal morbidity.

BACK TO TOP

## Mental Health of Asian American and Pacific Islander Military Veterans: Brief Review of an Understudied Group

Military Medicine
Tsai, Jack; Kong, Grace
November 2012

Abstract

The mental health of American military soldiers and veterans is of widespread concern; yet, there has been no prior review of studies on Asian Americans and Pacific Islanders (AAPIs) veterans. This article provides a brief, but comprehensive review of the mental health of AAPI veterans. An exhaustive literature search was conducted using the major medical and mental health literature databases. Of 13 identified articles, nine were empirical studies on either post-traumatic stress disorder among AAPI Vietnam veterans or health functioning of AAPI veterans based on national veteran surveys. Findings from these studies showed that some AAPI veterans who served during the Vietnam War encountered racism from fellow soldiers and race-related stressors were associated with more severe post-traumatic stress disorder symptoms. As a group, AAPI veterans were found to be physically healthier than other veterans, but reported poorer mental health and were less likely to use mental health services. However, these findings were limited by the paucity of studies on AAPI veterans and suggest a need for more research on this subpopulation.

## Introduction

The mental health and well-being of American military veterans is of wide public concern, but significant racial differences in mental illness and mental health care utilization have been found among veterans of different racial/ethnic groups.1–3 Reviews of studies related to the mental health of African American,4 Hispanic American,5 and Native American/Alaskan Native veterans6 have been conducted, but there has been no prior review of the mental health of Asian American and Pacific Islander (AAPI) veterans.

AAPIs are a diverse racial group consisting of over 60 different racial/ethnic groups that speak more than 100 different languages. Collectively, AAPIs currently constitute 5% of the U.S. population and have been the fastest growing minority group for over 2 decades.7 The population of AAPIs is projected to grow by more than 200% by the year 2050.8

Among AAPIs, 1.3% are military veterans.9 Although AAPIs have served in the military since the early 1800s, there has been little documentation of their contributions until the Civil War, followed by increased attention during World War II.10 Despite more than nearly 2 centuries of military service, little is known about AAPI veterans and their mental health.

AAPIs in the general population are often stereotyped as a “model minority,” and are perceived as a racial group with few social and psychological problems.11 Yet, studies have consistently found that AAPIs experience a range of mental health problems, use mental health services at a lower rate than the general population, and face cultural and institutional barriers to accessing mental health services.11–14 Moreover, some culturally adapted mental health interventions for AAPIs have been found to be more effective than general interventions.15

Studies have shown that veterans compared to the general population are at greater risk for a variety of mental health problems, including substance abuse, post-traumatic stress disorder (PTSD), homelessness, and mortality from suicide16–19 thus making them a potentially vulnerable group. However, the risks of mental health problems among AAPI veterans relative to other veterans or AAPIs in the general population have not been well-documented. Although studies on the mental health of AAPIs have been increasing,20 studies on AAPI veterans have not increased in proportion to the general AAPI literature or the veteran literature.

In this article, we conduct a brief, but comprehensive review of studies related to the mental health of AAPI veterans to summarize current knowledge and provide future directions for research on this understudied group.

## Methods

A comprehensive literature search was conducted from September 2011 to January 2012 in the MEDLINE, PubMed, PsycINFO, and Google Scholar databases using various combinations of keywords, including “Asian American veterans,” “Pacific Islander veterans,” “Hawaiian veterans,” and “Asian veterans” coupled with keywords like “mental health,” “psychiatric,” “substance abuse,” and “substance dependency” with no specified limit on date of publication. The bibliographies of articles found were also searched.

## Results

A total of 13 articles related to the mental health of AAPI veterans were found. Only nine articles presented empirical data, the other four presented case studies and conceptual frameworks. No experimental or longitudinal studies were found. The design and topic of the empirical studies were categorized as examination of PTSD among AAPI Vietnam veterans or analysis of national veteran surveys. Before the empirical studies in each of these categories are described, the nonempirical studies that were found will be briefly summarized.

### PTSD Among AAPI Vietnam Veterans

The four nonempirical articles found represent some of the earliest published work on the mental health of AAPI veterans. This work was focused specifically on AAPI veterans who served during the Vietnam War era, which is estimated to be 36,400 AAPI soldiers.21 The four nonempirical articles presented case studies and anecdotes describing the difficulties AAPI soldiers in Vietnam faced in having to fight Vietcong enemies who physically resembled them and at the same time, deal with discrimination they faced from their own American comrades for looking like the enemy.22,23 AAPI soldiers who served during the Vietnam War era were subject to racist behaviors and attitudes from both fellow Americans and enemy Vietnamese, which became a “painful and untenable position”24 that often threatened their physical and mental well-being.

Examination of PTSD Among AAPI Vietnam Veterans

Four empirical studies have specifically examined PTSD among AAPI veterans who served in the military during the Vietnam War era. Two of these studies examined two different samples of AAPI Vietnam veterans in Hawaii25,26 and the two other studies were conducted by Looet al27,28 who examined the relation between race-related stress and PTSD.

The first of the two Hawaii studies may be the first empirical study of AAPI veterans, which found that among a small sample of 44 AAPI Vietnam veterans, 57% reported they felt very similar to the Vietnamese and 53% reported they were mistaken for Vietnamese during the Vietnam War.25 This led to discrimination from commanding officers and, at times, life threatening and potentially traumatic experiences. The overall rate of PTSD in this sample (15–18%) was similar to those found in other Vietnam veterans, but Native Hawaiians and Pacific Islanders were at higher risk, with rates of 29 and 40%, respectively. The second Hawaii study26 used a larger sample of 202 AAPI Vietnam veterans and found results consistent with the first study; Native Hawaiian veterans had higher rates of partial and full PTSD than White veterans, whereas Japanese American veterans had lower rates. The strengths of both of these studies were that AAPI subgroups were analyzed and semistructured interviews were used as well as established PTSD measures. However, both of these studies were limited by small samples of AAPI veterans and the generalizability of their findings is unknown.

The two other empirical studies on Asian American Vietnam veterans were conducted by Loo et al, who collected and analyzed data on a sample of 300 AAPI Vietnam veterans to develop a Race-Related Stressor Scale (RRSS)27 and showed that race-related stressors contribute uniquely to PTSD symptoms and general psychiatric stress.28 They found that 77% of the sample reported exposure to at least one negative race-related event in the military during the Vietnam War and many of these events met criteria as a traumatic event for PTSD. These studies demonstrated, for the first time, the potential impact of negative race-related experiences of AAPI veterans on their psychopathology. A relatively large sample was used, appropriate measures were employed, and the methodology in these studies was generally sound. However, one limitation was the use of snowball sampling, which may have resulted in a sample that was not representative of all AAPI Vietnam veterans.

Analysis of National Veteran Surveys

In addition to studies on PTSD among AAPI Vietnam veterans, five empirical studies have examined cross-sectional national survey data on veterans of different racial/ethnic groups, including AAPI veterans. Although none have specifically examined PTSD rates, they have attempted to describe the characteristics of AAPI veterans relative to the rest of the veteran population. The sample sizes of AAPI veterans in these studies ranged from 145 to 2,044 AAPI veterans.

The first two of these studies29,30 analyzed data from the 2001 Veteran Identity Program Survey. The authors found that AAPI veterans reported the longest average length of military service and had one of the highest ratings of veteran identity compared to other racial/ethnic veteran groups. AAPI veterans were also more likely to be working, have higher incomes, be more highly educated, and have health insurance than other veterans. Despite this, AAPI veterans reported poorer mental health than African American veterans and lower use of Veterans Affairs (VA) health services than all other racial/ethnic veteran groups. These two studies describe some important but general differences in sociodemographic characteristics and mental health among AAPI veterans compared to other veterans using a national sample. However, detailed information on mental health was not collected and comparisons between AAPI veteran subgroups were not conducted.

The third study examined homelessness among AAPI veterans using administrative data over 4 years on users of VA homeless programs.31 Results showed that AAPI veterans were at lower risk of homelessness than other veterans, which may be attributable to their lower risk of alcohol abuse. However, homeless AAPI veterans were also found to have higher rates of psychotic illnesses. Nonetheless, the lower rate of alcohol abuse found among AAPI veterans was confirmed by the fourth empirical study,32 which analyzed data from the national VA Survey of the Health Experiences of Patients in 2005 and found that AAPI veterans were less likely to report using alcohol than other veterans. Although both of these studies have strengths in their large sample sizes, no formal diagnostic assessments or validated measures were used.

The fifth empirical study33 analyzed national veteran survey data using the 2001 National Survey of Veterans. The authors found that 20% of AAPI veterans used the VA, which was a rate similar to other veterans. Similar to the findings from the 2001 Veteran Identity Program Survey,29,30 AAPI veterans were physically healthier with higher incomes and more likely to receive VA benefits. However, again, AAPI veterans reported poorer mental health and were less likely to use VA and non-VA mental health services. Furthermore, the income distribution among AAPI veterans was bimodal, suggesting there are subgroups of AAPI veterans who have low income levels although data was not collected on specific AAPI veteran subgroups to confirm this. Similar to studies based on the 2001 Veteran Identity Program Survey, the weaknesses of these studies based on the 2001 National Survey of Veterans was that no detailed health information was collected and validated measures were not used. The strengths were that the findings are consistent with previous studies and a national sample of veterans was used.

## Discussion

To our knowledge, this is the first review of the literature on the mental health of AAPI veterans. Only nine empirical articles could be found, suggesting the mental health of AAPI veterans have largely been understudied by researchers and may reflect difficulties in

recruiting AAPI veterans for studies. Nonetheless, the existing empirical studies found on AAPI veterans indicate that AAPI veterans are a group that may need more clinical and research attention.

First, national surveys of the veteran population suggest AAPI veterans are physically healthier and more functional in terms of employment and income than other veterans. However, AAPI veterans report poorer mental health and are less likely to use mental health services both at the VA and at other mental health facilities than other veterans, suggesting AAPI veterans underutilize mental health services. This finding extends studies on AAPIs in the general population that have found lower rates of mental health service utilization among AAPIs compared to other race/ethnic groups12,34,35 and studies that have found AAPIs who do seek mental health care present with more severe, chronic psychopathology than other patients.36 Research on the general AAPI population has pointed to several factors to explain this, including cultural stigma about mental illness, somatization of mental illness, acculturation level, and perceived discrimination.11,37–39

It is unknown whether these factors are applicable to AAPI veterans as they may be different from other AAPIs in their access to mental health services (e.g., VA health services), level of assimilation to American culture, attitudes about mental health, and how they are treated by mental health providers. The need for attention to the mental health of AAPI military personnel has recently been underscored by preliminary results from the largest study of mental health risk conducted in the military called the Army Study to Assess Risk and Resilience in Servicemembers (Army STARRS), which found that AAPI soldiers have higher suicide rates than all other races/ethnicities, not only during deployment but also among those who have never been deployed.40 Although exact rates of suicide among AAPI soldiers have not yet been reported from the Army STARRS project, this pattern of suicides sharply contrasts with findings among the general AAPI population where rates of suicide are much lower than those of other races/ethnicities (AAPI men have the lowest rate compared to other men; AAPI women have second lowest rate compared to other women).41–43 This reverse difference in suicide rates between AAPI military personnel and AAPIs in the general population suggests the importance of focusing on the mental health of AAPIs serving the military.

Second, there was notable heterogeneity in PTSD found among veterans of different AAPI racial/ethnic subgroups. For example, Native Hawaiian and Pacific Islanders tend to report higher rates of PTSD than other Asian American veteran groups and White veterans. But, it is unclear exactly which AAPI racial/ethnic subgroups are particularly at-risk because there have been no comparisons using a wide range of AAPI racial/ethnic subgroups. It is also unclear whether these differences are specific to PTSD or also present in other mental health disorders. Nonetheless, these findings illustrate the heterogeneity that exists between different AAPI race/ethnic subgroups that may be masked by grouping them together under the broad category of AAPI veterans or in the "Other" race category. Studies on AAPI veterans should, when possible, analyze subgroup race/ethnic differences before grouping them together as one group, as has been suggested in studies on general samples of AAPI veterans.44–46

Third, empirical studies have found that AAPI veterans who served during the Vietnam War faced discrimination and racism from their fellow soldiers, which were associated with more severe PTSD symptoms and greater psychiatric distress.25,27,28 The possible effect of race-related stressors on PTSD is an important one, but research in this area has not continued with AAPI soldiers after the Vietnam War era. It is unclear the extent to which AAPI soldiers currently experience race-related stressors and their effects on mental health, although recent news reports indicate there may be race-related problems as two AAPI soldiers serving in Afghanistan committed suicide last year after reportedly being hazed and racially denigrated by fellow American soldiers.47,48

As should be evident in our review, there has been a limited depth and breadth of studies on AAPI veterans. Our review points to several areas for future study that may substantially contribute to the literature, which will now be discussed.

Directions for Future Research

No specific studies on AAPI veterans who served after the Vietnam War could be found. Research is needed on AAPI veterans who served during Operation Desert Storm and the more recent Operations Iraqi Freedom/Operation Enduring Freedom/Operation New Dawn in Iraq and Afghanistan. Longitudinal studies on rates of mental health problems and their trajectories among different veteran race groups would help gauge the effect of race/ethnicity. Additionally, collection of detailed health information of national samples of AAPI veterans beyond the basic information that can be garnered from broad national surveys would be invaluable. Particular attention should be given to rates and factors related to suicide.

A large gap in the literature exists in differences and similarities between AAPI veterans and AAPIs in the general population. Defining these differences and similarities is important in determining what studies on AAPIs in the general population extend to those in the military. For example, there have been various studies on barriers to seeking and accessing mental health services for AAPIs in the general population, but the question remains whether these findings are applicable to the AAPI veteran population.

One similarity between AAPIs in the general population and AAPI veterans appears to be the underutilization of mental health services. Greater efforts should be spent devising ways to increase access and reduce barriers to services for AAPI veterans. Examination of the needs of specific AAPI veteran subgroups may be informative in developing targeted interventions. Moreover, research on AAPI soldiers and veterans should not only sample mental health service users or rely solely on administrative program data, which exclude those who do not use services.

Finally, discrimination against AAPI soldiers in the military and against AAPI veterans in their civilian lives are rarely a topic of study. AAPI soldiers experienced racism from fellow soldiers during the Vietnam War, but studies are needed on the current military

environment toward AAPI soldiers. AAPIs, as a group, have experienced a long history of discrimination from various American institutions and policies,49,50 are often treated as a “model minority,”11 and have been “invisible” in public health policies and debates.45 Perceived discrimination can discourage AAPIs from seeking mental health51 and is linked to poorer mental health outcomes.39,51

Conclusions

There has been little research on the mental health of AAPI veterans as evidenced by the limited number of empirical studies. AAPI veterans constitute only a small proportion of the veteran and the AAPI population, but existing studies suggest AAPI veterans have poorer mental health than other veterans and are not receiving needed mental health services. It appears AAPI veterans are a group that has been understudied in relation to veterans of other race/ethnic veteran groups4–6 and the few studies that do exist are mainly limited to veterans who served in the Vietnam War. A better understanding of current AAPI veterans is needed to elucidate ways to improve access to and quality of the mental health services they deserve.

BACK TO TOP

***Sexual Health***

## Military Sexual Trauma Complicates PTSD for Female Veterans

Medscape
Troy Brown
9 Nov 2012

Editor's note: Female veterans who experience military sexual trauma (MST) often find it difficult to seek help; even with treatment, they can live with the symptoms of posttraumatic stress disorder (PTSD) for years, according to a presentation at the American Psychiatric Nurses Association (APNA) 26th Annual Conference, held October 7 to 10 in Pittsburgh, Pennsylvania.

Ursula Kelly, PhD, ANP-BC, PMHNP-BC, assistant professor at the Emory University School of Nursing and a nurse scientist at Atlanta Veterans Affairs (VA) Medical Center in Georgia, works with female veterans who have experienced MST. She is currently conducting research on factors that influence the ability of these women to seek treatment for PTSD.

Dr. Kelly spoke with Medscape Medical News about her presentation in a telephone interview.

Medscape: What prompted you to study this issue?

Dr. Kelly: We have a treatment team dedicated to military sexual trauma and we all observed that many patients come to us after having had PTSD symptoms for years, or even decades. I really wanted to know what kept them from coming in sooner, and what could we do to help women access care more easily.

Medscape: Why is it important for psychiatric nurses to know about this?

Dr. Kelly: The VA is the largest employer of nurses in the country. It's important for psychiatric nurses to learn that it's not strictly a psychological process; there are situational factors that influence [women's] treatment-seeking.

The VA is a very male-dominated environment. For women who experience sexual trauma while serving in the military and who have PTSD, there are incredible challenges for them — sometimes just walking into the building.

It would help for nurses to have a broader view of where women are coming from and what outside factors have influenced their decision to seek treatment or not.

Medscape: Why is your research focused on women?

Dr. Kelly: The number of veterans who are women continues to grow. The VA system was designed to treat the healthcare needs of men; women have specific healthcare needs, so there's a bit of a disconnect in some areas.

More importantly, women have historically been either invisible or ignored within the VA; there's been a historic lack of treatment available for women. [This] population...has not had [much] attention for a long time. That is changing, for sure. There's public discourse around women in the military and PTSD, and even military sexual trauma. It's an issue that people are just understanding is a problem, but we still need to learn how to help these veterans.

Medscape: What factors prevent women from seeking treatment?

Dr. Kelly: Very little research has been done in this area. In studies of male veterans, things like stigma around PTSD and negative views about mental health treatment are well documented; it's likely that may be the case for women. We're not done with the data collection or data analysis, but what we've found so far in this study is that, for women, the stigma is much less about PTSD and more about the MST.

The male-dominated environment is a very big issue, particularly for women who have been sexually traumatized in the military, most of whom have also been traumatized outside of their military service.

There are very high rates of interpersonal trauma, both in childhood and in adulthood, in this population of veterans. They're really coming in with a lot of cumulative trauma. We need to understand how that plays into coming in for treatment that is specifically billed as PTSD related to MST.

Medscape: What factors make it easier for women to seek treatment?

Dr. Kelly: A few factors have come up in our analysis. Some of the women who have sought treatment at our clinic didn't know that services were available. So for some women, just knowing that these services are available helps.

For other women, it was getting enough messages from people around them — whether it was family members, friends, or healthcare providers — that they needed help. For people who are experiencing a lot of symptoms of PTSD, and who have been for years, they may view their behavior as normal. "Doesn't everyone stay inside all day? Doesn't everybody check the perimeter of their house before they go to sleep?"

The other important factor is having someone in their life who is really supportive and encourages them to get treatment, so that it doesn't have to be a secret and so they get support and encouragement even when treatment gets tough, as it often does.

Medscape: How effective are current treatments?

Dr. Kelly: We don't know how effective treatment is for MST. For PTSD in general, there are evidence-based treatments that are primarily trauma-focused; however, only about a third of all veterans with PTSD actually seek treatment. Even among those who are receiving evidence-based treatment, about another third continue to suffer from PTSD symptoms for years. There's a fairly high dropout rate as well.

We do know that those treatments are effective for women who experience sexual assault as civilians.

There are additional components that need to be addressed when the PTSD is related to MST. Unlike civilian sexual assault, a woman who is raped in the military often experiences that within a unit; the rapist may be a peer or peers or may be her commanding

officer. The woman has to continue to live there, in the presence of the perpetrator, for however long she is there. That affects her ability to get help at the time because, if she reports it, she has to live with whatever consequences there might be. These tend to be very negative, very problematic and, for many women, serve as obvious barriers to them disclosing the MST, which means they're then living with all of that, sometimes for years. Women may also fear being seen as weak or a poor soldier for seeking help.

Also, being assaulted by someone whom you should be able to trust is more damaging than being assaulted by a stranger, psychologically and emotionally.

Medscape: How can nurses improve services and treatment for female veterans who have experienced MST?

Dr. Kelly: Nurses can improve services by making sure our veterans know they exist, and by understanding what's behind some of the behaviors they see in female veterans in a context that really frames them as indicative of someone who has PTSD and lots of trauma.

Some of the behaviors that result from symptoms of PTSD can be looked upon as highly problematic by healthcare providers. Veterans may appear really angry and be somewhat difficult to work with; they may miss lots of appointments. If nurses can understand that it's not that they're being disrespectful or irresponsible, it's because they have PTSD and couldn't get themselves out of the house, that will really help.

Understand that for some female veterans who have PTSD, the PTSD is related to sexual trauma; don't automatically assume that it's combat related. Understanding that can shift the way you approach a patient when you're talking about it.

Medscape: Do you have any future research planned?

Dr. Kelly: I hope to have IRB approval in a week or so for a really exciting study. We're going to provide a trauma-sensitive yoga intervention for female veterans with PTSD who experienced MST. There are evidence-based treatments, but they're all trauma-focused, which means it's very hard work; women have to revisit the trauma, talk about it, write about it, and work through it in an emotionally very difficult way. There are women who won't or can't go there. Because sexual trauma can cause women to be disconnected from their bodies and disconnected from themselves, yoga presents itself as an embodied mental health treatment that may provide a pathway to healing for some women that straight cognitive and emotional therapy does not.

It's important for the yoga to be trauma-sensitive, because for women who have experienced sexual trauma, some components of yoga could be very triggering — some of the poses, some of the language from instructors. It's important to structure sessions so that women feel in control of what's happening, not that they're being told to do this, then do that, where they might feel a loss of control.

Medscape: Where can nurses learn more?

Dr. Kelly: The best resource is the National Center for PTSD. It has information about PTSD across the board, not just specific to veterans or active military. It's a tremendous resource, for consumers and patients and for professionals and researchers.

Dr. Kelly has disclosed no relevant financial relationships.

BACK TO TOP

***Substance Abuse***

## A longitudinal assessment of alcohol intake and incident depression: the SUN project

BMC Public Health
Alfredo Gea (ageas@alumni.unav.es) Miguel A Martinez-Gonzalez (mamartinez@unav.es) Estefania Toledo (etoledo@unav.es) Almudena Sanchez-Villegas (asanchez@dcc.ulpgc.es) Maira Bes-Rastrollo (mbes@unav.es) Jorge M Nuñez-Cordoba (jnunezco@unav.es) Carmen Sayon-Orea (msayon@alumni.es) Juan J Beunza (jjbeunza@unav.es)
7 Nov 2012

Methods

We assessed 13,619 university graduates (mean age: 38 years, 42% men) participating in a Spanish prospective epidemiological cohort (the SUN Project), initially free of depression. They were recruited between 1999–2008 and biennially followed-up during 2001–2010. At baseline, a 136-item validated food frequency questionnaire was used to assess alcohol intake. Wine was the preferred beverage. Participants were classified as incident cases of depression if they reported a new clinical diagnosis of depression by a physician and/or initiated the use of antidepressant drugs. Cox regression and restricted cubic splines analyses were performed over 82,926 person-years.

Results

Only among women, an U-shaped relationship between total alcohol intake and depression risk was found (P=0.01). Moderate alcohol intake (5–15 g/day) was associated with lower risk (Hazard Ratio: 0.62; 95% Confidence Interval: 0.43-0.89). No association was apparent for higher intakes of alcohol or for any specific type of alcoholic beverage.

Conclusions

Moderate alcohol intake might protect against depression among women. Further confirmatory studies are needed.

Keywords

Alcohol intake, Depression, Cohort analysis, Mediterranean population

## Background

Unipolar depressive disorder is the most prevalent mental disease in the world and it is increasing steadily [1]. The prevalence of major depression may reach up to 21% in some populations [2]. Depression is the third leading cause of global disease burden measured in Disability Adjusted Life Years (DALYs), and the first one measured in Years Lost to Disability [3]. If nothing is done, depression will become the first leading cause of global disease burden (DALYs) in 2030 [3]. Besides, 2.25 million deaths per year in the world may be attributed to alcohol intake, even after subtracting the beneficial effects of moderate alcohol intake on the development of cardiovascular disease [4]. In addition, alcohol intake is the 8th global death risk factor and the 3rd risk factor for disease and disability measured in DALYs [4]. Some cross-sectional studies have shown high rates of comorbidity of depression and Alcohol Use Disorders (AUD), which includes alcohol dependence and alcohol abuse [5,6]. Although some cohort and case–control studies have investigated this relationship, the longitudinal association between the quantity of alcohol intake and clinically diagnosed incident depression has not been assessed so far. Moreover reverse causation is a very important threat for the validity of cross-sectional or short-term follow-up studies. Besides, alcohol intake enormously varies among different populations regarding the type of alcoholic beverage, quantity and drinking pattern. We aimed to prospectively evaluate the long-term influence of alcohol intake on the development of depression during a 10 years follow-up in a Mediterranean cohort, where red wine was the most consumed alcoholic beverage.

## Methods

### Subjects

The “Seguimiento Universidad de Navarra” (SUN) Project is a prospective dynamic cohort, which started in 1999, following the model of the Nurses‟ Health Study and the Health Professionals Follow-up Study [7]. Details of the design and methods of this cohort have been described elsewhere [8,9]. Biennial mailed questionnaires were used to obtain updated information from participants. Up to February 2008, 19,576 subjects were recruited. Of these, 17,462 were successfully followed-up (at least one follow-up questionnaire), achieving a retention rate of 89.2%. Subjects with energy intake out of predefined limits (800 Kcal/day or 4,000 Kcal/day in men and 500 Kcal/day or 3,500 Kcal/day in women) [10] were excluded (n=1,675). Other 1,844 subjects who reported prevalent, personal history of previous depression, or use of antidepressants at baseline, and 324 subjects who reported incident depression in the first follow-up questionnaire, up to 2 years of follow-up, were also excluded. The exclusion of incident cases at the second year of follow-up is discussed below.

Finally 13,619 participants were included in the analyses (Figure 1). The study was approved by the Institutional Review Board of the University of Navarra. Informed consent was implied by the voluntary completion of the baseline questionnaire.

Figure 1 Flow Chart of participants: The SUN project

Exposure assessment

Alcohol intake was assessed at baseline with a semi-quantitative food-frequency questionnaire (FFQ) that included questions on alcoholic beverage consumption during the past year (red wine, other wines, beer, and spirits). In the validation study for this questionnaire, the correlation coefficient for alcohol intake between the FFQ and four food records was 0.90 [11].

Participants were divided into four groups according to their baseline alcohol intake: abstainers, those who reported drinking less than 10 g/day of alcohol, those with an intake ranging from 10 and 25 g/day and the fourth group with an intake higher than 25 g/day. These groups were reclassified according to the results of the spline in 4 new categories: abstainers, less than 5 g/day, between 5 and 15 g/day, and more than 15 g/day. In the stratified analysis by type of beverage, alcohol intake was considered as a continuous variable, and adjustment for total alcohol intake was performed through the residuals method [12].

Outcome assessment

Incident cases of depression were defined as a positive answer by participants in any of the follow-up questionnaires (4th, 6th, 8th, or 10th year) to the question „Have you ever been diagnosed of depression by a medical doctor?" or a positive report after 4 or more years of follow-up habitual use of antidepressant drugs. We excluded the early cases (depression diagnosis made or antidepressant use reported within the first 2 years of follow-up), to avoid reverse causation bias. Antidepressant use was ascertained through an open question in which the participants reported their habitual medication use. This definition of incident depression was validated by a Psychiatrist in a sub-sample of our cohort using the Structured Clinical Interview for DSM-IV (SCID-I) as a Gold Standard, obtaining a specificity of 0.96; a

percentage of confirmed depression of 74% and a percentage of confirmed non-depression of81% [13].

Covariate assessment

We obtained the information about medical, socio-demographic, anthropometric, and lifestyle variables from the baseline questionnaire. Physical activity was assessed through a validated physical activity questionnaire [14]. Adherence to the Mediterranean Dietary Pattern (MDP was evaluated combining 8 items (fruits and nuts, vegetables, fish, legumes, cereals, dairy products, meat and meat products, and the ratio Monounsaturated Fatty Acids/Saturated Fatty Acids) according to the score proposed by Trichopoulou et al. [15], but excluding alcohol intake, that was not taken into account to build the MDP in order to avoid overlapping with our main exposure.

Statistical analysis

Cox regression models were used to assess the relationship between the four categories of alcohol intake and the incidence of depression. Hazard ratios (HR) and their 95% confidence intervals were calculated using the abstainers group as the reference category. The Cox model included age as the underlying time variable. Birth date was used as the origin variable. Entry time was defined as age at recruitment. Exit time was defined as age at diagnosis of depression for cases and for participants who did not develop depression as age when completing the last follow-up questionnaire or as age at death (whichever occurred first). For the multiple-adjusted model, the following potential confounders were considered: smoking, physical activity (MET-h/week), total energy intake (Kcal/day), body mass index (kg/m2), adherence to the MDP, marital status, and employment status. We evaluated the interaction between sex and alcohol intake by introducing an interaction term in the model. Although there was not an effect modification by sex, the analyses were also conducted separately for men and women. We conducted sensitivity analyses re-running all the models after: 1) including also early cases (up to the first two years of follow-up), 2) excluding late cases of depression (after 10 years), 3) excluding prevalent cases of other psychiatric diseases (insomnia, schizophrenia, anxiety, anorexia and bulimia, stress, obsessive compulsive disorder, bipolar disorder, phobias) at baseline, 4) excluding prevalent cases of cancer or cardiovascular disease (angina pectoris, coronary surgery including bypass, coronary angioplasty, stroke including thrombosis, embolism, and cerebral haemorrhage, paroxysmal tachycardia, atrial fibrillation, aortic aneurysm, heart failure, pulmonary embolism, peripheral venous thrombosis, intermittent claudication, and all types of cancer) at baseline, 5) depurating the abstainers group (excluding former drinkers and participants who do not drink due to medical causes), 6) excluding participants under anxiolytic drugs or other psychiatric medication at baseline, 7–9) excluding participants under anxiolytic, antipsychotic, antiepileptic or anticonvulsant medication during the follow-up period. Finally, we evaluated the potential non-linear association between alcohol intake and incident depression non-parametrically calculating the HR and 95% CI with restricted cubic splines [16], stratified by sex. Tests for non-linearity used the likelihood ratio test comparing the model with only the linear term to the model with the linear and the cubic spline terms. The results were adjusted for the same potential confounding factors as the main Cox regression analysis. We also conducted spline analysis on alcohol independently for each alcoholic beverage (beer, wine, spirits) as a continuous variable. All P-values were two-tailed and $P<0.05$ was considered significant.

Results

The main characteristics of the 13,619 participants (5,701 males and 7,918 females) categorized according to their alcohol intake are presented in Table 1. High alcohol intake (>25 g/day) was associated with being male (86% were male), older (mean age was 46 years), and with higher BMI (mean BMI 26 Kg/m2). A total of 459 incident cases of depression were identified during the follow-up period, which summed up 82,926 person-years.

The Cox regression analysis considering all the participants in the study categorized in four groups at cut-off points (>0, 10, and 25 g/day) showed an inverse association that was not statistically significant (Table 2).

We did not find any effect modification of alcohol according to the sex of the participants (P=0.644). Although the interaction was non-significant, we fitted again the Cox regression stratified by sex. The reasons for this sub-groups analysis are discussed below. Stratified analysis showed no association between alcohol intake and incident depression (Table 2). We carried out multiple

sensitivity analyses to rule out possible sources of bias in the estimation of the association between alcohol intake and depression (data not shown; available on request). We repeated the analyses after excluding prevalent and life-time cases of psychiatric diseases at baseline (n=85), after excluding prevalent or life-time cases of cardiovascular disease or cancer at baseline (n=410), after excluding later incident cases of depression reported after 10 years of follow-up (n=8), after depurating the abstainers group excluding former drinkers and participants who do not drink due to health reasons (n=227), after including earlier incident cases (n=324), after excluding participants under anxiolytic drugs or other psychiatric medication at baseline (n=324), and after excluding participants under anxiolytic, antipsychotic, or antiepileptic and anticonvulsant medication during the follow-up period (n=455, n=114, n=36 respectively). The associations between alcohol intake and depression observed in all these sensitivity analyses were similar to those obtained in the main analysis.

To account for a non-linear association, we performed restricted cubic spline analysis adjusted for the same potential confounding factors as the main Cox regression analysis. Among men, we found no statistically significant linear or non-linear association between alcohol intake and incident depression (P=0.21 and P=0.49 respectively). We found no relationship for any type of specific alcoholic beverage nor for total alcohol intake among men. Among women, we found a significant inverse relationship between total alcohol intake and the incidence of depression. This relationship was statistically significant for the nonlinear association (P=0.01). The U-shaped association that was found is graphically represented in Figure 2. Women with an alcohol intake ranging from 5 to 15 g/day showed a significantly lower risk of depression (the HR and the 95% CI were below the null value). In order to obtain estimates for specific categories of alcohol intake, we fitted another Cox regression analysis adapting the building of groups of alcohol intake to the shape of the dose– response curve obtained using the restricted cubic spline model: abstainers, <5 g/day, 5–15 g/day, and >15 g/day. This resulted in a HR and 95% CI of 0.62 (0.43-0.89) for the third category (5–15 g/day) as compared to abstainers but no association was found for the rest of categories (Table 3).

No associations were found for the consumption of specific beverage types and depression risk in the overall sample neither in the analyses separated by sex (data not shown).

## Discussion

The results of this prospective study suggest that whereas moderate alcohol intake (5–15 g/day) may protect women against the development of depression, higher amounts of alcohol intake may not confer any protection against depression.

Many studies have investigated the relationship between exposure to alcohol and depressive status. However, most of them have focused on subjects with AUD more than in subjects with alcohol intake levels in a lower range. Some other studies [17,18] have longitudinally investigated the relationship between alcohol dependence or alcohol abuse, but not categories of alcohol intake, and major depression. In fact, a recent review [19] reported a positive association between AUD and major depression.

However, longitudinal studies focusing on a normal range of alcohol intake even if they are follow-up studies, may still be affected by reverse causation bias because the follow-up is relatively short (only 1–2 years), or because they have not excluded prevalent cases of epression, or have assessed depression only on a narrow window of time during follow-up.

Moreover no Mediterranean cohort, with a higher red wine consumption, has ever assessed this association. Therefore the differences in the assessment of exposure and/or events, or in the type of patients evaluated make the comparison with these studies difficult. A study from the Health and Retirement Study [20] found an association between problematic alcohol intake and depression in a cohort of men aged over 50, but no association was found with non-problematic alcohol intake. Biennially for 6 years, alcohol intake was assessed through questionnaires, and incidence of depression was assessed using a symptoms scale with a cutpoint. However, 63% of the participants were classified as problematic alcohol drinkers even when they reported drinking less than 1 drink per week. In addition, some cases of depression might be underestimated in that study since they only evaluated self-reported cases of depression during the last 2 years of the study, but not those occurring in the first 4 years of follow-up. Finally, that study only included men since the exposure was problematic drinking, which is much more prevalent among men. Another study examined how alcohol use predicts changes in psychological symptoms among 393 adolescents followed up for 18 months [21] and found that initial levels of alcohol use did not predict changes in depression. Inclusion criteria for that study were to score one standard deviation above the school mean on one of the four subscales of the Substance use Risk Personality Scale: hopelessness, anxiety, impulsivity, and sensation seeking. In addition, depression was assessed through the Brief Symptom Inventory, not through a doctor made diagnosis. Finally, previous or prevalent cases of depression were not excluded since diagnosis of depression as such was not assessed. Paljärvi et al. [22] performed a two-wave 5-years follow-up study to determine which aspect of drinking pattern would be the best predictor for depressive symptoms. Selfreported depressive symptoms were measured with a questionnaire at baseline and at year-5. They found that binge-drinking and also higher categories of alcohol intake were directly associated with depressive symptoms. These results are in contrast with our findings, however Paljärvi el al. excluded abstainers from their analyses, and used the moderate alcohol consumers as the reference category. Skogen et al. [23] found an U-shaped association between alcohol intake and both anxiety and depression. They distinguished between the possible former drinkers and abstainers. However, these authors acknowledged that the direction of causality was not clear. Another study evaluated the influence of alcohol dependence on the first incident depressive episode in a 1-year follow-up [24]. Both disorders were investigated using Diagnostic Interview Schedule, according to DSM-III. Prevalent and previous cases of depression were excluded. Multivariate logistic regression was conducted separately for men and women, obtaining a stronger positive association for women. These results are not comparable with ours because we did not assess alcohol dependence as exposure. Some case–control studies have investigated the influence of alcohol intake on the development of depression [25-27]. For example, an study conducted by Armenian et al. [27] after an earthquake found a protective association for alcohol intake in cases of depression without comorbidity of other psychiatric disorders: OR (95% CI) =0.6 (0.3-0.9). The posttraumatic situation makes our results not comparable with the results of this study. Two other studies found contradictory results [25,26]. However, these case–control studies may also be affected by reverse causality bias.

Finally, some cross-sectional studies have found high comorbidity between both conditions [4]; however, their cross-sectional nature is an obstacle to assess causality, especially when it is well known that depression is a risk factor for AUD [17,18]. Our results

suggest an U-shaped association between alcohol intake and the development of depression, only for women. The non-significant association for heavy alcohol intake may be due to the very low average consumption and the scarcity of heavy drinkers in our cohort. This is consistent with Perreira et al. [20] who did not find any association for men. However, our results are not consistent with Mackie et al. [21] who did not find any association for men or women. These studies, and all cited above, are not generally comparable because of the differences in the alcohol and depression assessments [28]. Moreover, the different results between these studies may be due to the differences in the distribution of alcohol intake between populations. When the population has moderate to low average alcohol intake, the association tends to be inverse or null [20,21] and when the population has heavy alcohol intake, it tends to be positive [20,22]. When assessing the relationship between alcohol intake and the development of depression a careful consideration of the temporal sequence is needed. In order to reduce the risk of falling into reverse causation bias, early incident cases of depression were excluded. These cases may be sub-clinical cases of depression that are likely to increase their alcohol intake as a consequence of their recent depression. On the contrary, alcohol intake information was collected at baseline and some cases of depression were developed after 10 years of follow-up. Considering on the other hand that this late incident cases might not be influenced by alcohol intake at baseline, we excluded these later cases and the magnitude of the association did not change considerably. The development of depression may be influenced by the presence of other psychiatric disorders [29]. In order to control for this possible bias we conducted the analysis after the exclusion of prevalent and lifetime cases of psychiatric disorders. The association did not change substantially. Other important diseases like cancer [25,30] or cardiovascular disease may lead to the development of depression and also to changes in alcohol intake. After the exclusion of prevalent cases of both kinds of disorders, the association remained similar. Other potentially important issue is the comparison group. Abstainers might be a heterogeneous group. Those who do not drink may have different reasons not to do so. Some of them may be abstainer due to medical advice. Participants declaring being abstainers could have been former drinkers. These facts could bias the results towards a more protective association. However, after excluding former drinkers and those abstainers who did not drink due to medical advice, the association between alcohol intake and the development of depression did not change. We conducted the analysis separately for men and women although the interaction productterm was not significant. However, previous studies [31,32] found differences between both sexes, therefore we considered convenient to split our sample. Women seem to be more likely to the development of depression [33] and the same alcohol intake affects differently to men and women [34] due to their different body composition, affecting the volume of distribution, or different metabolism. Some mechanisms have been proposed to explain the deleterious effect of heavy or chronic alcohol drinking on brain function [35]; however, moderate alcohol drinking have a GABAergic effect, acting on GABAA receptors [36], that may prevent or counteract the effects of depression on this system [37].

Our study has some important strengths, such as its prospective design (avoiding reverse causation bias), data collection and data analysis. Other strengths are its large sample size, its high retention rate, the good adjustment for potential confounders, the existence of published validation studies for our methods, the high correlation observed between the FFQ and the food records for alcohol intake in the validation study, and the high educational level of the participants, achieving high quality information and high internal validity. In our analysis, we have carefully considered the reverse causation bias, used validated diagnosis of depression, taken into account other psychiatric disorders, analysed sex-alcohol interaction, excluded other important diseases, and depurated the abstainers group. A limitation of our study is the use of a self-reported clinical diagnosis of depression. We assume that a low

proportion of participants could misreport the diagnosis. Moreover, participants who reported habitual use of antidepressant drugs were also considered incident cases. Research suggests first treatment for depression typically occurs several years following the onset of depression. The median time from onset of depression to first treatment has been reported to be 8 years in the U.S. population [38]. Long lags in first contact for depression treatment have also been reported in other countries [39]. Thus, identifying incident cases of depression by having received a diagnosis or prescribed medication by a medical doctor may result in missed cases. This is especially true for cases recruited later in the study who were followed up for a shorter period of time than case recruited earlier in the study. These missing cases can translate in low sensitivity of our case ascertainment definition. However, our validation study found very high specificity (0.96) for the self-reported diagnosis of depression and theoretically with perfect specificity, nondifferential misclassification of disease, due to low sensitivity, will not bias the relative risk estimate [40].

A potential limitation is that few participants were heavy alcohol consumers, which may lead to a lack of statistical power to assess the relationship with depression for high levels of alcohol intake. Another limitation is that we do not have information about illicit drug use. However the SUN cohort participants are highly motivated responsible and health-conscious who voluntarily agreed to complete long and complicated questionnaires. We expect that the use of illicit drugs will be very low or even non-existent among them.

Conclusions

Moderate alcohol intake (5–15 g/d) among women might protect them against the development of depression, while high alcohol intake seems to confer no benefit. However, additional cohort studies are needed to confirm our findings.

Abbreviations

SUN, Seguimiento universidad de Navarra; HR, Hazard ratio; 95% CI, 95% Confidence interval; DALYs, Disability adjusted life years; AUD, Alcohol use disorders; FFQ, Food frequency questionnaire; SCID-I, Structured Clinical Interview for DSM-IV; MDP, Mediterranean Dietary Pattern; MET, Metabolic equivalent; BMI, Body mass index; OR, Odds ratio; GABA, Gamma aminobutiric acid.

Competing interest

There is no conflict of interest. There is no financial arrangement with any food or alcoholic beverages company.

Authors' contributions

AG conducted the literature review, participated in the design of the study, conducted the main statistical analyses and prepared the first draft of the manuscript. ET participated in the design of the study, conducted part of the statistical analyses and supervised the methods section. ASV participated in the design of the study, conducted part of the literature review, contributed to the interpretation of findings, the writing of the discussion section, and obtained funding. MBR participated in the design of the study, the interpretation of results, and obtained funding and administrative support. JMNC participated in the interpretation of results. CSO helped with the

literature review and with the statistical analyses. JJB directed and supervised the study, participated in the statistical analyses, the interpretation of results, and obtained funding and administrative support. MAMG was the founder and principal investigator of the SUN cohort, supervised all the steps in the statistical analyses and preparation of the manuscript, and obtained funding and administrative support. All authors took care of the critical revision of the manuscript for important intellectual content and approved the final version to be submitted for publication.

Acknowledgements

We are indebted to the participants of the SUN Study for their continued cooperation and participation. We thank to other members of the SUN Group: Alonso A, Benito S, de Irala J, de la Fuente-Arrillaga C, Delgado-Rodríguez M, Guillén- Grima F, Krafka J, Llorca J, López del Burgo C, Marti A, Martínez JA, Pimenta AM, Sánchez D, Seguí-Gómez M, Serrano- Martínez M, and Vázquez Z.

The SUN Project has received funding from the Spanish Government-Instituto de Salud

Carlos III (Grants PI01/0619, PI030678, PI040233, PI042241, PI050976, PI070240,

PI070312, PI081943, PI080819, PI1002658, PI1002293, RD06/0045, 2010/087, and

G03/140), the Navarra Regional Government (36/2001, 43/2002, 41/2005, 36/2008) and the

University of Navarra.

BACK TO TOP

## Energy Drink Consumption and Its Association with Sleep Problems Among U.S. Service Members on a Combat Deployment — Afghanistan, 2010

Centers for Disease Control and Prevention
Robin L. Toblin, PhD, Kristina Clarke-Walper, MPH, Brian C. Kok, Maurice L. Sipos, PhD, Military Psychiatry Br, Walter Reed Army Institute of Research; Jeffrey L. Thomas, PhD, US Army Medical Research Unit-Europe, US Army
9 Nov 2012

Beverages marketed as energy drinks have become a popular form of caffeine consumption targeted at young males, with some brands containing the caffeine equivalent of 1–3 cups of coffee or cans of soda (1). Energy drinks also include other ingredients intended to boost physical energy or mental alertness, such as herbal substances, amino acids, sugars, and sugar derivatives; however, caffeine is the main active ingredient (1). Approximately 6% of adolescent and young adult males in U.S. civilian and military populations consume energy drinks daily (2,3). These products generally are unregulated and can have negative side effects

(e.g., caffeine intoxication, overdose, withdrawal, and poor interactions with alcohol) (1). Paradoxically, excess consumption also can increase sleep problems and daytime sleepiness, which can impair performance (1). To determine the extent of energy drink use and the association with sleep problems and sleepiness during combat operations, Walter Reed Army Institute of Research analyzed data collected by Joint Mental Health Advisory Team 7 (J-MHAT 7) to Operation Enduring Freedom in Afghanistan in 2010. The analysis showed that 44.8% of deployed service members consumed at least one energy drink daily, with 13.9% drinking three or more a day. No differences by age or rank were found. Service members drinking three or more energy drinks a day were significantly more likely to report sleeping ≤4 hours a night on average than those consuming two drinks or fewer. Those who drank three or more drinks a day also were more likely to report sleep disruption related to stress and illness and were more likely to fall asleep during briefings or on guard duty. Service members should be educated regarding the potential adverse effects of excessive energy drink consumption on sleep and mission performance and should be encouraged to moderate their energy drink consumption in combat environments.

Mental Health Advisory Teams conduct comprehensive mental health surveillance of U.S. service members in combat environments and have administered the Deployment Well-Being Survey in Iraq during 2003–2009 and Afghanistan during 2005–2010 and 2012. The survey version used by J-MHAT 7 to collect data in Afghanistan during the summer of 2010 asked about demographic characteristics, deployment history, combat experiences, mental health, deployment stressors, family and relationship concerns, work environment, sleep difficulties and daytime sleepiness, health-care utilization, and various health behaviors, including energy drink consumption. The J-MHAT 7 survey was the first to inquire about the use of energy drinks.

In total, 1,249 service members were surveyed using a cluster sample of randomly selected U.S. Army and Marine combat platoons deployed to Afghanistan. All participants were male, because of the type of unit surveyed. Of those surveyed, 1,000 consented to have their data used for research purposes and 988 answered the following question: "How many energy drinks (e.g., Monster, Red Bull, 5-Hour Energy) do you use per day?" The six response options ranged from zero to five or more drinks per day. Service members also were asked about their use of sleep medication, average number of hours of sleep per day, concerns regarding lack of sleep, disruptions to sleep, and work impairment associated with sleepiness (Table 1). The number of sleep hours was dichotomized at ≤4 hours (reported by 24.2% of the persons sampled); in comparison, 50.2% of those sampled reported sleeping ≤5 hours. For comparison across sleep outcomes, energy drink use was divided into the following categories: no drinks, one to two drinks, and three or more drinks per day. These cutoffs were chosen because previous research demonstrated that 200 mg of caffeine, the equivalent of one to two energy drinks, improved cognitive performance in a military population (4). Prevalence rates of energy drink use are reported. Chi-square tests were used to determine significant differences between groups for sleep variables, using $p<0.05$ for significance. Post hoc analyses of the chi-square tests were conducted by examining discrepancies between observed and expected values for standardized residuals to produce z-scores and identify those cells contributing to the significant differences. The Sidak-Bonferroni correction was used to account for conducting multiple post hoc tests.

Service members surveyed were predominantly on active duty (93.2%), of junior enlisted rank (E1–E4; 71.2%), on their first deployment (60.8%), in the Army (75.5%), aged 18–24 years (66.6%), single (54.5%), not parents (70.9%), in the military <5 years

(81.2%), and had been on this deployment <6 months at the time of the survey (54.3%). The prevalence of daily energy drink use was 44.8%; 13.9% consumed three or more per day (Table 2). Of those reporting daily energy drink use, 56.6% consumed more than one energy drink per day. No associations were found between the proportion of service members reporting the number of drinks used per day (i.e., 0, 1, 2, 3, 4, or ≥5) and rank category, number of deployments, branch of service, age, marital status, or being a parent. In the same comparison, however, service members in the National Guard or Reserves were significantly more likely to use energy drinks than their active duty counterparts (p=0.002).

Service members who drank three or more energy drinks per day were more likely to report ≤4 hours of sleep on average per night (38.2%) than service members who drank one to two (18.4%) or zero (23.9%) energy drinks per day (Table 3). The groups did not differ in their levels of concern regarding not getting enough sleep. Service members drinking three or more energy drinks per day were significantly more likely than the other groups to report sleep disruption on more than half the nights in the past 30 days because of stress related to combat, stress related to personal life, and illness. However, no differences were noted in sleep disruption because of the sleep environment, high operational tempo, nighttime duties, or leisure activities. Service members who drank three or more energy drinks per day also were significantly more likely to report sometimes or often falling asleep while sitting in briefings or while on guard duty, but not while riding in convoys. No differences in energy drink consumption were found related to having had an accident or making a mistake that affected the mission because of sleepiness (Table 3). Despite a significant, omnibus chi-squared association, after post-hoc analyses were conducted, no differences were found in sleep medication use or receiving prescriptions for sleep medications while deployed by levels of energy drink consumption (Table 3). Across sleep disruption and daytime sleepiness outcomes, service members who consumed one to two energy drinks did not differ from those not consuming energy drinks.

## Editorial Note

Military and civilian findings show that more than half of adolescents and young adults drink at least one energy drink per month (5), with approximately 6% consuming energy drinks daily (2,3). In this study, 45% of service members reported consuming one or more energy drinks per day, a considerably higher prevalence. This might reflect the unique and extreme demands of a combat deployment and the widespread availability of energy drinks in the combat environment (e.g., free distribution in dining facilities and available for purchase in convenience stores). No differences in energy drink consumption by age or rank were observed, demonstrating the ubiquitous nature of energy drink consumption during deployment.

Consumption of three or more energy drinks per day was associated with negative sleep outcomes that included sleepiness on the job and sleeping ≤4 hours per night. This is a low number of hours of sleep even in the deployed environment, in which half of respondents averaged ≤5 hours of sleep. Although causality could not be ascertained from this cross-sectional study, this relationship is consistent with civilian studies demonstrating that caffeine use contributes to daytime sleepiness (6) and sleep problems (1,6), and that inadequate sleep and daytime sleepiness can impair work productivity (7). Further, this study suggests that high levels of energy drink consumption might indirectly impair performance in a military setting. Service members who consumed

three or more energy drinks per day reported significantly greater sleep disruption because of combat stress, personal issues, and illness, but not because of external factors. This is similar to results found in a civilian study in which caffeine use caused an increase in nocturnal worry and sleeplessness (8) and a military study that found that mental health symptoms increased energy drink use (9). Because inadequate sleep can considerably influence a person's health, excessive energy drink consumption might indirectly contribute to poor health.

The findings in this report are subject to at least five limitations. First, cause and effect cannot be determined because the data are cross-sectional. It is unclear whether service members with sleep problems used more energy drinks to stay alert, or if heavy use of energy drinks led to sleep disruptions; published studies suggest a cyclical combination of both (1,5). Second, the survey did not allow for a true estimate of caffeine intake. The caffeine content in energy drinks varies by the size of the can and milligrams of caffeine per ounce (1). Leading brands contain 80–160 mg of caffeine in their smallest containers, similar to 1–2 cups of coffee, with some brands containing up to 500 mg (1). In addition, the survey did not measure consumption of other caffeinated beverages (e.g., coffee, soft drinks, or tea). Third, the phrasing of the question about average number of energy drinks consumed per day might have resulted in an underestimate of energy drink use; a person who consumed several drinks a week, but did not consume them daily, might have answered zero to that question. Fourth, this study did not control for variables that might have confounded the relationship between energy drink consumption and sleep outcomes (e.g., mental health problems, physical injury, amount of time deployed, or peer group/unit effects). Nonetheless, survey data from questions about stress, illness, personal life, and leisure activities as reasons for sleep disruption might serve as proxies for those variables not analyzed. Finally, analyses did not control for sleep medication use, which also can cause daytime sleepiness. However, although groups differed in overall sleep outcomes, the groups did not differ in their prevalence of sleep medication use (approximately one in seven), suggesting that the main associations were not explained by use of sleep medication.

The widespread use of energy drinks across demographics and its association at high doses with sleep problems and work impairment, coupled with known associations between caffeine and sleep problems and sleepiness in the general population (1,6,7), support the need to educate service members about moderating consumption of energy drinks. Service members who used energy drinks in moderation (i.e., one or two per day) had similar levels of sleep problems and performance as those who did not use energy drinks. Based on the caffeine content of leading brands of energy drinks, this dosage is equivalent to the average caffeine consumption by men ages 20–29 years in the United States (10) and has been associated with cognitive performance (e.g., visual vigilance, reaction time, and alertness) (1,4). This also might explain the lack of a clear dose-response relationship between energy drink consumption and sleep problems.

The marketing of these types of drinks as energy boosters, together with their availability in the combat environment, makes it easy for service members to consume them in large volumes. Energy drinks are relatively new, generally unregulated, and lack warning labels. Service members should be educated that the long-term health effects of energy drink use are unknown, that consuming high doses of energy drinks might affect mission performance and sleep, and that, if used, energy drinks should be consumed in moderation.

Acknowledgments

Study participants. Paul Bliese, Barry Adams, Michael Slack, Robert Heinssen, Edward Edens, Derrick Polk, Matthew McGinnis, Joint Mental Health Advisory Team 7; Randall J. Nett, MD, Office of Public Health Preparedness and Response, CDC.

BACK TO TOP

## Behavioral Counseling After Screening for Alcohol Misuse in Primary Care: A Systematic Review and Meta-analysis for the U.S. Preventive Services Task Force

Annals of Internal Medicine
Daniel E. Jonas, MD, MPH; James C. Garbutt, MD; Halle R. Amick, MSPH; Janice M. Brown, PhD; Kimberly A. Brownley, PhD; Carol L. Council, MSPH; Anthony J. Viera, MD, MPH; Tania M. Wilkins, MS; Cody J. Schwartz, MPH; Emily M. Richmond, MPH; John Yeatts, MPH; Tammeka Swinson Evans, MOP; Sally D. Wood, BA; and Russell P. Harris, MD, MPH
6 Nov 2012

Abstract

Background
Alcohol misuse, which includes the full spectrum from risky drinking to alcohol dependence, is a leading cause of preventable death in the United States.

Purpose
To evaluate the benefits and harms of behavioral counseling interventions for adolescents and adults who misuse alcohol.

Data Sources
MEDLINE, EMBASE, the Cochrane Library, CINAHL, PsycINFO, International Pharmaceutical Abstracts, and reference lists of published literature (January 1985 through January 2012, limited to English-language articles).

Study Selection

Controlled trials at least 6 months' duration that enrolled persons with alcohol misuse identified by screening in primary care settings and evaluated behavioral counseling interventions.

Data Extraction
One reviewer extracted data and a second checked accuracy. Two independent reviewers assigned quality ratings and graded the strength of the evidence.

Data Synthesis
The 23 included trials generally excluded persons with alcohol dependence. The best evidence was for brief (10- to 15-minute) multicontact interventions. Among adults receiving behavioral interventions, consumption decreased by 3.6 drinks per week from baseline (weighted mean difference, 3.6 drinks/wk [95% CI, 2.4 to 4.8 drinks/wk]; 10 trials; 4332 participants), 12% fewer adults reported heavy drinking episodes (risk difference, 0.12 [CI, 0.07 to 0.16]; 7 trials; 2737 participants), and 11% more adults reported drinking less than the recommended limits (risk difference, 0.11 [CI, 0.08 to 0.13]; 9 trials; 5973 participants) over 12 months compared with control participants (moderate strength of evidence). Evidence was insufficient to draw conclusions about accidents, injuries, or alcohol-related liver problems. Trials enrolling young adults or college students showed reduced consumption and fewer heavy drinking episodes (moderate strength of evidence). Little or no evidence of harms was found.

Limitations
Results may be biased to the null because the behavior of control participants could have been affected by alcohol misuse assessments. In addition, evidence is probably inapplicable to persons with alcohol dependence and selective reporting may have occurred.

Conclusion
Behavioral counseling interventions improve behavioral outcomes for adults with risky drinking.

Primary Funding Source
Agency for Healthcare Research and Quality.

Alcohol misuse, which includes the full spectrum from risky or hazardous drinking to alcohol dependence (1Â -Â 3), is associated with numerous health and social problems and more than 85Â 000 deaths per year in the United States (4Â -Â 5). Alcohol misuse is the third leading cause of preventable death in the United States, after tobacco use and being overweight (6). It contributes to hypertension, cirrhosis, gastritis, gastric ulcers, pancreatitis, breast cancer, neuropathy, cardiomyopathy, anemia, osteoporosis,

cognitive impairment, depression, insomnia, anxiety, suicide, injury, and violence (7Â -Â 9). The definitions of the spectrum of alcohol misuse (that is, unhealthy alcohol use 1) continue to evolve. For this review, we use the definitions in Table 1 (5,10Â -Â 12).

About 30% of the U.S. population misuse alcohol, with most engaging in what is considered risky drinking (1). Recent U.S.-based data (13) revealed that 21.3% of primary care patients reported risky drinking.

Cross-sectional and cohort studies have consistently related high average alcohol consumption and heavy per-occasion use to short- or long-term health consequences (14Â -Â 15). A meta-analysis examining the association between all-cause mortality and average alcohol consumption (16) found that men who drank an average of at least 4 drinks per day and women who drank an average of at least 2 drinks per day had increased mortality relative to nondrinkers. The National Institute on Alcohol Abuse and Alcoholism has proposed guidelines (17) to limit the risks for drinking-related consequences. The maximum recommended consumption is 3 or fewer standard drinks per day (â‰¤7 drinks/wk) for adult women and anyone older than 65 years, and 4 or fewer standard drinks per day (â‰¤14 drinks/week) for men (15,17Â -Â 18). These guidelines do not apply to persons for whom alcohol intake is contraindicated, such as pregnant women, persons with alcohol dependence or medical conditions that can be worsened by drinking, or those receiving medications that interact with alcohol.

Behavioral counseling interventions include the range of personal counseling and related behavior-change interventions that are used to help patients change health-related behaviors (19). â€œCounselingâ€□ here denotes a cooperative method of work that demands active participation from both patient and clinician and aims to facilitate the patient's independent initiative (19). The goal of behavioral interventions for alcohol misuse is to eliminate risky drinking practices (for example, by encouraging fewer drinks per occasion or not drinking before driving) rather than to achieve abstinence.

For the Effective Health Care Program of the Agency for Healthcare Research and Quality (AHRQ) and to assist the U.S. Preventive Services Task Force (USPSTF) in updating its 2004 recommendation statement (20), we conducted a systematic review and meta-analysis of the effectiveness of screening followed by behavioral counseling, with or without referral, for alcohol misuse in primary care settings (21). The full report (21) addressed 7 questions (Appendix Table 1).

## Methods

We developed and followed a standard protocol. A technical report that details methods and includes search strategies and additional evidence tables is available at www.effectivehealthcare.ahrq.gov/reports/final.cfm.

Key Questions and Analytic Framework

The USPSTF and the AHRQ determined the focus of this review.

Data Sources and Searches

We searched MEDLINE, EMBASE, the Cochrane Library, CINAHL, PsycINFO, and the International Pharmaceutical Abstracts from 1 January 1985 to 31 January 2012, limited to English-language articles. The start date was selected on the basis of the earliest publication date found in previous reviews and expert opinion. We used Medical Subject Headings as search terms when available and keywords when appropriate, focusing on terms to describe relevant populations, screening, and behavioral interventions.

Study Selection

We developed inclusion and exclusion criteria with respect to populations, interventions, comparators, outcomes, timing, settings, and study designs (22Â -Â 23). For the question related to behavioral interventions, we included randomized, controlled trials of at least 6 months' duration that enrolled adults or adolescents with alcohol misuse identified by screening in primary care settings and that evaluated whether a counseling intervention improved behavioral or health outcomes.

Two investigators independently reviewed titles and abstracts, and then another 2 investigators independently reviewed the full text of all articles marked for possible inclusion during the initial review to determine final inclusion or exclusion. Disagreements were resolved with an experienced team member.

Data Extraction and Quality Assessment

We designed and used structured forms to extract pertinent information from each article, including information about the methods and populations, interventions, comparators, outcomes, timing, settings, and study designs. All data extractions were reviewed for completeness and accuracy by a second team member.

We assessed the quality (internal validity) of studies using predefined criteria based on those developed by the USPSTF (ratings of good, fair, or poor) (24) and the University of York Centre for Reviews and Dissemination (25). These included assessment of the adequacy of randomization, allocation concealment, similarity of groups at baseline, masking, attrition, and whether intention-to-treat analysis was used. Two independent reviewers assigned quality ratings for each study. Disagreements were resolved by an experienced member of the team.

Data Synthesis and Analysis

We stratified evidence by population (adults, older adults, young adults or college students, and pregnant women). Quantitative analyses were conducted of outcomes reported by a sufficient number of studies that were homogeneous enough to justify combining their results. We used random-effects models. For the outcome of alcohol consumption, the effect measure was mean difference between the intervention and control groups for change from baseline in drinks per week. The percentages of patients who had episodes of heavy drinking and those who achieved recommended drinking limits were compared (between intervention and control groups) with a risk difference. Because follow-up periods varied, the analysis for all-cause mortality was based on deaths per person-year and the comparison between intervention and control groups was calculated as a risk ratio. Analyses were conducted by using Comprehensive Meta Analysis, version 2.2.055 (BioStat, Englewood, New Jersey).

We used subgroup analyses to explore whether results differed by intensity, sex, country, deliverer of the intervention, or setting. The chi-square and I2 statistics were calculated to assess heterogeneity in effects between studies (26Â -Â 27). When quantitative analyses were not appropriate (for example, because of heterogeneity, insufficient number of similar studies, or insufficient or varied outcome reporting), we synthesized the data qualitatively.

To assess the differential effects of using more or less time and single or multiple contacts, we grouped interventions by intensity of counseling, as measured by the duration and number of contacts: very brief (â‰¤5 minutes, single-contact), brief (6 to 15 minutes, single-contact), extended (>15 minutes, single-contact), brief multicontact (each contact â‰¤15 minutes), or extended multicontact (some contacts >15 minutes).

We then graded the strength of evidence (SOE) as high, moderate, low, or insufficient on the basis of the guidance established for the Evidence-based Practice Center Program (Appendix Table 2) (28). Two reviewers assessed each domain for each key outcome, and differences were resolved by consensus.This report focuses on the key questions related to benefits and harms of behavioral interventions.

Role of the Funding Source

This review was funded by AHRQ. Staff of AHRQ and members of the USPSTF participated in developing the scope of the work and reviewed draft manuscripts. Approval from AHRQ for copyright assertion was required before the manuscript could be submitted for publication, but the authors are solely responsible for the content and the decision to submit it for publication.

Results

We included 38 articles reporting on 23 randomized, controlled trials (Appendix Figure 2). Sample sizes ranged from 72 to 1559, and study durations ranged from 6 to 48 months (Appendix Table 3). Eleven studies were done solely in the United States, 2 focused on older adults, 5 focused on young adults or college students, and 1 enrolled pregnant women. We identified no studies of adolescents.

Fourteen of the interventions (29Â -Â 50) were delivered by a primary care physician alone or with a health educator or nurse. Three (51Â -Â 54) were delivered by a nurse or physician assistant, 1 by a psychologist (55Â -Â 57), 2 by a researcher (58Â -Â 62), and 1 by unspecified interventionists (63). Two interventions in college students (64Â -Â 66) were conducted via a computer. Most trials tested brief multicontact interventions (31Â -Â 34,42,46,50Â -Â 51,53,64Â -Â 65) or brief interventions (29,49,52,58,60,62Â -Â 66); fewer tested very brief (45,63), extended (30), or extended multicontact interventions (38,45,48,55,60). Interventions were heterogeneous and included various counseling approaches, such as brief advice, feedback, or motivational interviews, and cognitive behavioral strategies, such as self-completed action plans, written health education or self-help materials, drinking diaries, or problem-solving exercises to complete at home (Appendix Table 4). Most comparator groups received screening or assessment followed by usual care or by provision of a general health pamphlet. A few studies included additional components in comparator groups that could have biased results toward the null, such as recording screening or assessment results on the chart (45) or forwarding them to physicians (60), advice from nurses on reducing drinking and a leaflet with benchmark alcohol guides (52), a pamphlet on the health effects of alcohol consumption (64Â -Â 66), or a booklet about preventing alcohol problems (48). We summarize the main findings by population and outcome and report the SOE for each.

### Screening

We found no studies meeting inclusion criteria that randomly assigned participants, practices, or providers to screening and a comparator (no studies addressing questions 1 or 3) (Appendix Table 1). We found adequate evidence that several screening instruments can detect alcohol misuse in adults with acceptable sensitivity and specificity (21). The full technical report includes additional details about the accuracy of screening tests.

### Effectiveness for Improving Intermediate Outcomes

Table 2 summarizes the results of meta-analyses for consumption, heavy drinking, and recommended drinking limits, by population. The Figure shows the forest plots for 12-month outcomes from our meta-analyses for adults. Overall, evidence supports the effectiveness of behavioral interventions for improving several intermediate outcomes for adults, older adults, and young adults or college students. For pregnant women, the included study (250 participants) (30) did not provide evidence of effectiveness for improving intermediate outcomes over 6 months or longer (low or insufficient SOE, depending on the outcome). Subgroup analyses identified no significant differences between men and women. Brief multicontact interventions had the best evidence of effectiveness across populations and outcomes and had follow-up data spanning several years. Meta-analyses of studies in adults found very brief

and brief single-contact interventions to be ineffective for some outcomes and less effective than brief multicontact interventions for others.

Effectiveness for Reducing Morbidity, Reducing Mortality, or Changing Other Outcomes

Table 3 summarizes results, by population. Our meta-analyses found no statistically significant reduction in all-cause mortality for adults (rate ratio, 0.64 [95% CI, 0.24 to 1.7]; 4 trials) or for all age groups combined (rate ratio, 0.52 [CI, 0.22 to 1.2]; 6 trials). Point estimates trended toward favoring interventions, but few studies reported mortality and few long-term data were available. No studies that enrolled pregnant women and reported these outcomes were found (insufficient SOE).

Potential Adverse Effects

We found no evidence of direct harms, aside from opportunity costs associated with interventions, which ranged from 5 minutes to 2 hours dispersed over several in-person or telephone visits (moderate SOE). We searched for evidence of potential adverse effects, such as illegal substance use, increased smoking, anxiety, stigma, labeling, discrimination, or interference with the physicianâ€“patient relationship. We found no evidence for most of these potential harms and very limited evidence reporting no difference between groups for smoking rates and anxiety (low SOE). Other than the results for opportunity costs, our results are limited by the few trials that reported any information; 5 of 23 reported smoking (29,33Â -Â 34,39,41,49Â -Â 50), and 2 reported anxiety (29,49).

Health Care System Influences

Where the study was conducted (United States vs. nonâ€“United States) had no impact on the effectiveness of interventions for consumption outcomes. Data showed a tendency toward greater reduction in consumption for interventions delivered in academic- or research-oriented settings than for those delivered in community-based settings (weighted mean difference, âˆ’5.0 drinks/wk [CI, âˆ’7.6 to âˆ’2.5 drinks/wk] vs. âˆ’3.2 drinks/wk [CI, âˆ’4.3 to âˆ’2.2 drinks/wk]; 3 vs. 7 trials). Interventions delivered mostly by primary care providers showed a tendency toward greater reduction in consumption than did those delivered primarily by research personnel (weighted mean difference, âˆ’4.0 drinks/wk [CI, âˆ’5.4 to âˆ’2.6 drinks/wk] vs. âˆ’3.0 drinks/wk [CI, âˆ’5.0 to âˆ’1.0 drinks/wk]; 7 vs. 2 trials). Our consumption meta-analysis included only 1 intervention delivered by a nurse (52), and the reduction was not statistically significant in that study (weighted mean difference, âˆ’0.2 drinks/wk [CI, âˆ’8.9 to 8.6 drinks/wk]). Two other studies, each of which provided insufficient data for our consumption meta-analysis, reported benefits of interventions delivered primarily by nurses (51) or

by nurses and physician assistants (53) for some consumption outcomes. In addition, 2 interventions conducted by computer reported some evidence of effectiveness for reduced consumption in college students (64Â -Â 66).

## Discussion

We found no studies that directly addressed our overarching question (key question 1)â€”no studies randomly assigned patients, practices, or providers to screening and comparator groups and subsequently provided interventions for those with positive screening results. All of the included studies randomly assigned patients after they had received positive screening results.

We found that behavioral counseling interventions improved drinking behavior outcomes (moderate SOE) and reduced hospital days (low SOE) for adults with risky drinking. For most health outcomes, available evidence either found no difference between intervention and control groups, such as for mortality (low SOE), or was insufficient to draw conclusions, such as for alcohol-related liver problems (insufficient SOE). Long-term outcomes from 2 studies (33,35Â -Â 37,39,42Â -Â 43) revealed that participants in the intervention groups maintained reductions in consumption or continued to reduce consumption, but differences between intervention and control groups were no longer statistically significant by 48 months. Studies identified delayed reduction in consumption in control groups that could reflect the natural history of alcohol consumption, the cumulative effect of follow-up with the health care system, differential attrition (if more participants lost to follow-up in the control group were risky drinkers), or (late) regression to the mean.

The evidence for effectiveness in adults is strongest for brief multicontact interventions. The effect sizes for these interventions were greater than those for other intensities (although CIs often overlapped). In addition, the best studies show that the effect of brief multicontact interventions remains for several years (35Â -Â 36,43) and also show improvement for some utilization outcomes, such as fewer hospital days (35Â -Â 36) and costs (benefitâ€“cost ratio of 39:1 over 48 months [CI, 5.4 to 72.5]) (36).

The brief multicontact interventions generally lasted 10 to 15 minutes per contact. All of the brief multicontact interventions in our meta-analyses of behavioral outcomes at 12 months were delivered by primary care providers, sometimes with additional intervention from a nurse or health educator. For example, the intervention in Project TrEAT (Trial for Early Alcohol Treatment) (33) included two 15-minute visits with a primary care provider 1 month apart and two 5-minute follow-up phone calls from a nurse 2 weeks after each visit. The intervention also included feedback about health behaviors, a review of problem drinking prevalence, a list of the adverse effects of alcohol, a worksheet on drinking cues, a drinking agreement or prescription, and drinking diary cards. Of note, 2 studies of brief multicontact interventions in adults, both of which provided insufficient data for our meta-analyses, reported

benefits of interventions delivered primarily by nurses (51) or by nurses and physician assistants (53) for some consumption outcomes.

Evidence suggests that very brief interventions (up to 5 minutes, single-contact) and brief interventions (up to 15 minutes, single-contact) are less effective or ineffective, depending on the outcome. Although extended multicontact interventions seem to be effective for improving intermediate outcomes, we found no evidence that they are more effective than brief multicontact interventions.

The only included study that enrolled pregnant women (250 participants) (30) found no difference in reduced consumption between groups but did find higher rates of continued abstinence among women who were abstinent before the assessment in the intervention group than among those in the control group. Our searches identified other studies focusing on pregnant women that did not meet our inclusion criteria (67Â -Â 84). Several took place in such settings as jails or specialized drug and alcohol treatment centers (75), and others lacked a control group or followed participants for fewer than 6 months (73,84). Several of these studies reported benefits of interventions, including reduced consumption (73,84), reduced risk for an alcohol-exposed pregnancy (75), higher rates of abstinence (79), and better fetal and newborn outcomes (higher birth weights and lengths and reduced fetal mortality rates 79).

We have described several categories of alcohol misuse (such as risky or hazardous use and alcohol dependence). These categories are not all discrete (an individual may meet the definition for more than one). Included trials generally enrolled participants with risky or hazardous drinking, but the trials used varying terminology to describe the populations and often enrolled heterogeneous samples. Nevertheless, most investigators excluded participants with alcohol dependence or constructed their inclusion and exclusion criteria to substantially limit the number of such participants. Our best assessment is that our overall findings apply to risky or hazardous drinkers but not to persons with alcohol dependence. It is uncertain whether our findings apply to harmful drinkers or persons with alcohol abuse.

All interventions required support systems to provide screening; screening-related assessment; and in some cases, provider prompting. Screening assessments were often multistep processes that included interviews with research personnel that lasted up to 30 minutes. Less time would be required for screening and screening-related assessments in primary care practice. We estimate that 5 to 10 minutes would be required for persons who had positive screening results, with most of the time used to assess whether such persons have alcohol abuse or dependence (and should probably be referred for specialized treatment) as opposed to risky or hazardous drinking (for which behavioral counseling interventions in primary care may be effective). Nevertheless, support systems are probably required for effective screening and intervention. In addition, most interventions required training providers or staff.

It is unclear whether our findings apply to persons with certain comorbid conditions, and some researchers have suggested that brief behavioral interventions may be ineffective or less effective in people with comorbid psychiatric conditions. A subgroup analysis from a German study (56) found that brief interventions did not reduce drinking among 88 participants with comorbid anxiety or depression. Although most trials in our review did not exclude persons with depression, anxiety, or chronic pain, it is unclear how many participants with these conditions were included in most trials.

A previous systematic review (85) found no evidence of efficacy for brief behavioral interventions in patients with alcohol dependence in primary care settings. Our review also found no such evidence. Included studies that enrolled more than 10% of participants with alcohol dependence reported interventions to be ineffective or less effective than studies that did not enroll alcohol-dependent participants.

Screening for alcohol misuse will inevitably identify some alcohol-dependent individuals; thus, providers and those making recommendations need information about whether effective interventions are available for alcohol dependence. If complete abstinence is used as an outcome, 15% to 35% of patients have been reported to achieve 1 year of sobriety after such treatment approaches (86) as pharmacotherapy, motivational enhancement therapy, cognitive behavioral therapy, 12-step facilitation, and therapy at alcoholism-treatment centers. Similar sobriety outcomes at 3 to 5 years or longer have been reported (9).

Our review has limitations. First, the scope of our review was limited to primary care settings. Second, most evidence involved self-report of alcohol use. Investigators in some trials verified self-reported use with other persons (such as family members). Self-report of alcohol use has been found to be accurate if collected carefully (87Â -Â 88). Third, the assessments conducted in the included trials could have concealed benefits of interventions (and biased results toward the null) by causing behavior changes. Control participants generally reduced alcohol consumption. Possible explanations include increased awareness of drinking, discussions with their provider about drinking that were prompted by the screening questions, receipt of some minimal intervention (control groups in the included studies often received some printed educational materials), or regression to the mean. A recent systematic review (89) concluded that answering questions on drinking in brief intervention trials seems to alter subsequent self-reported behavior, potentially generating bias by exposing nonintervention control groups to an integral component of the intervention. Finally, publication bias and selective reporting may be present.

In conclusion, behavioral counseling interventions improve intermediate outcomes, such as alcohol consumption, heavy drinking episodes, and drinking above recommended amounts (moderate SOE) and may reduce hospital days (low SOE) for adults with risky or hazardous drinking. For most health outcomes, available evidence found no difference between intervention and control groups,

such as for mortality (low SOE), or was insufficient to draw conclusions about the effectiveness of behavioral interventions, such as for alcohol-related accidents or quality of life (insufficient SOE). Brief multicontact interventions (about 10 to 15 minutes per contact) have the best evidence of effectiveness for adults.

BACK TO TOP

## Performance enhancing drug abuse and cardiovascular risk in athletes: implications for the clinician

British Journal of Sports Medicine
Peter J Angell; Neil Chester; Nick Sculthorpe; Greg Whyte; Keith George; John Somauroo
November 2012

Abstract

The use of performance-enhancing and social drugs by athletes raises a number of ethical and health concerns. The World Anti-Doping Agency was constituted to address both of these issues as well as publishing a list of, and testing for, banned substances in athletes. Despite continuing methodological developments to detect drug use and associated punishments for positive dope tests, there are still many athletes who choose to use performance and image enhancing drugs. Of primary concern to this review are the health consequences of drug use by athletes. For such a large topic we must put in place delimitations. Specifically, we will address current knowledge, controversies and emerging evidence in relation to cardiovascular (CV) health of athletes taking drugs. Further, we delimit our discussion to the CV consequences of anabolic steroids and stimulant (including amphetamines and cocaine) use. These drugs are reported in the majority of adverse findings in athlete drug screenings and thus are more likely to be relevant to the healthcare professionals responsible for the well-being of athletes. In detailing CV health issues related to anabolic steroid and stimulant abuse by athletes we critique current research evidence, present exemplar case studies and suggest important avenues for on-going research. Specifically we prompt the need for awareness of clinical staff when assessing the potential CV consequences of drug use in athletes.

Introduction

In a bid to improve performance and/or aid recovery, various pharmaceutical products have been used, both openly (legally) and in a clandestine manner against the rules of governing bodies, by a broad array of athletes. The World Anti-Doping Agency (WADA)

strictly regulates the use of pharmaceutical products in competitive sport. WADA produced and regularly updates the World Anti-Doping Code that includes a prohibited drug list. This list dictates what is and is not acceptable, from a doping perspective, within sport. The list has various subsections with some drugs banned both 'in' and 'out' of competition, while others are banned 'in' competition only (eg, ephedrine and the cannabinoids). The WADA list of prohibited substances is further broken down into several subcategories of specific products, including anabolic agents, peptide hormones and growth factors; β-2 agonists; hormone and metabolic modulators; diuretics and masking agents; stimulants; narcotics; cannabinoids; and glucocortocosteroids. Despite strict rules and punishments being in place in an attempt to limit doping offences, there continues to be those who choose to try and gain an unfair advantage in sport by taking drugs.

As well as the concerns of WADA in defending the spirit of free and fair competition, awareness is also raised when there are known health consequences of drug abuse. The negative health consequences of doping for athletes and the education of support staff are the focus of this narrative review. Clearly, doping and athlete health is a vast area and so two points of delimitation are made up-front. First, we concentrate on cardiovascular (CV) health consequences of drug use. Second, this review focuses on anabolic agents and stimulants. According to WADA's adverse analytical findings report from 2010, anabolic agents accounted for around 60% of adverse findings, with stimulants contributing around 10%. Consequently, the clinical support teams working with athletes must understand the potential CV health consequences when athletes abuse these drugs. Finally, we will discuss the CV effects of cocaine use, as it is one of the most widely used recreational drugs detected in athletes. As well as reflecting on previous data related to the CV health consequences of anabolic steroid, stimulant and cocaine use we have attempted to provide extra context and information in the form of brief case-study exemplars on anabolic steroid and cocaine abuse.

## Anabolic agents

There are a number of drugs that are used in an attempt to increase lean muscle mass. Of these, the most well known is the steroid hormone testosterone (T) and the various analogues that are based around testosterone, usually referred to as androgenic anabolic steroids (AS). In addition to AS there are non-steroid agents that are used in an attempt to generate the same anabolic effects. These include the β-2 agonist, clenbuterol (which is additionally used as an anorectic agent to reduce body fat), human growth hormone (HGH) and insulin/insulin-like growth factors. Other growth factors are commonly used in between courses of anabolic agent use and these include human chorionic gonadotropin (HCG) and erythropoietin (EPO) and more recently selective androgen receptor modulators. Whatever the classification, the purpose of anabolic and growth factors is to stimulate skeletal muscle growth and promote rapid recovery following intensive training. From a clinical perspective it is also worth noting that EPO is predominantly used to boost endurance exercise performance and in general has not crossed over to amateur and recreational sports performers.

AS, HGH, HCG and insulin/insulin-like growth factors, however, are routinely used by professional, amateur and recreational athletes. Consequently, the likelihood of clinicians coming across users of these particular drugs is much greater.

It is worthy to note that the classification of AS covers a number of structural variants. Classically, AS are classified as water-soluble orally active forms (17-α-alkylated) and lipid-soluble parenteral forms (17-β-esterified). In addition, they are often also classified as either testosterone-based, dihydrotestosterone-based (DHT) or 19-nortestosterone-based (Nandrolone) all of which have differing properties and expected side effects. The situation is further complicated by belief among users, often stemming from anecdotal advice, that some AS are better for predominantly 'bulking' (eg. Deca-Durabolin) while others are better suited to losing body fat or 'cutting' (eg. Winstrol). Users will often use these different forms of AS in varying quantities. The use of AS is also characterised by periods of use followed by periods of abstinence, or 'cycles'. This helps to maximise the effects of the drugs while also limiting the negative consequences and allowing the body to normalise following an 'on' cycle. Furthermore, users will often supplement their cycles with additional pharmaceutical agents both when bulking (eg, Insulin, human growth hormone) and when losing body fat (clenbuterol, cytomel, 2,4, dinitrophenol). Finally, there are a surprising number of drugs used to attempt to limit side effects of AS use or normalise the hypothalamo-pituitary-gonadal (HPG) axis following an AS cycle. These include estrogen receptor antagonists (tamoxifen), selective estrogen receptor inhibitors (clomifene), aromatase inhibitors (arimidex), 5-α reductase inhibitors (finesteride) and HPG axis stimulators such as HCG.

Evidence of athlete use of AS has been available since the 1950s with AS contributing to c. 60% of adverse findings according to recent WADA reports. In the general population there are data showing an increase in the prevalence of AS use. Despite such widespread use there is still some controversy as to the CV health consequences of taking AS. Large sample epidemiological evidence of the CV health consequences of long-term AS use is lacking, likely because of the reluctance to admit use and/or possession. In addition, evidence for a link between AS use and CV disease outcomes or end-points is mostly limited to case study reports. Published case studies include AS use associated with myocardial infarction, stroke, embolism and other CV health issues. Although caution should be expressed in implying cause and effect from case studies, they can provide direction for case series and experimental studies as well as informing/educating clinical practitioners.

Cardiovascular events and risk factors from AS use

Significant research attention has focused on the impact of AS use on CV disease risk factors namely blood pressure, lipid profile, left ventricular (LV) mass, cardiac function and arterial function. Elevated systemic arterial blood pressure is associated with an increased CV disease risk. Compared to healthy controls, AS users have increased resting, and exercise systolic blood pressure. Conversely, other studies have not observed increased blood pressure in AS user. Differences in the training level of the participants along with age could be responsible for the differences seen in these studies.

AS have also been associated with negative alterations in lipid profiles. Changes reported include a decrease in high-density lipoprotein (HDL), an elevation in low-density lipoprotein (LDL) and reduced apolipoprotein levels, possibly through up-regulation of hepatic triglyceride lipase. The changes in lipid profiles indicate an increase in atherosclerotic risk. Increases in homocysteine, a naturally occurring amino-acid thought to have a role in vaso-control, and C-reactive proteins (CRP), an acute-phase protein that rises in response to inflammation, have been implicated as risk factors for CV disease. Grace and Davies demonstrated a significant increase in CRP in AS users. While Zmuda *et al* observed no significant increases in homocysteine in a group of AS users, Graham *et al* noted a significant elevation in homocysteine in AS users as well as those who had abstained from AS use for 3 months, indicating a possible effect of AS on vitamin B absorption. Previous studies have also suggested a possible link between AS use and thrombotic risk through alterations in haemoglobin levels.

An increase in LV mass is an independent risk factor for CV disease. AS use has been associated with an increase in LV mass, but there is conflicting data. There are some data in AS users that suggest a reduction in systolic cardiac function although this is not a consistent finding between studies. A reduction in diastolic function has been observed more frequently and it has been suggested that a reduction in myocardial relaxation/elastance is associated with AS use. AS use has also been associated with reduced endothelial function in conduit arteries. Ebenbichler *et al* and Sader *et al* noted a reduced flow-mediated dilation in AS users as well as a reduced vasodilator response to glyceryl-trinitrate.

There is a growing evidence base that AS use can have a negative effect on multiple CV disease risk factors. We present a case study exemplar to illustrate the broad effect of AS use on the CV system.

Case study 1

A 25-year-old bodybuilder who was enrolled into a research study gave a detailed history of prolonged AS use. He self-reported no cardiac health problems. The participant was 1.93 m and 127 kg with an LV mass of 218 g on cardiac MRI. While absolute LV mass was high this was normalised when indexed for fat free mass. Left ventricular (LV) ejection fraction was normal (63%), while right ventricular (RV) ejection fraction was slightly reduced at 49%. LV diastolic function, measured using ultrasound echocardiography, was moderately depressed (significant increases in late atrial filling and tissue velocities). Heart rate was 75 beats per min and the ECG was unremarkable. Blood pressure was 131/71 mm/Hg. Total cholesterol was within normal clinical limits (4.6 mmol/l), but high-density lipoproteins were significantly reduced in the participant at 0.46 mmol/l. Low-density lipoproteins were above clinically acceptable levels at 3.77 mmol/l. The participant had an elevated level of alanine transaminase indicative of reduction in liver function. In addition, both testosterone and sex-hormone binding globulin were well below those expected within a male, at 2.13 and 6 nmol/l, respectively. Gamma glutamyltransferase was not assessed. Based on changes to numerous CV risk factors, this athlete was informed of the test outcomes and was directed to seek a general practitioner consultation.

While recent data have started to paint a clearer picture of some of the negative CV consequences of AS use, longitudinal data are still lacking making the long-term chronic effects of AS use difficult to ascertain. There is also uncertainty in relation to the impact of withdrawal of AS use on CV risk factors. Inherently, the study of AS use is complicated by many factors. Specifically, most studies recruit diverse cohorts of self-selected AS users. In these groups there is likely great heterogeneity in total AS dose, poly-drug regimens employed as well the difficulty in verifying the 'true' dosages used.

There is little direction given to clinicians and health practitioners regarding the identification of clandestine AS use. One of the reasons for athletes engaging in illicit AS use is that there are few 'tell tale' signs particularly in those sports where elevated strength and/or fat free mass is relatively commonplace. AS use occurs predominantly but not exclusively in strength and/or power athletes. Of the nine adverse findings for AS by the UK Anti-Doping Agency, three were from professional rugby players. It is worth noting that clinicians are far more likely to encounter recreational athletes using AS, training for predominantly aesthetic purposes. AS users may have rapid increases in lean mass, excessive hypertrophy of the trapezius, the shoulder musculature and a large neck circumference possibly due to higher androgen receptor levels. Other signs of AS use include puffy swelling around the face and acne which occurs in approximately 50% of users (including acne fulminans and acne conglobata). Consequently, clinicians may be suspicious of AS use in muscular individuals with acne (and/or scarring from previous eruptions) extending beyond the face to the back, chest and upper arms. Another common side effect of AS use is the development of female breast tissue (gynaecomastia) because of the aromotisation of AS to estrogen. The size and rate of development will vary depending on the type and duration of AS exposure with cases ranging from small areas of puffy tissue behind the nipple to easily recognisable breast tissue. It must be stressed however that these are general signs that may indicate AS use, but clinicians should be wary of unfairly targeting individuals who are simply prone to acne, or have a genetic basis for a muscular shoulder girdle.

Stimulants—amphetamines

Stimulants consist of psychoactive drugs, including amphetamines and its derivatives, as well as cocaine, caffeine and nicotine and are used to increase psychological activity, thereby inducing improvements in mental aptitude, physical function or both. Stimulants used for performance enhancement are typically those that affect the central nervous system (CNS); however, concurrent effects on the CV system are common. Amphetamines affect the CNS and autonomic nervous systems causing tachycardia and vasoconstriction as well as having psychological effects, increasing mental alertness and decreased fatigue. Cocaine has similar affects to amphetamines and also exerts its effect through the CNS causing an increase in the release of catecholamines, such as norepinephrine, as well as causing an increase of circulating dopamine. Ephedrine and its isomers are sympathomimetic agonists at both α- and β-adrenergic receptors and also have potent CV effects including increasing heart rate and blood pressure through vasoconstriction. The CV effects of stimulants have been shown to have a detrimental effect on CV health.

There have been numerous reported cases of serious adverse CV events, including fatalities, linked with the use of ephedrine's. The use of ephedrine has also been linked to cardiomyopathy and stroke while pseudoephedrine use has also been linked with stroke and coronary artery spasm with myocardial infarction. Further, there are numerous case reports of significant cardiovascular events following the administration of over-the-counter stimulants. As with AS use, the interpretation of such case reports is difficult, since it may be that serious adverse events are as a consequence of preexisting medical conditions combined with drug use. What case studies do promote is further experimental study. For example, doses equivalent to three to four times greater than the recommended therapeutic dose of pseudoephedrine have raised diastolic blood pressure above 90 mm Hg. These results were in accord with two other studies: Bye *et al* reported significant increases in heart rate and systolic blood pressure following relatively high doses of pseudoephedrine (120 and 180 mg); Empey *et al* noted that doses of 120 and 180 mg produced statistically significant increases in both pulse and systolic blood pressure. The clinical relevance of these blood pressure changes is not known.

Reports of the CV effects of sympathomimetics in therapeutic doses have been conflicting. Bye *et al* observed that a single dose of ephedrine (25 mg) significantly elevated both heart rate and systolic blood pressure while a single therapeutic dose of pseudoephedrine (60 mg) significantly elevated only systolic arterial blood pressure. Bright *et al* and Empey *et al* noted little change in CV function following therapeutic doses. Increased blood pressure has been demonstrated in cases whereby ephedrine's have been co-administered with a moderate dose of caffeine. Caffeine is thought to exacerbate the action of ephedrine's since it too may cause vasoconstriction through antagonism of adenosine and release of catecholamines.

All stimulants structurally related to amphetamine can cause catecholamine-mediated cardiotoxicity. Increased catecholamine levels can lead to vasoconstriction, vasospasm, tachycardia and hypertension and it is as a result of these responses that oxygen supply to the heart is compromised and hypertrophy, fibrosis and necrosis can result. Clearly, such conditions develop over time, as a consequence of chronic exposure to amphetamines and the repercussions may include myocardial infarction, aortic dissection and sudden cardiac death.

Little has been done by way of measuring both the acute/chronic effects of stimulant use/abuse on a range of CV risk factors. While significant CV events are associated with their use, the effect on long-term atherosclerotic risk even with moderate usage, through negative alterations in known CV risk factors is worthy of further examination. With regard to amphetamines, emphasis has largely focused on the adverse effects relating to the CNS; however, it is clear that their effects can be wide-ranging and no less significant.

Stimulants—cocaine and others

Cocaine is one of the most frequently used recreational drugs worldwide. The prevalence of use in the UK has been steadily increasing over recent years with 6.6% of 16–24 year olds admitting to regular use and figures suggesting over 1 million current

cocaine users in the UK. Despite limited use as a performance-enhancing drug, there is still a relatively high level of adverse findings for cocaine use in athletes. According to WADA around 11.3% of adverse findings for stimulants were as a result of cocaine use. Athletes are not immune to the lure of social drugs like cocaine.

The use of cocaine has been associated with acute and chronic cardiovascular disease. Cocaine inhibits norepinephrine reuptake in the sympathetic system leading to overstimulation and may cause release of catecholamines from central and peripheral stores. Acute coronary syndromes (including myocardial ischemia and infarction) are the commonest cardiac events secondary to cocaine abuse. This may be due to coronary artery spasm, increase in myocardial oxygen consumption from increases in heart rate and blood pressure, and a prothrombotic state. Most myocardial infarctions occur in the absence of atherosclerotic coronary disease and are unrelated to the dose and frequency of cocaine use. Cocaine abuse may however also lead to premature coronary disease sometimes with quite rapid onset. Cocaine abuse may also lead to coronary artery aneurysms, aortic dissection, rupture, vasculitis and stroke.

Arrhythmias are not common with cocaine use, but sinus tachycardia, sinus bradycardia, supraventricular, ventricular arrhythmias and bundle branch blocks have been reported. A dilated cardiomyopathy can be caused by cocaine use and a cocaine-induced myocarditis has been reported at postmortem in 20–30% of cases. Myocarditic changes, however, may be fully reversible if identified early and abstention from further cocaine abuse occurs. In the case study detailed the ECG changes with a significant cardiac event are shown in a young professional athlete who has a cocaine abuse history.

Case study 2

We briefly describe the case of a 27-year-old professional skater with a history of cocaine abuse. He developed retrosternal chest pain and ECG changes of inferolateral ST elevation after taking cocaine. His symptoms settled with nitrates and he developed T-wave inversion on his ECG the following day with a rise in Troponin I of 18.5 µg/l. He had normal left ventricular systolic function on echocardiography. On cardiac catheterisation he had a smooth 40% stenosis in the ostial left anterior descending artery with the other coronary branches normal. He was initially treated with antiplatelets, heparin and a β-blocker, but as it was felt that he had sustained a myocardial infarction due to coronary artery spasm, the β-blocker was substituted with a calcium channel blocker prior to discharge from hospital. He remained pain-free when further reviewed at 3 months.

While detection of cocaine use, unless openly admitted, can be problematic for the practitioner there are certain psychological and physiological indications that may help inform the healthcare provider. Cocaine use can produce euphoria, decreased appetite and need for sleep, but may also cause anxiety, irritability, paranoia and hallucinations. There may be associated tachycardia, sweating,

pupil dilatation and nausea. Withdrawal after chronic use mainly causes psychological symptoms including depression, anxiety and increased sleep.

Other common stimulants used by athletes include caffeine and nicotine (from smoking tobacco, chewing tobacco and smoking cessation products). Both caffeine and nicotine have been shown to have cardiac effects. Caffeine is used by athletes as a stimulant to decrease the feelings of fatigue and elevate heart rate and neural activity. The effect of caffeine will increase myocardial $O_2$ and has been shown to increase blood pressure as well as raises the level of free-fatty acids in the blood, thereby increasing blood viscosity. If used in moderate doses there are minimal long-term side effects to caffeine use; however, extremely high doses have been associated with cardiac arrhythmias and events in otherwise healthy individuals. Cigarette smoking has long been known to promote atherosclerosis, with an increased risk of sudden death, myocardial infarction and stroke. The use of nicotine therapies such as transdermal patches, inhalators and gum as well as the common use of chewing tobacco in some sports has also been associated with an increased cardiovascular event risk. Again, much of these data have come from case study and anecdotal data, with larger studies not observing a significant increase in risk of cardiovascular events. Like most other stimulants further research is required and given that both caffeine and nicotine are not completely banned admission and monitoring of use may be easier than with other drugs.

Summary

This review has attempted to describe available data and exemplar case studies detailing the potential CV health issues surrounding the use of some of the most common drugs detected in doping screens on athletes. Negative CV health effects do occur but more controlled research trials are required alongside longitudinal studies of chronic drug use. Available data demonstrate enough concern in relation to CV health and drug use that healthcare practitioners should be fully educated to help detect signs and symptoms and support athlete treatment and recovery where appropriate.

BACK TO TOP

***Suicide***

## Increase in state suicide rates in the USA during economic recession

Lancet
Aaron Reeves a, David Stuckler a b, Martin McKee b, David Gunnell c, Shu-Sen Chang c d, Sanjay Basu
6 Nov 2012

Evidence from European countries indicates a significant rise in suicides from the economic recession, totalling more than 1000 excess deaths in the UK alone.1—3 Among the worst affected economies in Europe, such as Greece, suicides have risen by more than 60% since 2007.2 Thus far, there has been little or no analysis of US mental health data, mostly owing to delays in data availability.

Here, we extend our previous analyses of recessions and suicides in Europe1, 3, 4 to assess trends in all 50 US states. We use data on suicide mortality rates from 1999 to 2010 from the Centers for Disease Control and Prevention. Unemployment data come from the Bureau of Labor Statistics. Time-trend regression models were used to assess excess suicides occurring during the economic crisis—ie, deaths over and above the level that would be expected if historical trends continued (see appendix for methodological details). Although there are concerns that suicide data are under-reported in the USA, these biases are likely to have been consistent over this relatively short period, although they might lead to a conservative estimate of the mental health effects of the crisis.

In the years before the onset of the crisis (from 1999 to 2007), the suicide mortality rate in the USA was rising on average at a rate of 0·12 per 100 000 per year (95% CI 0·09—0·14; figure). Coinciding with the onset of the recession, the suicide rate accelerated. There were an additional 0·51 deaths per 100 000 per year (95% CI 0·28—0·75) in 2008—10. This acceleration corresponds to an additional 1580 suicides per year (95% CI 860—2300). Thus, during the recessionary period after 2007, there were an estimated 4750 excess suicide deaths (95% CI 2570—6920).

Next we investigated the association between rising unemployment and suicide mortality rates. In Europe, we previously noted that a one percentage point rise in unemployment was associated with a rise in the suicide rate of 0·79% (95% CI 0·16—1·42; p=0·016).5 Our findings in the USA are slightly higher: a one percentage point rise in unemployment is associated with a 0·99% increase in the suicide rate (95% CI 0·60—1·38, p<0·0001), which is closer to the association estimated when there were no labour market protections (1·06%). The magnitude of these effects is slightly larger than for those estimated previously in the USA,6 which might indicate that previous studies have not investigated periods of high unemployment or that this recession might be exerting more negative effects on mental health than previous downturns.

Since the rate of unemployment between 2007 and 2010 in the USA increased from 5·8% to 9·6%, our model indicates that the rise in US unemployment during the recession is associated with a 3·8% increase in the suicide rate, corresponding to about 1330 suicides. In other words, rising unemployment could account for about a quarter of the excess suicides noted in the USA during this time.

Looking across US states between 1999 and 2010, we found that the strongest correlation between unemployment and suicides was in Texas (r=0·91), but overall the correlations were statistically indistinguishable between the north, south, east, and west, or when disaggregating states by Democrat and Republican governors (appendix). Small numbers of suicides in small populations limit a state-by-state comparison for all 50 states. Similar patterns were seen if absolute numbers of suicides were used instead of overall rates.

Suicide is a rare outcome of mental illness; these data are likely to be the most visible indicator of major depression and anxiety disorders, as seen in primary-care settings in Spain7 and in the Greek population.8 The pattern of accelerating suicides noted in the USA mirrors that recorded for economic reasons in Italy.4

Future research should explore other risk factors such as foreclosures and job and income losses, and modifying factors such as gun control policies, access to the means of self-harm, and vulnerable groups, which could explain the remaining portion of the suicide rise observed during the recession.

Our findings have immediate implications for policy. Given that some countries have avoided increases in suicides despite significant economic downturns, there is a clear need to implement policy initiatives that promote the resilience of populations during the ongoing recession. Active labour market programmes—projects that immediately help the unemployed find social support and new work opportunities (even part time)—and mental health prevention programmes seem to mitigate significantly the negative mental health effects of recessions.5, 9 The fact that countries such as Sweden5 have been able to prevent suicide rises despite major recessions reveals opportunities to protect Americans from further risks of suicide during the continued economic downturn.

We declare that we have no conflicts of interest.

BACK TO TOP

## Sharp Increase in US Suicides Blamed on Financial Crisis

Medscape
Caroline Cassels
5 Nov 2012

A sharp increase in US suicide rates is being blamed on the country's current economic crisis.

A new analysis published online November 5 as Correspondence in the *Lancet* shows that after 2007, deaths by suicide quadrupled.

"In the run-up to the US presidential election, President Obama and Mitt Romney are debating how best to spur economic recovery. Missing from this discussion is consideration of how to protect Americans' health during these hard times," lead author Aaron Reeves, PhD, of the University of Cambridge in the United Kingdom, said in a statement.

The authors note that this research supports previous data from Europe, where it is estimated that the economic recession is responsible for "1000 excess deaths in the UK alone."

They add that among the worst affected economies in Europe, such as Greece, suicides have risen by more than 60% since 2007.

"Suicide is a rare outcome of mental illness, but this means that these data are likely the most visible indicator of major depression and anxiety disorders among people living through the financial crisis, as revealed by recent research in Spain and Greece," Dr. Reeves added.

However, the investigators note that there has been "little or no analysis of US mental health data."

For the study, the researchers analyzed suicide and mortality data from the Centers for Disease Control and Prevention. They also analyzed unemployment data from the Bureau of Labor Statistics.

Time-trend regression models were used to assess excess suicides over and above the level that would be expected if historical suicide trends continued.

The investigators found that between 1999 to 2008, the annual suicide rate was rising by 0.12 deaths per 100,000 population. However, coinciding with the onset of the recession, the suicide rate accelerated, and there was an additional 0.51 deaths by suicide per 100,000 population from 2008 to 2010. The researchers report that this increase corresponds to an additional 1580 suicides per year.

"Thus, during the recessionary period after 2007, there were an estimated 4750 excess suicide deaths."

Next, they examined the relationship between rising unemployment and suicide mortality rates and found that a 1 percentage point rise in unemployment was associated with a 0.99% increase in the suicide rate.

During 2007-2010, the US unemployment rate increased from 5.8% to 9.6%. The researchers report that the rise in unemployment was associated with a 3.8% increase in the suicide rate, corresponding to approximately 1330 suicides.

"In other words, rising unemployment could account for about a quarter of the excess suicides noted in the USA during this time," they write.

The investigators note that it is possible to avoid increased rates of suicide in tough economic times. They note that some countries, such as Sweden, have avoided an increase despite a major economic downturn.

"Active labour market programmes — projects that immediately help the unemployed find social support and new work opportunities (even part time) — and mental health prevention programmes seem to mitigate significantly the negative mental health effects of recessions.

"The fact that countries such as Sweden have been able to prevent suicide rises despite major recessions reveals opportunities to protect Americans from further risk of suicide during the continued economic downturn," the authors write.

BACK TO TOP

***Technology***

BACK TO TOP

***Traumatic Brain Injury***

## Herbicide Exposure/TBI Combination Triples Parkinson's Risk

Medscape
Pauline Anderson
13 Nov 2012

Both a traumatic brain injury and exposure to the herbicide paraquat increase the risks of developing Parkinson's disease (PD), but the combination of the 2 triples the risk, a new study suggests.

The study is important because it shows what epidemiologists have known for years: that risk factors should be looked at in combination, and in vulnerable populations, study author Beate Ritz, MD, PhD, chair, Department of Epidemiology, Center for Occupational and Environmental Health, Fielding School of Public Health, University of California at Los Angeles, told Medscape Medical News.

"For example, someone who has already suffered a head injury might need to stay away from the next neuro toxic agent more than his neighbor does" because that head injury may have made him more vulnerable to an additional insult, said Dr. Ritz.

The study was published in the November 13 issue of Neurology.

Widely Used Herbicide

The study included 357 patients diagnosed with incident idiopathic PD within the previous 3 years and 754 controls, living in 3 mostly rural agricultural counties of California from 2002 to 2011.

From medical history interviews, researchers determined whether participants had ever had a TBI, defined as a head injury with loss of consciousness for more than 5 minutes. They documented the age at the time of the event and whether the participant was hospitalized for the injury.

Investigators also calculated paraquat exposure of each participant up to 1999 (before PD was diagnosed in most patients) using a sophisticated geographic information system that takes advantage of California pesticide records.

From addresses and amounts of pesticide applied per acre (within 500 meters of a workplace or home), researchers estimated an average study period exposure. Participants were considered exposed to the pesticide if they had an average study period exposure greater than 0 at both work and at home.

After adjusting for age, sex, smoking status, race, county, and educational level, the study found a 100% increased risk of developing PD among those who had experienced a TBI (adjusted odds ratio [aOR], 2.00; 95% confidence interval [CI],1.28 - 3.14). The effect estimate was stronger for women (aOR, 2.61; 95% CI,1.32 - 5.16) than for men (aOR, 1.71; 95% CI, 0.94 - 3.11).

TBI induces an inflammatory cascade and accumulation of α-synuclein and tau, 2 proteins that are major components of the hallmark PD Lewy bodies. TBI may also contribute to PD by disrupting the blood-brain barrier and mitochondrial function.

Up to 41% of the study population had been exposed to paraquat. The study found that exposed participants had close to a 40% higher chance of developing PD than those not exposed (aOR, 1.36; 95% CI, 1.02 - 1.81).

In animals, paraquat, one of the most widely used herbicides in the world, has been shown to cause dopamine neurons to die, said Dr. Ritz.

More Than Additive

Adding the risks from the 2 exposures resulted in more than an additive risk. Study participants who had experienced a TBI and been exposed to paraquat had a 3-fold increased risk of having PD compared with those who had never experienced a TBI or been exposed to the pesticide (aOR, 3.01; 95% CI, 1.51 - 6.01).

The physiologic process triggered by a head injury may increase the vulnerability of neurons to insults from neurotoxic pesticides, with the combination increasing the risk for PD more than each exposure on its own, said Dr. Ritz.

As she explained, we are all born with a certain number of dopamine neurons and have to lose 60% to 80% of these neurons before we become parkinsonian. A TBI may increase the vulnerability of dopaminergic neurons to additional insults because the neurons that are left have to work overtime to compensate for the earlier insult, said Dr. Ritz.

"When you have one thing happen and a certain number of neurons die, the other neurons may become more vulnerable to the next hit. You could imagine this biologically because those neurons have to work harder; have to do their dead comrades' job."

Or, in the case of long-term pesticide exposure, the dopamine neurons are continually trying to "defend themselves against injury," a situation that causes the neurons to be constantly stressed and more vulnerable to attack, said Dr. Ritz.

Confirms Animal Research

The current study results mimic those produced in laboratory animals and confirm the researchers' previous findings but with a larger control group and inclusion of exposures both at residences and workplaces. They previously reported a 2- to 3-fold increase in risk of developing PD with exposure to specific types or classes of pesticides, especially for combined exposure to paraquat and maneb, and for persons who carry genetic polymorphisms in susceptibility genes.

A limitation of the study was its inability to address temporality of TBI and paraquat exposure because pesticide records were available only after 1974. Some study participants had experienced TBI at a young age, when pesticide exposures were not yet recorded in the California system. (In animal models, the 2 exposures had to be within a certain time frame, said Dr. Ritz.)

The results may not apply to regions with low pesticide exposure. And because TBI information was collected through interviews, a bias may have been introduced as a result of differential recall.

PD, which affects 1% to 2% of the population older than age 65 years, is the second most common neurodegenerative disorder after Alzheimer's disease.

Two-hit Scenarios Common

Commenting on the findings for Medscape Medical News, Anna DePold Hohler, MD, associate professor of neurology, Boston University School of Medicine, Massachusetts, and member of the American Academy of Neurology, said the results were interesting but not surprising.

"Two-hit scenarios are common in neurological disease," she said.

Although cumulative injury and the length and quantity of parquet exposure may increase the risk of developing PD, the underlying genetic predisposition to PD may also be a factor in determining who will go on to develop the disease, said Dr. Hohler.

The study findings "provide a source of inspiration" for future research, she added. They suggest, for example, that it may be possible to track high-risk individuals for longitudinal study and biomarker analysis to better understand the earliest phases of PD, said Dr. Hohler. "This knowledge can be leveraged into the neuroprotective treatments of the future."

Studies that address the etiology of PD are particularly useful in understanding the mechanism of disease in order to develop a cure, she said.

This work was supported by the National Institute of Environmental Health Science and the National Institute of Neurological Disorders and Stroke; in addition, initial pilot funding was provided from the National Institutes of Health and a pilot grant by the American Parkinson Disease Association. Dr. Ritz and Dr. Hohler have disclosed no relevant financial relationships.

BACK TO TOP

## The Influence of Sleep and Mood on Cognitive Functioning Among Veterans Being Evaluated for Mild Traumatic Brain Injury

Military Medicine
Brigid Waldron-Perrine, PhD; Adam P. McGuire, BS; Robert J. Spencer, PhD; Lauren L. Drag, PhD; Percival H. Pangilinan, MD; Linas A. Bieliauskas, PhD
November 2012

ABSTRACT

Objective: Veterans undergoing evaluation for mild traumatic brain injury commonly report insomnia, psychiatric symptoms, and cognitive dysfunction. This study examines the effects of self-reported amount of sleep and subjective sleep quality on neuropsychological test performance. Method: 262 veterans were seen for neuropsychological assessment in a Veterans Affairs traumatic brain injury clinic. All participants completed measures of depression, anxiety, and sleep satisfaction, and also estimated the number of hours they slept the night before the assessment. Factor scores of attention/concentration and memory were created using factor analyses. Data were analyzed with linear regression. Results: Depression and anxiety were significantly correlated with sleep satisfaction and predictive of cognitive ability. Both sleep satisfaction and hours slept were significantly correlated with memory, but not attention. After controlling for the effects of depression and anxiety, hours slept but not sleep satisfaction was predictive of memory test performance. Conclusions: Perceived sleep quality is heavily influenced by psychiatric symptoms; therefore, veterans' report of sleep satisfaction may merely reflect their overall level of distress. Sleep quantity, however, appears to uniquely contribute to memory performance. Thus, assessment of sleep is important and provides clinicians with useful information, especially among individuals with psychiatric comorbidities.

## INTRODUCTION

Research shows associations among sleep duration, cognition, and mental health with the consistent finding that poor sleep is associated with both negative affect and less efficient cognitive functioning. Because sleep dysfunction is also closely related to mood, another important but less extensively investigated issue concerns the degree to which there are significant cognitive effects of poor sleep beyond what can be expected from depression and anxiety alone.

### Sleep and Cognition

Although findings vary with regard to which cognitive processes are directly affected by poor sleep, in general, those individuals who report poor sleep quality and/or quantity display poorer cognitive performances compared to those describing an adequate amount of sleep. Poor sleep influences various aspects of cognitive functioning, including decreased concentration, poor memory, and protracted processing speed. Inadequate sleep may diminish alertness and attention, which in turn affects many other cognitive abilities such as processing speed, memory, and executive functioning.

### Sleep and Mood

Sleep disturbance is often comorbid with mood and psychiatric illness and may be a cause or consequence of distress. Individuals who report highly distressed sleep also report greater tension and depression than do those with nondistressed sleep. Poor sleep

quality is also strongly associated with lower levels of positive affect. Sleep problems are part of the diagnostic criteria for mental disorders such as major depressive disorder (MDD) and post-traumatic stress disorder (PTSD). Common symptoms associated with MDD or anxiety disorders include insomnia, hypersomnia, difficulty falling asleep, and difficulty staying asleep. In general, sleep disturbances are consistently related to psychiatric symptoms, but the directionality of the relationship is as yet unclear.

The Confound: Mood, Sleep, and Cognitive Functioning

Distress has a strong influence on cognitive functioning, regardless of sleep. Depression and anxiety are consistently associated with decreased attention, poorer memory, and executive dysfunction. Importantly, however, research in this regard typically lacks adequate methodological or statistical controls for confounding variables such as poor sleep.

Given the demonstrated association between mood dysfunction and cognitive impairment, clinicians should consider the degree to which mood symptoms influence the relationship between sleep and cognitive functioning. Currently, the potentially complex relationship among these constructs is not well understood and has been the subject of few neuropsychological investigations.

Measuring Sleep in Neuropsychological Assessment

Performance on neuropsychological tests are typically adjusted for potentially modifying influences (e.g., age, sex, and education), formally or informally, when interpreting test data. Some factors known to influence cognitive functioning, such as psychiatric problems, are routinely assessed in neuropsychological evaluations via questionnaires and interview. Conclusions about cognitive functioning are thus based on objective test results as well as knowledge of the psychiatric and emotional context.

Recognizing that sleep may negatively affect cognitive functioning, clinicians should routinely measure aspects of sleep in order to understand the degree to which sleep may influence neuropsychological test results. Although objective methods (e.g., polysomnography, actigraph, and sleep diaries) are useful for measuring of sleep quantity and quality, these are not feasible in routine clinical practice. More practical alternatives include self-report questionnaires measuring sleep satisfaction or estimates of sleep quantity over a specified time frame. The Insomnia Severity Index (ISI) is a brief but well-validated measure of insomnia impact with demonstrated internal consistency (Cronbach's $\alpha$ = 0.76 at baseline and 0.78 at follow-up) and concurrent and construct validity. The authors of this validity study note that factor analysis using principal components analysis yielded three related components of the measure: impact, severity, and satisfaction. Although prone to self-report bias, simply asking patients to quantify their sleep and report their level of sleep satisfaction can be quickly accomplished during a standard clinical evaluation using this instrument.

Multidomain Neuropsychological Assessment in a Veteran Population

Adequate assessment of sleep and knowledge about the relationships among sleep, psychiatric functioning, and cognition allows clinicians to interpret test data more accurately. For this reason, the combined influence and unique contributions that disordered sleep and psychiatric symptoms have on cognitive functioning needs to be more fully understood. This is especially important in returning Veterans being evaluated for mild traumatic brain injury (mTBI), a population in which both sleep disturbance and psychiatric symptoms are prevalent.

mTBI comprises approximately 80% of all TBIs, and symptoms of mTBI are generally considered to be reversible. Although most individuals fully recover, others report lingering symptoms, such as physical problems (including sleep dysfunction), cognitive deficits, and behavioral changes often referred to collectively as postconcussive syndrome (PCS). Notably, PCS symptoms overlap considerably with psychiatric symptoms and may not directly reflect brain injury sequelae.

Sleep dysfunction is nearly ubiquitous among Veterans referred for assessment of mTBI, with 94% reporting some degree of impairment. Therefore, understanding the relationships between sleep, mood, and cognition is imperative to fully conceptualizing a patient's unique constellation of challenges and providing appropriate treatment recommendations.

This study examined the influence of sleep and psychiatric functioning on objective and subjective cognitive functioning in a sample of Veterans evaluated for mTBI. Previous research has found self-reported sleep quality to be modestly associated with cognitive performance but more strongly associated with subjective cognitive perception. Thus, both subjective report of cognitive symptoms and objective cognitive test performance were examined as both are relevant when drawing conclusions about the influence of sleep on cognition.

In this study, sleep was measured by subjective ratings of sleep quality (i.e., self-report ratings of sleep satisfaction) and by the patient's estimation of the amount of sleep from the night before testing (i.e., sleep quantity). We hypothesized that both sleep quality and quantity would have a significant effect on cognitive abilities. The primary aim of this study was to establish the extent to which sleep quality and quantity exert unique effects (i.e., not accounted for by mood) on cognitive functioning or subjective experience of cognitive efficiency. Additionally, we investigated whether self-report of sleep functioning is a useful method that allows for such influences to be detected.

## METHOD

### Participants

Participants were 452 Veterans referred for neuropsychological assessment within a polytrauma/TBI clinic at a Midwest Veterans Affairs (VA) as part of standard clinical care. Neuropsychological evaluations were conducted by trained neuropsychological technicians or a licensed psychologist with experience in neuropsychological assessment. Of note, this sample is also the basis for Spencer et al and Drag et al. Participants were excluded (*N* = 47) if they had conditions assumed to produce significant cognitive impairments, such as non-TBI neurologic disease or moderate or severe traumatic brain injury (defined by greater than 30 minutes loss of consciousness [LOC] or 24 hours of post-traumatic amnesia [PTA] or disorientation). To ensure the validity of cognitive testing results, participants were also excluded if they failed a commonly used embedded measure of effort towards cognitive tasks (Reliable Digit Span [RDS]). To maximize the validity of our data and minimize measurement error, we then further excluded individuals if they failed a separate stand-alone measure of effort (Test of Memory Malingering [TOMM]), 125 Veterans demonstrated suboptimal effort on RDS (28%); of those who passed RDS and were administered the TOMM on the basis of clinical judgment, 18 displayed suboptimal effort. Thus, 143 of the 452 Veterans screened (31.6%) displayed suboptimal effort on one or more of these commonly used measures. The finding is generally consistent with recent estimates that 17 to 58% of Veterans referred for similar evaluations fail measures more sensitive to effort than to neurological insult. Data from the remaining 262 participants (95.8% men) were included in the analyses.

Measures

Sleep Satisfaction

Each patient was administered the ISI, a 7-item questionnaire which uses a five-point scale (0 [none]–4 [very severe]) to identify insomnia. Three questions pertain to problems with falling asleep, staying sleep, and waking up too early. The remaining four questions pertain to the impact that the individual's sleep pattern has on aspects of their quality of life. Items include “how satisfied are you with your sleep,” “to what extent do sleep problems interfere with your daily functioning,” “how noticeable is this impairment to others” and “how worried are you about your sleep pattern.” Scores of 15 and above reflect significant sleep impairment. Cronbach's $\alpha$ was 0.90 in the present sample. Although not part of the ISI, patients were also asked to quantify the amount they slept the night before.

Anxiety and Depression

The Hospital Anxiety and Depression Scale (HADS) is a 14-item questionnaire, with seven questions that reflect symptoms of anxiety and seven that reflect symptoms of depression. Participants answered each item by selecting a categorical response reflective of severity of the specific symptom on a four-point scale (0–3). Total scores were calculated for anxiety and depression, with scores of 8

to 10 indicating mild distress, 11 to 14 moderate distress, and 15 and above severe distress. Cronbach's $\alpha$ was 0.84 for anxiety and 0.83 for depression in the present sample.

Neuropsychological Assessment

Shipley Institute of Living Scale-Vocabulary Test

The Shipley Institute of Living Scale-Vocabulary subtest is composed of 40 target words, to which participants indicate the word most similar in meaning from among four alternatives. General cognitive ability before injury has potential to contribute to memory and attention functioning postinjury and should be accounted for in analyses. Vocabulary tasks and other such measures of crystallized intelligence are known to be predictive of overall intelligence quotient and robust to neurological damage. Such measures are therefore commonly used in neuropsychological assessments as estimates of baseline cognitive functioning/premorbid intelligence.

Injury Characteristics

A structured interview following the VA Comprehensive TBI Evaluation template was used to collect information about injury characteristics. Each veteran was asked in an open-ended interview format to report on time spent unconscious or disorientated/confused. Targeted follow-up questions were asked as needed, based on clinical judgment and need for clarification. Data on injury characteristic were recorded categorically (e.g., <1 minute or 1–30 minutes for LOC and <30 minutes or 30 minutes to 24 hours for disorientation and PTA) as required by the VA system.

Wechsler Adult Intelligence Scale-IV Digit Span

Wechsler Adult Intelligence Scale-IV (WAIS-IV) Digit Span is a test of auditory attention. Participants were given sets of digits to repeat forward, backward, or to sequence, with increasing lengths of digit strings as each series progresses. The total of correctly repeated sequences for all three trials was used to derive a scaled score based on normative data (mean = 10; SD = 3). RDS was calculated based on the sum of the longest span of digits consistently correct within a set for forward plus backward administrations.

Trail Making Test

The Trail Making Test is a test of visuospatial attention and processing speed. In part 1 (Trails A), patients connect numbered circles in ascending numerical order as quickly as possible. In part 2 (Trails B), patients connect circles in ascending order, alternating sequentially between letters and numbers. This task requires both attention and intact executive function. Total time to complete

each task was recorded, and T-score conversions (mean = 50; SD = 10) were made using normative data. The Trails B test has been shown to be among the most sensitive neuropsychological tests to impairment in mTBI.

Repeatable Battery for the Assessment of Neuropsychological Status (RBANS): Story Memory Subtest

The RBANS Story Memory is test of memory for prose. There are two learning trials (i.e., immediate recall) and a delayed recall after 20 minutes. Scaled scores are derived from normative data and based on the total of the two immediate recall trials and the delayed recall trial.

Rey Complex Figure Recall

The Rey Complex Figure Recall (RCFR) is derived from the Rey Complex Figure Test. In this test, patients are asked to copy and then subsequently recall a geometrically complex figure. Following a 3-minute delay, patients redraw the figure from memory (RCFR), which is scored using a 36-point scale.

Neurobehavioral Symptom Inventory

The neurobehavioral symptom inventory is a 22-item questionnaire of postconcussive symptoms. Patients are asked to rate their perceived severity of disturbance caused by each symptom in the past 2 weeks from 0 (no problems) to 4 (very severe problems). Ratings on items pertaining to memory, concentration, and processing speed were summed to form an overall cognitive symptom score (Cronbach's $\alpha$ for these items = 0.86).

Statistical Analysis

To minimize the number of variables used in the analyses and to simultaneously increase power, separate linear composite scores were created for memory and attention tests. Separate factor analyses were performed in which the memory tests were included in one factor analysis and attention tests were included in a second factor analysis. For each factor analysis, regressed factor scores were saved and used in subsequent analyses. Digit Span and Trails A & B, tests purported to measure aspects of attention, were included in an Attention Factor, and RBANS Story Memory and RCFR, tests purported to measure aspects of memory, were included in a Memory Factor.

Hierarchical multiple regression analyses conducted using SPSS determined the proportion of the variance in three cognitive outcomes (global memory performance, global attention performance, and self-report of cognitive deficit) accounted for by demographic, injury-related, sleep, and mood variables. Demographic and injury-related characteristics were entered into each

model first to account for their influence on cognitive outcome. Thus, for each outcome, age and estimated premorbid ability (Shipley Vocabulary T-score) were entered on step 1. Perceived sleep quality (ISI total) was entered in step 2, sleep quantity (hours slept) was entered in step 3, and mood symptoms (HADS anxiety and depression scores) were entered in step 4.

Before analysis, all variables were examined for accuracy of data entry and violations of the assumptions of multivariate analysis. Inaccuracies were corrected. No extreme outliers were detected and no transformations were required. Tolerance statistics did not yield evidence of problematic multicollinearity. Given the relatively large sample size, we used listwise deletion of missing data in the individual analyses to maximize the validity of the original data.

## RESULTS

### Descriptive Statistics of Demographic, Injury, and Cognitive Variables

Participant's responses on the ISI show a mean of 16.12 total score (SD = 6.43), which is above the 15-point cutoff indicative of probable sleep disturbance; 55% of the sample reported sleep disturbance in this clinical range. Average scores on the HADS were 11.56 (SD = 4.40) for depression symptoms and 8.42 (SD = 4.39) for anxiety symptoms. In terms of symptoms of anxiety, 20% of the participants reported minimal symptoms, 22% mild symptoms, 29% moderate symptoms, and 29% severe symptoms. In terms of symptoms of depression, 44% reported minimal symptoms, 27% reported mild symptoms, 23% reported moderate symptoms, and 6% reported severe symptoms.

Pearson product–moment correlation coefficients were calculated to examine the associations among the sleep, mood, and cognitive variables presents the correlations among all variables and demonstrates multiple relationships of statistical significance. Perceived poor sleep quality showed a moderate correlation with self-report of cognitive deficit, $r$ (223) = 0.43, $p < 0.001$, and a large association with depression, $r$ (241) = 0.54, $p < 0.001$, and anxiety symptoms, $r$ (240) = 0.54, $p < 0.001$, as measured by the HADS. Additionally, perceived poor sleep quality was negatively correlated with performance on neuropsychological assessments measuring memory, $r$ (225) = −0.20, $p < 0.001$, but not attention, $r$ (229) = −0.06; $p = 0.18$, capabilities.

Sleep quantity (hours slept) showed a small but significant positive association with memory test performance, $r$ (223) = 0.23, $p < 0.001$. It also demonstrated a small but significant negative correlation with self-report of cognitive deficits, $r$ (222) = −0.17, $p < 0.01$, and small to moderate correlations with depression symptoms, $r$ (240) = −0.22, $p < 0.001$, anxiety symptoms, $r$ (239) = −0.27, $p < 0.001$, and perceived sleep quality, $r$ (239) = −0.53, $p < 0.001$.

Several significant correlations were also found between patients' self-reported cognitive complaints and the sleep and distress variables. Self-report of cognitive problems was associated with depression, $r$ (234) = 0.59, $p < 0.001$, and anxiety, $r$ (233) = 0.56, $p < 0.001$, symptoms. Self-report of cognitive problems also demonstrated a negative correlation with memory test performance, $r$ (208) = −0.26, $p < 0.001$, but not attention, $r$ (212) = −0.06, $p = 0.21$.

Significant relationships were found between memory test performance and depression, $r$ (225) = −0.26, $p < 0.001$, and anxiety, $r$ (225) = −0.26, $p < 0.001$, symptoms. A correlation was also found between attention test performance and depression, $r$ (229) = −0.18, $p < 0.05$, and anxiety, $r$ (228) = −0.11, $p < 0.05$, symptoms.

Prediction of Memory

Hierarchical Multiple Regression Analysis: Memory Factor ($N$ = 213)

In the prediction of memory functioning, all individual steps were significant ($p < 0.05$), as was the total model ($R^2 = 0.13$, $F$ [6, 207] = 5.22, $p < 0.001$). Although perceived sleep quality was initially significant when added to the model ($t$ [210] = −2.58, $p < 0.05$), in the presence of the addition of number of hours slept, this variable was no longer uniquely significant ($t$ [210] = −1.20, $p = 0.23$). Notably, evaluation of the unique contributions of each variable (semipartial correlations) indicates that in the context of adding anxiety and depression symptoms to the model, number of hours slept last night continued to yield unique power in the prediction of objective memory test performance ($t$ [210] = 2.07, $p < 0.05$). Depression ($t$ [210] = −2.02, $p < 0.05$) also had unique predictive power.

Prediction of Attention

In the prediction of overall attention, steps including demographic characteristics ($R^2$ change = 0.05, $F$ [2, 214] = 6.14, $p < 0.01$) and psychiatric distress ($R^2$ change = 0.05, $F$ [2, 210] = 5.45, $p < 0.01$) were significant, but the additions of perceived sleep quality ($R^2$ change = 0.00, $F$ [1, 213] = 0.15, $p = 0.70$) and sleep quantity ($R^2$ change = 0.00, $F$ [1, 212] = 0.03, $p = 0.86$) did not add significantly to the predictive power. The total model accounted for only 10% of the variance in attention functioning ($R^2 = 0.10$, $F$ [6, 210] = 3.96, $p < 0.001$). Evaluation of the unique contributions of each variable (semipartial correlations) indicates that only estimated premorbid intellect ($t$ [210] = 3.10, $p < 0.01$) and mood symptoms ($t$ [210] = −2.47, $p < 0.05$) uniquely predicted overall attention.

Prediction of Self-Reported Cognitive Problems

In the prediction of self-reported cognitive problems, the total model accounted for 46% of perceived cognitive deficit ($R^2 = 0.46$, $F$ [6, 205] = 28.60, $p < 0.001$). Although perceived sleep quality was initially significant when added to the model ($t$ [205] = 7.02, $p < 0.001$), with the addition of actual hours slept and anxiety and depression symptoms, this variable was no longer uniquely predictive

($t$ [205] = 1.56, $p$ = 0.12). Anxiety ($t$ [205] = 5.08, $p < 0.001$, $sr^2 = 0.07$) and depression ($t$ [205] = 5.19, $p < 0.001$, $sr^2 = 0.07$) each accounted for significant unique variance in self-report of cognitive problems.

## DISCUSSION

This study examined relationships between subjective sleep quality and quantity, cognitive functioning, and psychiatric symptomatology. Consistent with other studies, depression and anxiety were prominent in this veteran population being assessed for mTBI, with mean scores falling in the mild-moderate range. Similar to other investigations, we found that both sleep and psychiatric symptoms were individually predictive of cognitive test performance and self-report of cognitive problems. However, when the impacts of sleep and distress on cognitive functioning were examined simultaneously, the unique influence of sleep-related variables on cognitive functioning was substantially reduced. Once mood and anxiety symptoms were accounted for, the number of hours slept was uniquely predictive of memory but not attention or self-report of cognitive symptoms. Subjective sleep quality was not uniquely predictive of any cognitive outcome beyond the effects of mood and anxiety. Thus, the personal experience of poor sleep satisfaction may be reflective of general distress.

Not surprisingly, an estimate of general intellectual ability was predictive of current cognitive functioning. Moreover, symptoms of depression and anxiety were uniquely predictive of self-reported cognitive complaint, which is likely at least in part because of the fact that both are self-report measures of distress (see discussion of shared method variance in “limitations” below). Symptoms of both depression and anxiety were also predictive of objectively measured memory performance whereas depression, but not anxiety, negatively impacted objectively measured attention. This finding is contrary to the hypothesis proposed by Killgore that attention is the underlying cognitive process affected by poor sleep that leads to other cognitive deficits. However, our findings that the influence of mood on cognition is as strong as or stronger than the influence of sleep are consistent with Killgore's assertion that lack of sleep may affect cognitive functioning through disruption of systems responsible for emotional regulation.

Although this study establishes the unique and combined effects of sleep and emotional distress on cognition, sleep disturbance is a symptom in many diagnosed mood disorders and a large body of literature has supported the relationship between the physiological components of sleep and mood disorders and between subjective sleep quality and affect. Although the findings of this study support the notion that sleep disturbance may be a result of psychiatric problems and that patient's satisfaction with quality of sleep is influenced by mood symptoms, it is the reduced quantity of reported sleep which appears to affect aspects of cognitive functioning.

### Clinical Implications

Our findings, supported by previous literature, clearly demonstrate that both sleep and mood influence cognition. Although the influence of sleep on cognitive functioning (specifically memory) was small, its significance underscores the need to account for this potential influence in neuropsychological evaluations. Clinicians should be continuously aware of the myriad ways, direct and indirect, in which psychiatric problems may influence cognitive functioning.

Assessment of sleep may help identify additional treatment opportunities. Insomnia is often successfully treated using cognitive behavioral and/or pharmacological methods, and if adequately treated, could result in a corresponding improvement in both objective cognitive problems and self-reported cognitive deficits. As demonstrated by this study, simply asking the patients how many hours they slept last night is a low technological method that has demonstrated validity as a measure of sleep quantity.

In summary, sleep quantity does have a small but significant effect on memory performance and should be quantified during neuropsychological assessments. Although sleep satisfaction was not found to be uniquely predictive of neuropsychological test performance, this does not suggest that it is unimportant or that it should not be measured. Similar to other subjective measures (such as self-reported cognitive complaints and psychiatric symptoms), perceived sleep quality can contribute to the clinician's understanding of the patient's current state from an emotional perspective.

Limitations

In general, subjective reports are but one method of assessing patients' capabilities or experiences, and report of distress on one self-report measure is likely to be associated with report of distress on others (i.e., shared method variance). Because the ISI measures sleep quality as perceived by the patient, it may not be consistent with sleep quality as measured by other objective methods (e.g., physiologically based sleep study). Even report of sleep quantity is subjective, and may be subject to biases. Similarly, subjective experience of cognitive symptoms does not necessarily mean that an individual has objectively defined attention or memory problems. In fact, in the present study and consistent with previous research findings from our population, patients' perception of their own cognitive impairment was not closely related to their objective cognitive functioning. Thus, although subjective perception of dysfunction is important, in general, objective corroboration of self-report is ideal. Use of more technologically advanced measures, such as actigraphs would improve our ability to understand the true impact of objectively measured sleep quality on cognition and its relationship with emotional distress.

Participants in the present study were typically young, male Veterans with sleep and mood problems undergoing evaluation for TBI. Whether the findings of the present study apply to other populations will need to be established in the future to fully understand the interplay of mood, sleep and cognitive factors across populations.

Future Research

Further research using experimental study designs and objective measures of sleep quality is needed to confirm and extend the present findings. Additional research is also needed to examine the effects of poor sleep on attention. It is possible that no effects were observed in this study because of the tests' low difficulty level, but higher level cognition may be particularly vulnerable to sleep disturbance. Research is also needed to examine the cognitive-enhancing potential of sleep interventions.

General cognitive ability before injury has potential to contribute to memory and attention functioning postinjury. Additional research is needed to examine the influence of other factors that may influence sleep, mood, and cognitive functioning. Given the probable complex relationships among pain, PTSD, substance use, sleep, and cognition, it is likely that pain, PTSD, and substance use also influence the relationship between sleep and cognitive functioning as seen in neuropsychological testing. Thus, these constructs and their relationships to one another should be the focus of future investigations.

BACK TO TOP

***Medical Protocol and Training***

## Mobile Learning Module Improves Knowledge of Medical Shock for Forward Surgical Team Members

Military Medicine
Carl I. Schulman, MD, PhD; LTC George D. Garcia, MC USA; Mary M. Wyckoff, PhD; Robert C. Duncan, PhD; Kelly F. Withum, BS; Jill Graygo, MA, MPH
November 2012

ABSTRACT

Objective: Acute trauma care is characterized by dynamic situations that require adequate preparation to ensure success for military health professionals. The use of mobile learning in this environment can provide a solution that standardizes education and replaces

traditional didactic lectures. Methods: A comparative evaluation with a pre–post test design regarding medical shock was delivered via either a didactic lecture or a mobile learning video module to U.S. Army Forward Surgical Team (FST) members. Participants completed a pretest, were randomly assigned to treatment group by FST, and then completed the post-test and scenario assessment. Results: One-hundred and thirteen FST members participated with 53 in the mobile learning group and 60 in the lecture group (control). The percent mean score for the mobile learning group increased from 43.6 to 70 from pretest to post-test, with a scenario mean score of *M* = 56.2. The percent mean score for the control group increased from 41.5 to 72.5, with a scenario mean score of *M* = 59.7. The two-way analysis of variance mean score difference was 26.4 for the mobile learning group and 31.0 for the control, *F* = 2.18, (*p* = 0.14). Conclusions: Mobile learning modules, coupled with a structured assessment, have the potential to improve educational experiences in civilian and military settings.

## INTRODUCTION

Mobile learning presents unique opportunities and challenges in a variety of health care settings, including hospitals, field response, and austere environments. The use of mobile learning for trauma can provide a solution that standardizes education, replaces traditional didactic lectures, and facilitates “just-in-time” communication at the point-of-care. Residents in surgical specialties are required to perform a rotation in a trauma and critical care department in order to garner hands-on trauma patient treatment and management experience for board certification as general surgeons.

Since the educational needs of the trainee are subordinate to the needs of the patient in crisis-oriented emergency and critical care settings, it is difficult for educators to facilitate learning and for residents to gain hands-on experience. Additionally, with the growing trend of minimally invasive procedures, opportunities for clinical learning and practice have decreased by as much as 30%. Researchers found that 64% of the physician's informational needs were not being met during their teaching rounds. These limitations have challenged educators to find innovative solutions that seek to overcome limited faculty resources and time constraints while also improving the quality of medical education as a whole.

E-learning in general and mobile learning in particular offer models of learning through which caregivers in chaotic and austere environments can have ongoing access to informational resources, especially at the point-of-care. E-learning refers to the general use of electronic or World Wide Web-based technology to deliver an array of solutions that enhance knowledge and performance. Mobile learning is a subtype of e-learning that uses personal digital assistants to bring the latest information to the point-of-care, with or without Internet access. These comprehensive technologies target both knowledge delivery and learning enhancement in order to build knowledge and skills. This is especially promising when these technological advancements utilize both multimedia instructional

methods and content. Mobile learning allows students to access the information according to their own schedules and provides additional opportunities to review material as needed.

Researchers have recently begun to assess the utility of new education technologies. Some studies suggest that integrating e-learning technologies into an interprofessional health science courses did not improve outcomes as compared to traditional face-to-face group meeting. However, several studies have found that learning improvements, resulting from new educational platforms (i.e., technological advancements, mobile learning, and e-learning), are equivalent to those resulting from traditional lecture formats. One study found that medical students' performance on a urology knowledge acquisition examination, following a computer aided learning software program and after a standard lecture format, was statistically similar. The focus of this evaluation demonstrates statistical equivalence between mobile- and lecture-based learning in military trauma settings.

## METHODS

U.S. Army Forward Surgical Team (FST) members rotating through the University of Miami/Ryder Trauma Center at the Army Trauma Training Center (ATTC) from August 2010 to March 2011 were invited to participate in the evaluation. One FST per month trains at the ATTC for a duration of 2 weeks before their international deployment. Consenting participants were randomly assigned to either the control group, which consisted of the traditional didactic lecture on medical shock, or the mobile learning group, which viewed the shock mobile learning lecture on an iPod Touch (Apple, Cupertino, California). The participants were randomly assigned to the mobile learning or control group as an entire FST to prevent video information sharing while they were living in such close quarters.

Department of Surgery faculty and trauma fellows created the module content, which was then refined and edited by our Director of Surgical Education. The preproduction final product includes a PowerPoint (Microsoft, Redmond, Washington) presentation with an associated script. The production company uses these components, along with a narrated voice file of the script, to assure proper medical pronunciation. The production company then works closely with the Director of Surgical Education to create a high-quality multimedia presentation complete with professional voice, soundtrack, and live video or animation as necessary. An iterative process and further information specific to particular anatomy or physiology are often required to enable the production company to understand the material and transform it into an engaging learning module.

This evaluation used a pre–post test design to measure subject matter knowledge, in this case the classification and treatment of medical shock. The pretest and post-test was created by an expert panel of trauma and critical care clinicians that used the shock

module script to derive the questions. The pretest and post-test consisted of the same 10 multiple choice questions (Appendix A). The patient case scenario assessment consisted of five multiple choice questions that included two case scenarios with physical descriptions of the patients (Appendix B). All research assessments were pilot-tested with a sample population of physicians.

The pretests were disseminated and collected by a research assistant immediately before participants viewed either the mobile learning module or the actual live lecture. Once all pretests were collected, students participating in the didactic lecture remained in the lecture hall with their peers and viewed the PowerPoint lecture given by a Department of Surgery faculty member. The length of the lecture was approximately 45 minutes. The PowerPoint lecture and script were identical to the mobile learning module, but without the animation or multimedia content. Students in the intervention group were given an iPod Touch with headphones and were instructed to access and view the entire shock module one time. The length of the shock mobile learning module was 10 minutes and 31 seconds. Upon completion, the iPod was collected by the research assistant. Immediately following either the lecture or viewing of the mobile learning module, the research assistant again disseminated and collected the post-test and a patient case scenario assessment that was administered to measure application of knowledge. All data collected were coded and nonidentifiable. Informed consent was obtained from all the participants, and the research study was exempt approved by the Department of Defense and University Institutional Review Boards.

Databases were created using Excel (Microsoft, Redmond, Washington) and data were analyzed using SAS 9.2. Two-way analysis of variance tests were performed that compared the pre- and post-test score differences between the mobile and control groups, significance considered at $p = 0.05$ level. Mean scores are represented in percentages with ±SD.

## RESULTS

A total of 113 FST members from the ATTC participated in the study. There were 53 FST members in the mobile learning group and 60 in the didactic lecture control group. Descriptive statistics and a means comparison were used to analyze the data. The percent mean score (SD) for the mobile learning group increased from 44 ± 23 to 70 ± 20 from pretest to post-test, with a scenario assessment mean score of 56 ± 26. The percent mean score for the control group also increased from 42 ± 21 to 73 ± 18 from pretest to post-test, with a scenario assessment mean score of 60 ± 29. There was only one FST member from the control group that had a decrease in score from pretest to post-test (10%) with no mobile learning group participants decreasing their scores.

A 2-way analysis of variance was performed on the pretest and post-test score differences in order to compare the mobile and control groups. For the mobile group, the mean difference between the pretest and post-test score was 26.4 ± 18.0. For the control

group, the mean difference between the pretest and post-test score was 31.0 ± 18.0. Comparing the mean differences in mobile and control groups resulted in an *F* value of 2.18, which was not significant (*p* = 0.14).

Results Analysis by Position

In the mobile learning group, there were 18 medics, 24 nurses (Registered Nurses [RNs], Licenced Practical Nurses [LPNs], and Certified Registered Nurse Anesthetists [CRNAs]), 1 physician assistant (PA), and 10 technicians (operating room and surgical). The 60 participants in the control didactic lecture group had a similar breakdown of 22 medics, 21 nurses (RNs, LPNs, and CRNAs), 5 PAs, and 12 technicians (operating room and surgical).

The means for pretest and post-test for the 18 medics in the mobile learning group were *M* = 41.7 ± 17.6 and 65.6 ± 19.8, respectively, with a scenario assessment mean score of *M* = 52.2 ± 24. The mean difference between the pretest and post-test score was 23.9 ± 14.2. The means for pretest and post-test for the 22 medics in the didactic lecture control group were *M* = 36.4 ± 16.8 and 65.9 ± 19, respectively, with a scenario assessment mean score of *M* = 56.4 ± 29.4. The mean difference between the pretest and post-test score was 29.5 ± 18.1. Comparing the mean differences in mobile and control groups resulted in an *F* value of 0.98, which was not significant (*p* = 0.32).

The means for pretest and post-test for the 24 nurses in the mobile learning group were *M* = 50.8 ± 25.7 and 78.8 ± 15.4, respectively, with a scenario assessment mean score of *M* = 69.2 ± 22. The mean difference between the pretest and post-test score was 27.9 ± 20.8. The means for pretest and post-test for the 21 nurses in the didactic lecture control group were *M* = 52.4 ± 18.4 and 80.9 ± 14.5, respectively, with a scenario assessment mean score of *M* = 72.4 ± 18.4. The mean difference between the pretest and post-test score was 28.6 ± 17.4. Comparing the mean differences in mobile and control groups resulted in an *F* value of 0.01, which was not statistically significant (*p* = 0.90).

The means for pretest and post-test for the 10 technicians in the mobile learning group were *M* = 25.0 ± 8.5 and 54.0 ± 19, respectively, with a scenario assessment mean score of *M* = 30.0 ± 19.4. The mean difference between the pretest and post-test score was 29.0 ± 17.9. The means for pretest and post-test for the 12 technicians in the didactic lecture control group were *M* = 22.5 ± 10.6 and 62.5 ± 16.6, respectively, with a scenario assessment mean score of *M* = 31.7 ± 24.8. The mean difference between the pretest and post-test score was 40.0 ± 18.0. Comparing the mean differences in mobile and control groups resulted in an *F* value of 2.04, which was not statistically significant (*p* = 0.16).

The mean for pretest and post-test for the one PA in the mobile learning group was *M* = 90.0 and 100, respectively, with a scenario assessment mean score of *M* = 80.0. The mean difference between the pretest and post-test score was 10.0. The means for pretest

and post-test for the five PAs in the didactic lecture control group were $M = 64.0 \pm 16.7$ and $90.0 \pm 0$, respectively, with a scenario assessment mean score of $M = 88.0 \pm 11$. The mean difference between the pretest and post-test score was $26.0 \pm 16.7$. Comparing the mean differences in mobile and control groups resulted in an $F$ value of 0.66, which was not significant ($p = 0.42$).

There were no statistically significant differences found between the mobile and control groups in any of the different study populations. This suggests the two learning modalities are equally effective. Comparison of mean score differences by group and position. There were no statistically significant differences found between mobile and control groups in each of the different study populations. This suggests that the two learning modalities are not different.

## DISCUSSION

The current data suggest that mobile learning modules are equivalent to traditional didactic lectures in trauma and critical care for FST members. Both groups performed better on the post-test than the pretest with a nonsignificant mean score difference. The benefit of mobile learning is found when comparing the time needed for education. Traditionally, this type of education is delivered as a didactic modality, which consumes valuable time that could be spent on hands-on simulation and training activities. For comparison, many traditional didactic lectures are 30 to 60 minutes, whereas the mobile learning module is 10 minutes. The mobile learning modules can be paused and rewatched and do not require an instructor, thus making it an ideal solution for educating the FST members in austere deployment settings.

Trauma and critical care education currently face the constant challenge of increasing time constraints as a result of clinical hour limitations, increased amounts of information required to be retained, and the need for immediate ability to access information in a trauma and critical care situation (“just-in-time learning”). Properly designed mobile learning modules can help to mitigate these significant challenges. In addition, best practices like limiting written text on screen, audio capability, and high-quality graphics are essential to the success of any mobile module.

This prospective study was subject to certain limitations that were primarily methodological in nature. The participants were randomly assigned to the mobile learning or control group as an entire FST to prevent video information sharing while they were living in such close quarters. In addition, our study did not include a knowledge retention test, which would ideally be administered 1 month after intervention, to determine levels of information retention. This follow-up testing was prevented because of the international deployment of the FST after the training. Further data collection would allow researchers to better understand how this type of

learning compares to traditional didactic lectures and how much knowledge is retained over time. Future research will include observational clinical research to evaluate if mobile learning modalities change outcomes in clinical practice applications.

CONCLUSIONS

Although technological advances have allowed for the creation and advancement of multimedia learning modules that can be reliably accessed by mobile technologies, additional evaluation and standardization is necessary. Both validity studies that measure whether multimedia learning modules are actually teaching what they are intended to teach and curricula that address what e-content should be included are still needed. Given the wide range in terms of quality of e-content currently available, knowledge acquisition and learner satisfaction are two key outcomes that must be measured when assessing this type of education.

The goal of this investigation was to determine whether clinical trainees can appropriately use information provided via mobile learning to achieve a higher standard-of-care. The data compiled from this research suggest that FST members learn the medical shock module equally well with either the mobile learning modules or the traditional didactic lecture. This suggests that mobile learning modules are an effective means of providing the same knowledge in about one-quarter the amount of time needed to provide it in traditional didactic lectures. Mobile learning modules, coupled with a structured assessment, have the potential to improve educational experiences in civilian and military settings.

BACK TO TOP

## Beyond Battlemind: Evaluation of a New Mental Health Training Program for Canadian Forces Personnel Participating in Third-Location Decompression

Military Medicine
Zamorski, Mark A.; Guest, Kim; Bailey, Suzanne; Garber, Bryan G.
November 2012

Abstract

Introduction: Battlemind training, which improves postdeployment well-being, has been part of Canada's postdeployment Third-location Decompression (TLD) program since 2006. In 2010, a new educational program drawing on Battlemind was implemented to make it more consistent with Canada's current mental health training strategy. Methods: Subjects consisted of 22,113 Canadian personnel returning from Afghanistan via TLD in Cyprus; 3,024 (14%) received the new program. Pre-/post-training attitude and self-efficacy questionnaires assessed the impact of the training. In addition, a quasi-experimental approach used questionnaires administered at the end of TLD to compare the satisfaction, attitudes, and self-efficacy under the old vs. new program. Results: Pre-/post-training questionnaires showed medium to large positive effects of the training on targeted attitudes and self-efficacy (Cohen's d = 0.44–1.02). Participants completing the new program were more satisfied with the educational program (adjusted odds ratio = 3.2), perceived the TLD to be more valuable (odds ratio = 1.7), and had at least certain more favorable post-TLD attitudes and self-efficacy (d ranging from 0.00 to 0.29). Conclusion: All of these findings point to the superiority of the new program. However, quasi-experimental approaches are bias-prone, and it is unknown whether these advantages will translate into meaningful improvements in well-being.

## Introduction

Military personnel returning from deployment must readapt to their home environment, psychologically, physically, and socially.1 Most are glad to be home, but many find parts of this transition to be at least temporarily distressing.1,2 In addition, an important minority go on to have serious long-term mental health problems.3 Some have argued that difficult early homecoming experiences are an important etiological factor in these postdeployment disorders.4,5

In recognition of the challenges and potential toxicity of the transition home, military organizations have developed policies and programs to make the process of transition easier and to mitigate mental disorders. The two most commonly used approaches are third-location decompression (TLD) programs6 and psychoeducational programs such as the U.S. Army's Battlemind training, which has been shown to improve postdeployment well-being.7 The Battlemind program is a cognitive and skill-based program that reframes transition difficulties as a failure to adapt skills learned in combat to the home environment.

Since 2006, Canadian personnel returning from combat and peace support operations in Kandahar Province, Afghanistan, have completed a 5-day TLD program in Cyprus.8 In addition to having opportunities for rest and recreation, all TLD participants have received a clinician-delivered version of the postdeployment Battlemind training program.7 Participants also attended two 1-hour “elective” presentations on a variety of topics relevant to reintegration. Personnel are strongly supportive of the TLD concept, find the Canadian TLD program to be valuable, and perceive it to help make the transition easier for them.8 About three quarters of participants were satisfied with the TLD educational program, and satisfaction with the program actually increased over the first 4 to 6 months after return from deployment.8About three quarters also felt that the TLD experience helped them realize that there was nothing wrong with getting help with a mental health problem8; this was a central message of the Battlemind program.7

Notwithstanding the apparent success of the TLD program, a variety of factors drove Canada to change its TLD education program:

— Some participants found the Battlemind video used during the training to be “too American” in its themes and settings;
— Participants going through TLD for a second time complained that the training was getting “stale;”
— Canada's career- and deployment-cycle mental health training program was using increasingly divergent training paradigms from Battlemind;
— Canada needed a more flexible postdeployment psychoeducational program that was suitable for noncombat operations;
— Recent research has increased awareness of barriers to care other than stigma that could be targeted in training9;
— The use of electives prevented training from being delivered to cohesive groups, where it was expected to be most effective7; and
— The Canadian Forces (CF) moved to a joint delivery model for its mental health training, in which both a clinician and a nonclinician provide training together.

In response, the CF developed a new 4-hour educational program for TLD. This article presents its results and compares these to those of TLD's using the old educational program. Specifically, it will explore:
— Pre-/postsession changes in self-efficacy and attitudes;
— Differences in satisfaction with the TLD program among those receiving the new vs. old program; and
— Differences in post-TLD self-efficacy and attitudes among those receiving the new vs. old program.

## Methods

### Subjects

The subjects were 22,113 CF members returning from 6- to 9-month deployments in Kandahar Province, Afghanistan, who completed a 5-day TLD program in Cyprus from 2006 to 2010. The old and new programs were received by 19,089 and 3,024 personnel, respectively.

### TLD Program

Canada's TLD program has been described elsewhere.8 In brief, Day 1 consisted of arrival procedures, a short orientation briefing, and individual free time. Days 2 and 3 consisted of educational sessions (either the old or new program) followed by individual free time or group outings. Day 4 featured individual free time or outings, and participants flew home on Day 5. TLD was held in a four-star resort in varying locations in Cyprus.

### Old Educational Program

On Day 2, a version of the Battlemind training program was delivered by a mental health clinician. A video with vignettes of four soldiers experiencing transition or mental health problems was shown, interspersed with didactic material and group discussion on how to recognize transition problems, how to use self and buddy care where appropriate, and how to know when professional care is needed. The video was developed by the U.S. Army's Walter Reed Army Institute for Research and was intended for use 3 to 6

months after return, but it was used during TLD because of its engaging nature. Battlemind sessions lasted approximately 60 minutes and occurred in groups of 30 or fewer personnel. To the extent possible, training occurred within platoons or among personnel who worked together.

On Day 3, participants attended one of between four and seven "electives." Attendance was mandatory, but participants could select the two sessions they preferred. The offerings varied over time, but the most popular offerings were:
— "Coping with Stress and Anger," a clinician-delivered, cognitive-behaviorally oriented anger and stress management program;
— "Healthy Relationships," a clinician-delivered, interpersonal therapy-oriented program intended to help participants prevent, recognize, and solve family conflicts;
— "Post-deployment Reintegration from the Veteran's Perspective," a presentation by veterans who were part of a peer support program for personnel with operation-related mental health problems; and
— "Leadership after the Action," a presentation by clinicians or personnel selection officers on recognizing and supporting those with postdeployment mental health problems.

The electives lasted 60 minutes each and consisted of didactics and group discussion, with a maximum of 30 participants per group. No attempt was made to deliver the electives in cohesive groups. Trainers were not necessarily selected on the basis of their instructional abilities, and they received varying durations of training (between a few hours and 1 week).

New Educational Program
Key differences between the old and new programs are summarized in .
The fundamental goals of the new program remained to ease the process of transition and to facilitate care-seeking for those who need it. These goals were met differently, though, via the following learning objectives: — Identify the difficulties and accomplishments of the mission;
— Understand the physiological effects of stress;
— Appreciate the physiological decompression process;
— Recognize the common transitional phase challenges during reintegration;
— Identify key reintegration strategies in returning home from deployment;
— Identify positive coping strategies to assist during reintegration;
— Identify negative coping strategies;
— Recognize behavioral signs that may indicate external support is required;
— Challenge barriers to seeking mental health care; and
— Identify informal and formal mental health resources.

Training was delivered in two 2-hour blocks, on Day 2 and Day 3. The new TLD program was delivered in a group setting with an average group size of 27 learners (range = 2–50). Where possible, those who worked together while deployed received training together. The training included one individual exercise and three small-group exercises.

A mental health clinician (largely social workers and mental health nurses) and a nonclinician (either a veteran with a history of service-related mental health problems or line personnel with a special interest in mental health, similar to the Master Resilience Trainers10 used in the United States) provided the training as a pair. Instructors underwent a minimum of 30 hours of training on psychoeducation in general and on the new TLD program in particular. Successful completion of a practical competency test was required.

Evaluation Process

For the new program only, participants completed identical 15-item pre-/postsession evaluations covering attitudes and self-efficacy that were tied to the program's objectives. Five-point Likert scales were used, with 1 = “strongly disagree” and 5 = “strongly agree.”

For both programs, participants completed an overall evaluation form at the very end of TLD that captured satisfaction with the TLD and with different aspects of it (notably the educational component) and key attitudes and self-efficacy using items obtained from the developers of the Battlemind program.7 Most of the mental health care barrier items have been used in earlier research.9,11–13 Program satisfaction was measured using a four-point “forced choice” Likert scale (with no middle category), and attitudes and self-efficacy were measured using a five-point Likert scale. To simplify analysis of satisfaction data, the response categories were dichotomized into satisfied vs. unsatisfied.

Analysis

Analysis was done using SPSS for Windows, version 15.0. Univariate association of categorical variables was explored using the $\chi^2$ test. The association of satisfaction with the new vs. old program was explored with binary logistic regression.

In order to simplify analysis and presentation of results, exploratory principal components analysis with a varimax rotation was done on the 15 presession knowledge and attitude items to identify items for calculation of subscales. Cronbach's a was used to measure subscale reliability. Pre-/postsession differences and differences in attitudes and self-efficacy for the old vs. new program are expressed as Cohen's d, a standard measure of effect size for continuous data; confidence intervals (CI) for d and significance testing was done using the Z distribution. Differences in overall satisfaction with the TLD program under the old vs. new program were assessed using Somer's d, a nonparametric measure of ordinal association.

Ethical Aspects

Completion of all questionnaires was voluntary and anonymous. As routine evaluation of an educational program, approval by a Research Ethics Board was not required.

## Results

For the new program, pre-/postsession knowledge and attitude questionnaires were received from 2,952 and 2,950 respondents, respectively (98% response rate). For the new and old programs, overall TLD evaluation forms were received from 2,935 and 14,253 respondents, respectively (97% and 75% response rates). A leading factor in the lower response rate for the old program was the loss of more than 2,000 evaluation forms in a single shipment from theater. Characteristics of the participants are shown in Table II; there were small differences in the sociodemographic and military characteristics of those who underwent the old and new programs, with those receiving the new program being slightly younger, more likely to be Reservists, and more likely first-time deployers ($p < 0.001$ by $\chi 2$ test).

Principal components analysis of the 15 pre-/post-session items yielded three factors accounting for 58% of the variance. The first factor (44% of the variance, 6 strongly loading items, $\alpha = 0.87$) represented confidence in knowledge and abilities for transition, and included items such as “I can identify positive and negative ways of coping with transition challenges,” “I understand the common experiences that occur during transition from deployment to home,” and “I am confident in my ability to recognize when to get a transition problem checked out.” The second factor (8% of the variance, 5 items, $\alpha = 0.83$) captured mental health literacy with items like: “I would know what to expect in terms of treatment if I were to seek help for a mental health problem,” “I understand the impact of seeking mental health care on my CF career,” and “I can identify a number of different ways to get help for mental health or transition problems in the CF.” The final factor (7% of the variance, 3 items, $\alpha = 0.73$) represented the sense of personal responsibility toward others regarding their mental health with the following items “It's my responsibility to help a buddy with a mental health problem” and “It is the responsibility of leaders to encourage CF members to get help when they have mental health problems.” A cross-loading item that was conceptually related was also included in this last subscale (“I am confident in my ability to help CF members get assistance for a mental health problem”). Subscales for each of these factors representing the mean score of loading items were calculated for each respondent. A single item (“If I had a mental health problem, I would definitely get professional help for it”) cross-loaded on all three factors and was not used in any subscale.

Presession attitudes and self-efficacy were largely favorable, though mental health literacy showed the lowest average score (3.6 out of 5). All three knowledge and attitude factors saw statistically significant improvements in response to the new program with the effect sizes (Cohen's d) for confidence in knowledge and abilities, mental health literacy, and sense of responsibility for others being 0.66, 1.03, and 0.44, respectively.

As shown in Figure 1, the new program was associated with significantly higher satisfaction with the educational program ($p < 0.001$). The “strongly agree” category showed a particularly strong shift in favor of the new program (39% vs. 16%). Figure 2 suggests that

this increased satisfaction with the educational component enhanced the perceived value of TLD as a whole ($p < 0.001$). All measured sociodemographic and military characteristics had a statistically significant univariate relationship with satisfaction with the educational program and/or perceived value of TLD (Table IV ). Logistic regression models confirmed that the new program was independently associated with increased satisfaction with the educational program (adjusted odds ratio [OR] = 3.8, 95% CI = 3.2–4.5) and greater perceived value of the TLD (OR = 1.7, 95% CI = 1.5–2.0).

Sociodemographic and military characteristics showed different patterns with respect to satisfaction with the educational program and with TLD as a whole. Women were more likely than men to be satisfied by the educational program and to find TLD valuable. Officer rank was associated with lower satisfaction with the educational program but greater perceived value of TLD. Previous deployers were equally satisfied with the educational program but found TLD as a whole to be less valuable. No significant interactions were seen between the new educational program and any of the sociodemographic or military characteristics (results not shown).

Respondents who had underwent TLD in Cyprus previously (N = 1,491) were asked to compare the value of their current TLD experience with their previous one. As shown in Figure 3 , the 984 respondents who completed the old educational program twice tended to find the current experience less valuable. In contrast, the 507 who completed the old program followed by the new program strongly favored the current experience.

The new program was also associated with more favorable attitudes toward mental health care (Table V ). In particular, negative attitudes toward care (e.g., “I don't trust mental health professionals” and “Mental health care doesn't work”) were less prominent in those who received the new program (Cohen's d = −0.29 and −0.27, respectively). Other attitudes (most involving stigma, such as “I would be concerned about what others might think”) showed little or no difference between the two programs. The new program was also associated with greater self-efficacy surrounding mental health (Cohen's d = 0.20–0.34); these items covered self-efficacy both for identifying fellow CF members with mental disorders and for self-management of mental health problems. The univariate differences remained significant after adjustment for potential confounders (sex, age, component, rank, and deployment history) using univariate analysis of variance (results not shown).

## Discussion

### Key Findings

This article demonstrates the superiority of a new TLD educational program over an earlier program including Battlemind, at least when it comes to very short-term outcomes. Three pieces of evidence are presented: first, pre- and postsession attitude and self-

efficacy questionnaires established that the program resulted in medium to large changes in the domains of confidence in mental health knowledge and abilities (Cohen's d = 0.66), mental health literacy (d = 1.02), and sense of responsibility toward others (d = 0.44). Second, overall evaluation forms completed at the end of TLD showed that the new program was independently associated with a significantly increased odds of satisfaction with the educational program (adjusted OR = 3.2), and this satisfaction translated into increased odds of perceiving the TLD as a whole to be valuable (OR = 1.7). Those who had also undergone TLD under the old program clearly preferred the new one, on average. Finally, overall evaluation forms also showed that those who completed the new program had significantly more positive attitudes toward mental health care, though the magnitude of this effect was small (Cohen's d = −0.29 to −0.18). Stigma-related attitudes showed little or no difference between the two programs (d = −0.15 to 0.02). The new program was associated with a small but significant advantage with respect to self-efficacy for helping others with mental health problems and for self-management of mental health problems. Gains in self-efficacy for self-management occurred alongside gains in knowing when to seek care, suggesting that messages on resilience and care-seeking are not incompatible.

Although the new program did not appear to be superior with respect to its impact on mental health stigma, it was superior in its apparent impact on negative attitudes toward care (e.g., “Mental health care doesn't work”). This is an important finding because these attitudes are emerging as a much stronger predictor of care-seeking than stigma.9 These negative attitudes received much greater emphasis in the new program.

The absence of significant interactions between key sociodemographic and military characteristics and the new program suggests that it has broad appeal and relevance: All subgroups benefited similarly from the new program. In short, all indicators point toward a significant advantage of the new program over the old one for all participant groups.

Comparison With Other Literature

The only real point of comparison for the present study is the rigorous group randomized trial of the postdeployment Battlemind training program done by its developers.7 One hour of Battlemind delivered at the time of redeployment was superior to a conventional lecture on stress both in terms of training satisfaction and with respect to psychological well-being, 4 to 6 months later. For those with heavy combat exposure only, the investigators found well-being effect sizes (Cohen's d) ranging from 0.06 to 0.30. These effect sizes cannot be compared to any of the effect sizes in our sample, which deal with much shorter-term outcomes where higher effect sizes are achievable.

Although we earlier emphasized differences between our new program and the old program rooted in Battlemind, our “new” program did draw heavily from key elements of Battlemind. Key similarities included:

— Emphasis on identifying when help is needed;

— Reframing transition difficulties as consequences of appropriate adaptations to combat;
— Expectation of normal transition and reintegration for most;
— Use of military relevant examples;
— Skills-oriented;
— Strengths-based;
— Projective emphasis on unit cohesion and buddy support; and
— Use of sound principles of adult education.

In addition, our new program incorporated an element from another effective Battlemind program, Battlemind debriefing,7,15 in that it started with a brief review of the difficulties encountered during the deployment. This exercise served as an ice breaker, acquainted the instructors with the unique experiences of each group, and permitted adaptation of the program to their specific needs (such as those related to noncombat operations).

Limitations

The foremost limitation of any pre-/postintervention evaluation strategy is that differences between the two groups other than the intervention could account for its apparent benefits. For example, the differences in response rate between the old and new programs could have introduced bias. We controlled for a limited number of confounders, but other factors may have been at play. The individuals best suited to judge the likelihood of bias are those closest to the project, namely the authors of this article. Although we can hypothesize any number of differences in, say, leadership, deployment experiences, and other factors, we cannot, in all honesty, identify any differences substantial enough to account for the sizable apparent advantage of the new program over the old one. We thus believe that a good fraction of the observed advantage reflects a real difference in the short-term educational impact of the program.

If one accepts these differences as real, a more important question arises: Will the short-term improvements in satisfaction, attitudes, and self-efficacy translate into improvements in long-term well-being and functioning? Military leaders want healthy, productive personnel, not just satisfied and confident learners. Battlemind has documented small but meaningful advantages in terms of well-being at 4 to 6 months after return, whereas the benefits of our new program are unknown. Whether the favorable changes we document here will translate into more meaningful changes in well-being hinges first upon the persistence of those effects over time. Unfortunately, the effects of health education (and education in general) are often transient. However, our TLD program has short-term outcomes in mind, namely successful reintegration over the weeks and months following return, so even a transient effect may be perfectly satisfactory. Our larger mental health training program also offers opportunities for reinforcement during other training

offered across the career and deployment cycle. The translation of benefits in terms of attitudes into benefits in terms of well-being also hinges upon the extent to which those attitudes mediate well-being. Such mediation is likely but unproven.

One final major limitation: We cannot say which specific aspect of our new program (Table I) accounted for its apparent advantage. Was the team delivery model the active ingredient? Its less combat-centric content? Its delivery in more cohesive groups? Its slightly longer duration? Its greater interactivity? Better instructors, perhaps? Or simply the novelty of the new program? All of these are plausible, and none of them are testable with our data.

There are also some technical limitations in this analysis: Assumptions required for some of the statistical analyses may have been violated. We could not control for instructor or multilevel effects in training outcomes, and these may have been significant.7 Use of logistic regression for other than rare outcomes (e.g., training dissatisfaction) has limitations.16 Some of our respondents had participated in TLD earlier, hence their responses on the evaluation of the new program violated the assumption of statistical independence of the observations. We were not able to do paired comparisons for the pre- and post-tests; instead, we could only document differences at the cohort level. Finally, we did not adjust for multiple comparisons. We acknowledge these limitations but judge them to be much smaller threats to the conclusion as to the advantage of the new program over the old one. These would be of greater concern if the observed differences were less dramatic than they proved to be.

## Conclusions

The development and validation of Battlemind was a landmark event in military mental health training, and we are deeply indebted to its creators. The meaningful gains in terms of well-being after a brief, group intervention for unselected learners made us look at military mental health training in a different light. When we selected Battlemind as the cornerstone of our TLD program, we endeavored to alter it as little as possible, with the intention of retaining as much of its efficacy as we could. At that time, we were not certain of what parts of the program were the active ingredients or if the whole was greater than the sum of its parts.

Over time, our understanding of deployment-related mental health problems and barriers to mental health care has progressed. Evidence is accumulating that psychoeducation and resilience training may actually be broadly effective.17,18 In response to these changes, Battlemind itself has evolved in “Battlemind Resilience Training,”19 and other resilience training has grown up around it in the form of the Comprehensive Soldier Fitness strategy.20 Arguably, this article amounts to the testing of our new program (one with Battlemind firmly in its genome) against an increasingly obsolete version of the original program. However, many of the changes we made in our program mirror those in the mental health training in the U.S. Army. For example, Comprehensive Soldier Fitness relies

on nonclinicians as trainers,10 and the training as a whole is less combat-centric.21 As such, the apparent success of our new program may have something to say about the effects of similar changes elsewhere.

Turning away from Battlemind, even as much as we have, has been a difficult decision. However, we are confident that we have retained enough of its essence to preserve its demonstrated benefits. Moreover, this evaluation suggests that our new program has, if anything, added to those benefits.

BACK TO TOP

## Multidisciplinary Response to the Escherichia coli O104 Outbreak in Europe

Military Medicine
LTC Charles C. Dodd, VC USA; CDR Michael J. Cooper, USPHS United States Army Public Health Command Region-Europe, CMR 402, APO AE 09180, Landstuhl, Germany
November 2012

Abstract

The 2011 outbreak of Escherichia coli (E. coli) O104 in northern Germany resulted in over 4,100 illnesses, 900 cases of hemolytic uremic syndrome, and 50 deaths. The U.S. Army's Public Health Command Region-Europe established a multidisciplinary advisory team to prevent E. coli O104 exposure in the Department of Defense (DoD) population. This decentralized, interagency team engaged European public health authorities and coordinated control measures including risk communication. Following German public health investigations, the DoD advisory team compiled information from available official reports, assessed risk, and published guidance to halt the local procurement and supply of suspect foods for all DoD installations in Europe. Advisory team members processed outbreak information, adjusted advisories, and coordinated response efforts. The advisory team quickly identified authoritative information sources, coordinated case definitions, and streamlined potential case reporting. Timely and accurate risk assessment, management, and communication were vital in protecting the DoD population during this outbreak. There were no cases in DoD-related personnel.

Introduction

Shiga toxin-producing Escherichia coli (STEC) remain a significant cause of foodborne illness.1 Since the emergence of Escherichia coli (E. coli) O157 in the early 1980s,2,3 non-O157 STEC have become an increasing foodborne illness burden.4–6,7 Outbreaks of rare STEC serotypes may create diagnostic challenges that hinder public health responses to outbreaks. On May 19, 2011, the Health and Consumer Protection Agency in the Hamburg area of northern Germany reported three pediatric hemolytic uremic syndrome (HUS) illness cases that were suggestive of enterohemorrhagic E. coli infections.8 The Robert Koch Institute (RKI) investigated the following day and confirmed that there were several HUS cases in the area that also included an unusually high number of adults.9 The RKI soon determined that the cluster of HUS cases was caused by E. coli O104; isolates were later determined to be E. coli O104:H4. In the following days, the Hamburg area of northern Germany was identified as the epicenter of the outbreak; an additional cluster of outbreak cases occurred in the Bordeaux area of France.10 Initially, outbreak investigations by German public health authorities revealed epidemiologic links among the outbreak cases and the consumption of raw tomatoes, lettuces, and cucumbers. Later, contaminated sprouts were identified as the most probable cause. The outbreak lasted from May 1 to July 26, 2011. The STEC strain appeared particularly virulent and there were an unexpectedly high proportion of HUS cases.9 As shown in the Table I , the 2011 outbreak of Escherichia coli O104:H4 in Europe resulted in over 4,000 human illnesses, 890 cases of HUS, and 50 deaths.

During the outbreak, multiple public health agencies attempted to prevent the spread of E. coli O104:H4 infections. Initial case–control studies by the RKI suggested that vegetables were the source of the outbreak.11 Unfortunately, the complex supply and distribution system of vegetables within Germany and the European Union added to the difficulty of tracing potentially contaminated foods. One large population at risk was the U.S. Department of Defense (DoD), which procures local-sourced food for over 200,000 of its personnel and their family members in the region. The U.S. Army Public Health Command Region-Europe (PHCR-E), which is responsible for food protection and infectious disease surveillance for DoD personnel throughout Europe, initiated a DoD-wide unified response in order to prevent spread of the outbreak to the DoD's geographically dispersed population throughout Germany and Europe.

The E. coli O104:H4 outbreak presented several unique challenges to an effective DoD public health response. Because a rare serotype was involved, commercial laboratory assays specific for E. coli O104:H4 were not available, leading to uncertainty of the size of the outbreak.12 Early in the outbreak, case definitions used by the German public health authorities were unclear and reporting delays were evident.13 The U.S. Army PHCR-E was forced to make outbreak response decisions in the face of uncertainty, which is often inherent to second-hand reports of aggregated data from outbreak investigations. Because of the severity and rapid growth of the outbreak involving a rare STEC serotype, news media-reported information was occasionally conflicting and difficult to confirm by the DoD. Hence, the DoD response required tailored solutions to these unique problems. The purpose of this manuscript

is to describe the DoD's multidisciplinary, multiagency response to the E. coli O104:H4 outbreak and suggest best practices for future outbreak responses.

Multidisciplinary Ad Hoc Advisory Team

When the RKI publicly announced the E. coli O104:H4 outbreak, the U.S. Army PHCR-E began assessing risk to DoD personnel and established a multidisciplinary advisory team to monitor and prevent E. coli O104 exposure in the European DoD population. This decentralized, interagency team of physicians, epidemiologists, laboratory diagnosticians, food safety specialists, logisticians, public affairs officers, environmental health officers, and veterinarians representing all European DoD organizations engaged German and European public health authorities, coordinated control measures, and communicated risk to military and civilian stakeholders across Europe. This ad hoc advisory team provided guidance for U.S. Army, Air Force, Navy, Marine, and Public Health Service entities throughout the region. Food-related personnel from the U.S. Defense Commissary Agency, Defense Logistics Agency, and the Army and Air Force Exchange Service were strategically included in the team since effective control measures would require unified mitigation actions involving DoD food supply routes.

The PHCR-E ad hoc advisory team met daily and processed outbreak information, adjusted advisories, and coordinated response efforts. A 24/7 teleconference line was established to facilitate interagency involvement and allow expert panel discussions in real-time as outbreak information became available. The advisory team secretary electronically disseminated meeting minutes daily to all team members in order to invite corrections and clarify current tasks and concerns. All team members maintained their preoutbreak job assignments and normal operations; no personnel were able to devote all of their work efforts toward the outbreak response. Daily media releases were coordinated with a DoD public affairs officer and timed according to daily reporting deadlines of available media channels. The advisory team disseminated outbreak response information through official command channels of the DoD in the form of concise, one-page situational reports. News media reports were read and assessed daily in order to better understand public perception and consumer attitudes during the outbreak. This feedback was necessary in order to adjust DoD risk communication techniques and assess control measure effectiveness.

The PHCR-E ad hoc advisory team served a crucial role in balancing various stakeholder differences with the need to provide a unified outbreak response within the DoD. Food logistics stakeholders had various levels of risk aversion, food protection resources, and logistic distribution networks. Daily advisory team meetings allowed these individual stakeholders to voice concerns or complaints and achieve timely solutions. This participatory environment empowered the team to examine 2nd and 3rd order effects of planned advisories and control measures and allowed preemptive corrections. Also, public health and logistical stakeholders were able to present a unified message to DoD food customers, creating an atmosphere of transparency and trust.14

Although epidemiologic information and diagnostic data were limited, the advisory team agreed upon two primary information sources (RKI and the World Health Organization), traced suspect food pedigrees, pursued harmonized case definitions, and streamlined potential case reporting within DoD medical treatment facilities. Medical personnel in DoD facilities across Europe immediately began to look for potential cases among the DoD population and ensure outbreak awareness by all health care providers. Upon public release of initial outbreak investigation results by the RKI, the PHCR-E ad hoc advisory team assessed the risk to DoD personnel and family members and published guidance to halt the local procurement and supply of raw cucumbers, lettuces, and tomatoes for all European DoD installations. This action was based upon RKIs initial case–control study that revealed associations between HUS patients and the consumption of tomatoes, lettuces, and cucumbers. In these initial exploratory interviews, only 25% of HUS patients had consumed sprouts. Consequently, low-exposure foods like sprouts were not included in initial statistical models in order to reduce the risk of sporadic correlations. Subsequent, more comprehensive case–control studies using univariable and multivariable analyses of data from larger study populations revealed sprouts as the likely cause of the outbreak.

Outbreak Response Challenges

The PHCR-E ad hoc advisory team encountered several unique challenges during the outbreak. During a large epidemic involving a rare, virulent strain of STEC, prompt and accurate risk communication was necessary in order to effectively prevent exposure within the DoD population.15 Of the U.S. DoD agencies involved, only the U.S. Army, Air Force, and Navy had their own medical assets. To achieve an effective outbreak response, cohesive advisories and control measures were necessary across a complex intertwined system of DoD personnel, food supply, health care, and news media systems. During the outbreak, DoD stakeholders desired a single DoD source of information; the ad hoc advisory team fulfilled this role by providing transparent and timely risk communication in the form of official advisories, as well as nontraditional forms such as food safety informational sheets that targeted the general populace (Fig. 1 ).16 The advisory team aspired to be the preeminent comprehensive source of outbreak information. Many DoD personnel and family members first learned about the E. coli O104:H4 outbreak when DoD food retail and food service establishments posted notices and removed suspect food items from the supply system. This expedient and active form of risk communication increased consumer perception of DoD food retail and food service establishments while counteracting dramatized and less-accurate reports that had begun to circulate in local, civilian news media.10

Another significant challenge to an effective outbreak response was the wide geographic distribution of outbreak cases and potentially contaminated foods across multiple countries that have different languages, regulatory agencies, and food production systems. Although the source of contaminated food was eventually defined, initial outbreak responses included multiple-country surveillance efforts to detect contaminated foods and outbreak sources. This increased level of surveillance and open communication

of presumptive associations created political tension among European Union countries and inflamed food production industries. As seen during this outbreak, initial consumer perceptions can override attempts to retract false information, and presumptive risk communication during an outbreak can wrongfully damage food production markets.17

Suspect cases within the DoD were classified by diagnostic algorithms using proxy test protocols. From June 2 to July 1, 2011, DoD medical personnel identified 43 suspect cases. Since a commercially available and validated test specific for the E. coli O104:H4 serotype did not exist, DoD cases of hemorrhagic colitis were only considered suspect cases if they were culture negative for Yersinia spp., Vibrio spp., and E. coli O157. Suspect case samples were then tested by polymerase chain reaction for Shiga toxin 1 (stx1), Shiga toxin 2 (stx2), and intimin (eaeA). Since E. coli O104:H4 isolates were reported to be stx1-negative, stx2-positive, and eaeA-negative, tested samples having this virulence factor profile were considered to be a probable outbreak case.18 No samples from suspect DoD personnel met this criterion. Although this testing algorithm was only a proxy classification for the E. coli O104:H4 serotype, it was the only commercially available diagnostic protocol used during the outbreak.19

In contrast, the case definition employed by the World Health Organization and RKI had three categories: possible epidemic case, probable epidemic case, and confirmed epidemic case.13 A possible epidemic case was any person who developed diarrhea or HUS on or after May 1, 2011, and the fecal sample had laboratory evidence of Shiga toxin production. A probable epidemic case included the above criteria plus a history in the previous 14 days of being in a country (Germany) where another person had acquired the infection and had consumed food in Germany or had close contact with a confirmed case. A confirmed case was defined as any person who was a possible epidemic case and from which had been isolated E. coli O104:H4 or E. coli O104 while fulfilling the criteria of a probable case. The confirmed case category of this case definition required a laboratory protocol specific for the E. coli O104 serotype, which was not directly available to DoD laboratories. Hence, the aforementioned proxy protocol was used within the DoD medical system to identify probable cases.

## Outbreak Response Concepts

Although risk is commonly defined as a combination of the probability and severity of an adverse event, the ad hoc advisory team's risk assessments included a third dimension: food product need. Although product necessity does not directly affect risk, it is critical to risk management in relation to consumer behavior. The consumer need for suspect food items during the outbreak largely impacted advisories and control measures that the advisory team implemented during the outbreak. In accordance with RKI's initial advisory not to consume raw cucumbers, lettuces, and tomatoes of European origin, the advisory team implemented the same advisory for all DoD personnel and family members in Europe even though some sources of cucumbers, lettuces, and tomatoes were of less risk because they originated from sources less likely to be involved in the outbreak. Regardless, because these products were not essential to DoD consumers until the source of the outbreak was confirmed, the advisory team decided to impose a wide safety

net in order to prevent exposure of contaminated foods in the Europe-wide DoD population. Although food pedigrees could be traced, cross-border transport of suspect foods during acute changes in agricultural markets was plausible. Low market demand for suspect foods in areas close to the outbreak could create a surplus of food types associated with the outbreak, encouraging producers and processors to transport these products to other market areas where consumers would be more likely to feel safe consuming the products. Hence, the advisory team considered the third dimension of food product need, as well as unusual food distribution patterns, when assessing and managing risk during the outbreak.

Several guides exist for the conduct of outbreak investigations, but best practices for public health response to outbreaks that are investigated and analyzed by other agencies have seldom been defined. Had the DoD encountered outbreak cases among the DoD population, detailed coordination between numerous United States, German, and European public health agencies would have been necessary. The PHCR-E ad hoc advisory team was faced with making several decisions in the face of having only summarized outbreak investigation data that were reported to the public. In order to reduce uncertainty and improve alignment of efforts, the ad hoc advisory team members engaged public health officials at the RKI telephonically on multiple occasions in order to clarify issues and confirm media reports. These interactions were crucial for the advisory team's decision-making process.

Transparent, clear, and accurate risk communication is critical during effective outbreak responses. The ad hoc advisory team's communication of risk through diverse DoD organizational channels and DoD public media channels was necessary. Early during the outbreak response, gaps were noted in the dissemination of credible information to the DoD population within Europe. Social media and e-mail transmissions within DoD communication systems conveyed deficiencies regarding public understanding of E. coli O104 and the epidemiology of the outbreak; these information gaps were mitigated by revising electronic dissemination lists within and among DoD organizations and aggressively disseminating unified, accurate messages via e-mail advisories, electronic food safety infosheets, news releases, and interviews to the DoD's American Forces Network. Competent public affairs officers enabled the advisory team to target various audiences, including personnel living off military installations, and release updated information before daily news media submission deadlines. Expedient, accurate, and transparent assessments of the outbreak situation, while striving to clearly communicate practical prevention methods, were key in mitigating the risk of exposure.

## Conclusion

Few peer-reviewed publications describe infectious disease outbreak responses when public health professionals may not have primary access to outbreak investigation data. In the face of the devastating 2011 E. coli O104:H4 outbreak in Europe, the U.S. Army Public Health Command Region-Europe's ad hoc advisory team was able to engage competent German and European public health authorities. A focused epicenter from a point source aided the public health response. Expedient and clear risk communication

should remain a critical focus in public health responses to foodborne illness outbreaks. Complex food distribution systems and diagnostic hurdles in foodborne illness outbreaks demand tailored solutions. Effective outbreak responses within complex organizations require the leverage and coordination of multidisciplinary, interagency assets in order to achieve unified risk mitigation in the face of uncertainty. Ad hoc response teams provide a suitable platform to communicate and act quickly, responsibly, and cohesively to prevent the spread of foodborne illness during outbreaks.

BACK TO TOP

## Medical Evacuation and Triage of Combat Casualties in Helmand Province, Afghanistan: October 2010–April 2011

Military Medicine
LCDR Jonathan E. Clarke; Lt Col Peter R. Davis
November 2012

ABSTRACT

Medical evacuation of combat casualties in Operation Enduring Freedom-Afghanistan is achieved primarily by helicopter, because of distances involved as well as ground-based threats. In Helmand Province, evacuation from the point of injury may occur on a variety of helicopter evacuation platforms with disparate levels of attendant medical expertise. Furthermore, triage to a medical treatment facility may involve varying echelons of care before definitive management. Consequently, considerable differences in medical care may be encountered between point of injury and definitive treatment. We discuss the role of helicopter-based medical evacuation in Helmand, Afghanistan, as well as triage and timelines to the most appropriate medical facilities. Based on our experience and available evidence, we have made recommendations to regional commanders which favor the utilization of prehospital critical care teams aboard helicopter-based evacuation platforms and direct triage to the highest echelon of care available when feasible.

INTRODUCTION

Medical evacuation (MEDEVAC) and triage of combat casualties in Helmand Province, Afghanistan, involves a joint service, multinational system, with a wide range of capabilities, delivering care in a hostile, austere environment. Decisions regarding MEDEVAC tasking and triage of combat wounded to the most appropriate medical facility are driven by a number of influences. These include mechanism or type of injury, patient stability for transport, evacuation timelines, availability of assets, and prevailing

tactical and weather conditions. Additionally, historical precedents and biases along the decision chain play considerable roles. Such decisions occur in a complex medical battle space, and evidence to support or refute current practice is continually evolving as lessons learned from Operation Iraqi Freedom (OIF) and Operation Enduring Freedom (OEF) are identified.

Several controversial decisions regarding MEDEVAC tasking and implementation, as well complications during primary and secondary transport, occurred in Helmand Province during the period between October 2010 and April 2011. These events were reviewed by a multinational group of peers, and evidence-based recommendations were made to regional commanders in an effort to effect performance improvement. Based on a review of pertinent literature and our own experiences at UK Role 3 Hospital Bastion, we discuss the role of helicopter-based MEDEVAC in Helmand Province, Afghanistan, as well as triage and timelines to the most appropriate medical facilities, in keeping with the motto of "right patient, right care, right time."

## DISCUSSION

### The "Golden Hour" in Military Trauma

The concept of a "golden hour" in trauma has circulated for over 35 years and has long influenced the doctrine of military medicine despite a lack of clear supporting evidence. In OEF-Afghanistan, the U.S. doctrinal timeline for evacuation of "urgent" casualties to a NATO Role 2 or 3 medical facility is less than 60 minutes, despite a published NATO standard of 90 minutes (Control and Coordination of Medical Evacuation within Regional Command Southwest [RC(SW)]. Annex B (Ed 2) to RC(SW) SOP 6001, June 2010). Any mission which exceeds this self-imposed 60-minute limit is subject to a delayed mission report. For reasons we will describe, it is our opinion that strict and arbitrary adherence to this "golden hour" may be at odds with optimal patient care in the current conflict in Helmand Province and may have the potential to result in additional morbidity and mortality. This may occur by suboptimal tasking of MEDEVAC assets, triage to a less-resourced medical facility and/or aviation mishaps.

Recent literature has refuted the concept of the golden hour, and multiple studies have largely failed to demonstrate an association between prehospital times and mortality in either the civilian or military setting. Military trauma differs from its civilian counterpart in several important ways, but relevant to this argument is the temporal distribution of death.

On the battlefield, historical evidence suggests a bimodal distribution of death, with the first peak secondary to catastrophic injuries within 15 minutes of wounding. The second peak of death is reported to occur between 60 and 180 minutes after wounding. It is unclear what effect implementation of current Tactical Combat Casualty Care guidelines and liberal tourniquet use might have on this figure in current conflicts. Immediate surgical intervention within the first 15 minutes after wounding has been shown to have no impact on survival; so, there appears to be no role for placement of extremely far-forward surgical assets. Therefore, emphasis has

been logically placed on evacuating early survivors to the closest medical facility, and these capabilities are generally disbursed such that evacuation to medical care occurs within 60 minutes, or during the “golden hour.” This model persists despite growing evidence that longer evacuation timelines may be acceptable.

Echelons of Care on the Modern Battlefield

Initial experience with dynamic, large scale invasion and conventional warfare in Iraq resulted in the concept of forward deployed medical and surgical teams to stabilize casualties close to the point of injury before evacuating to more definitive care, which often required prolonged evacuation timelines. “These units were designed and staffed to support maneuver warfare in an immature theater where Role 3 medical facilities were not readily available (HR Bohman, personal correspondence).” The birth of highly mobile Forward Surgical Teams and Surgical Shock Trauma Platoons proved successful in this environment and outcomes were reported as equivalent to Role 3. NATO classification of these units would be Role 2+, with limited surgical and holding capacity, whereas Role 3 is defined as a field hospital, with specialist surgical and other advanced capabilities.

As the battlefield matured and Role 3 field hospitals became more readily accessible, it became increasingly difficult to justify forward basing valuable surgical assets close to troop buildups. Nonetheless, this principle, whose foundation lies on the concept of a “golden hour,” has largely persisted in OEF-Afghanistan, where Role 2s are often forward deployed alongside combat troops.

Importantly, Role 2 medical assets are managed by operational commanders rather than an integrated, theater-wide trauma system, and placement is often independent of Role 3 hospital proximity. This has occasionally resulted in the two medical capabilities being placed within as little as 10 to 15 minutes of each other. We express concern that commanders may be tempted to position Role 2s for “peace of mind” without clearly understanding their value or their place within a coordinated trauma system, which currently provides rapid access to more capable Role 3s. When confronted with evidence refuting their utility in certain circumstances, doctrinal change to this approach has been met with resistance.

The utility of Role 2s in a mature battlefield is now being more closely scrutinized, especially considering the minimal difference in evacuation timelines from point of injury directly to Role 3s. We have found this difference in Helmand Province to be generally less than 20 minutes using available helicopter-based MEDEVAC, and total evacuation times remain less than 1 hour in most cases (Patient Evacuation Coordination Center RC(SW) unclassified data). In our experience, we have found that a delay in definitive care of 1 or more hours may occur when the patient is initially stabilized at Role 2 and secondarily transferred to Role 3. We believe this is largely unjustified when considering short evacuation timelines and the disparity in capabilities which exist between Role 2 and Role 3.

Controversies in Defining the Utility of Role 2

Role 2 capabilities are variable and dependent upon the composition of personnel and supplies, but resources are typically quite limited. As an example, the nominal requirement for blood stocks at typical Role 2 facility is 50 units of packed red blood cells, often with no other blood component therapy. A walking blood bank is also available, but its use incurs a delay of at least 30 to 40 minutes. This is often more than the time required for helicopter transfer to Role 3, during which transfusion of up to 8 units of blood products may be accomplished en-route.

Personnel experience at Role 2 is also variable, and trauma teams may be required to care for the most critically injured patients imaginable, with potentially insufficient recent experience and resources at their disposal. There exists a paradox, whereby these casualties require the most skilled and experienced clinicians and nursing staff, as well as the most rehearsed resuscitative teams; yet, they often operate in an environment which does not foster skills maintenance, with low patient volumes and acuity. We feel that the experience, skill, and teamwork required for optimal patient outcomes in severe polytrauma cases can only derive from recent exposure to a large number of seriously wounded patients with adequate resources. It is our opinion that these capabilities do not always exist at Role 2s.

Yet, despite scrutiny, there appears to be at least a limited role for Role 2 employment in the mature battlefield. Eastridge previously described comparable outcomes between a U.S. Army Forward Surgical Team (Role 2+) and the Combat Support Hospital (Role 3) in Iraq. Civilian data also supports the utility of Role 2 equivalent facilities in long-distance rural settings where stabilization by Advanced Trauma Life Support-trained providers is accomplished before transfer to a level 1 trauma center. This scenario may be analogous to stabilization at a remote Role 2, with secondary retrieval to Role 3 in less than 4 hours. However, this comparison is not clearly applicable to the current outlay of medical resources in Helmand.

Historical military evidence most clearly supports the utility of Role 2 in the immediate stabilization of hypotensive victims of isolated penetrating neck and torso trauma. These patients are known to benefit from early surgical intervention, and Chambers' and Bohman's experience in Fallujah, Iraq supports this conclusion. Just how early such surgical intervention is needed is less clear, but Bohman suggests 30 minutes as a limit in 3% of cases. However, in OEF-Afghanistan, patients with isolated penetrating neck and torso trauma account for less than 10% of combat casualties, and only 10% of these patients may have injuries which are ameliorable to immediate damage control surgery. This yields very little utility for far-forward deployment of valuable surgical resources on the battlefield, and we argue that this capability may not be worth the associated costs in some cases.

Arguments for Triage Directly to Role 3

We feel that there are often convincing reasons to bypass Role 2 and triage the most critically injured polytrauma patients directly to Role 3 field hospitals. Historical evidence, personnel experience, resource management considerations, and evacuation timelines all seem to subjectively favor direct triage to Role 3 in many of our cases in Helmand. There is currently little objective evidence to resolve at what distance or evacuation timeline triage should proceed to Role 2 versus directly to Role 3 in the current setting, and this remains a subject for future study.

Civilian literature generally supports the transfer of trauma patients to the highest level of care available within the region, the level 1 trauma center, often bypassing lower echelons of care. In the military setting, this capability is represented by the Role 3 field hospital, whereas Role 2s may represent less-resourced community hospitals. We feel the current military trauma system in Afghanistan lacks the thorough integration between echelons of care which is present in civil practice. It is our opinion that the military system stands to gain efficiency and improvement in trauma outcomes by revising and integrating the placement of medical facilities of all echelons of care across the battle space.

We find the most compelling argument for direct triage to Role 3 considers the material resources available to Role 2 versus Role 3 facilities. In our experience at UK Role 3 Hospital Bastion, a victim of dismounted improvised explosive device (IED) blast with multiple extremity amputations and severe polytrauma may require resuscitative procedures by a team of more than 10 physicians.

We believe that the complexity and resource-intensive nature of such resuscitations exceeds the capabilities of most Role 2s and would render them materially incapable of managing the next critical casualties. However, this type of injury pattern and resuscitation is common in our experience at Bastion, where victims of IED-associated blast receive an average of 43 units of blood products during primary resuscitation

Initial stabilization of such patients at Role 2 is not known to improve outcomes, and we feel that, at best, it only serves to delay definitive treatment and exhaust Role 2 resources.

Another, often overlooked, consideration for triage directly to Role 3 is minimization of risk to MEDEVAC crews and aircraft. Initial stabilization at Role 2, followed by secondary transfer to Role 3, results in two separate MEDEVAC missions, with a doubling in exposure to a hostile environment and the potential for aircraft mishap and loss of life and property. Additionally, secondary transfers are often associated with a downgrade in the level of care provided to the trauma patient, with subsequent risk of clinical deterioration.

Integration of Prehospital Critical Care into the Military Trauma System

The keystone of an integrated trauma system is the provision of appropriate prehospital care, and evidence clearly supports this. In OEF-Afghanistan, prehospital care and transport to field hospitals is primarily accomplished by helicopter-based MEDEVAC. Within NATO, the skill mix of MEDEVAC crews varies across evacuation platforms and according to the participating nations' own MEDEVAC doctrine, but in Helmand Province, the spectrum of helicopter-based prehospital medical capabilities may be summarized as follows:

- (1) 1 × Emergency Medical Technician-Basic aboard U.S. Army “Dustoff” UH-60A Blackhawk.
- (2) 2 × Emergency Medical Technician-Paramedic aboard U.S. Air Force “Guardian Angel” HH-60 Pavehawk.
- (3) Critical care transport team (1 × Emergency Physician or Anesthesiologist, 2 × Paramedic, and 1 × Emergency Nurse) aboard UK Medical Emergency Response Team (“MERT”) modified CH-47 Chinook.

In the civilian setting, several authors have noted that a higher level of prehospital clinical capability leads to decreased mortality despite longer evacuation timelines. These findings are in keeping with the UK military experience in Helmand and drive the physician-led UK MERT model. Using this model, during a 3-month period in 2009, the rate of unexpected survivors as a result of en-route, physician-led critical care and direct triage to Role 3 was 14.9 per 100 patients, calculated by Trauma Injury Severity Score methodology (JTTC Database Academic Department of Military Emergency Medicine, Birmingham, England.). During the same period, the rate of unexpected survivors by U.S. MEDEVAC was 4.8 per 100 patients (MS Support to CJCS Guidance, June 2010). Notably, during this period, the mean time to Role 3 for the most seriously injured British casualties in Helmand was 99 minutes, in contrast to significantly shorter current timelines.

The crew composition of UK MERT is similar to that of the recently described U.S. Navy Mobile Trauma Bay, as is the clinical capability to provide lifesaving interventions such as RSI, intubation and mechanical ventilation, tube thoracostomy, and hemostatic resuscitation with blood products. Thus, UK MERT may be considered analogous to a helicopter-based Mobile Trauma Bay and is capable of providing critical care within a mean time of less than 40 minutes after 9 line MEDEVAC request.

These (Role 2 equivalent) interventions continue en-route to the receiving medical facility, with an average on-scene time of less than 60 seconds (Patient Evacuation Coordination Center RC(SW) unclassified data).

Additionally, the flight physician is able to provide input as to whether evacuation should proceed to Role 2, Role 3, or another theater-level asset such as a receiving neurosurgical facility. We feel this comprehensive package of clinical capability, experience, and judgment surpasses the capabilities of crews aboard traditional U.S. tactical MEDEVAC platforms, and they should not be

judged as equivalent capabilities (Davis PR, Griffiths A, Nadin MN: Delivering Tactical Prehospital Critical Care—The Emergency Medical Response Team, 2006).

Controversies in Prehospital Care

Debate continues to flourish surrounding the philosophies of “scoop and run” versus “stay and play” during the prehospital response, and there are varying opinions as to whether a physician should be core to the MEDEVAC team. However, various reviews of the literature seem to indicate more favorable outcomes for victims of polytrauma, severe traumatic brain injury, and major thoracic injury when prehospital critical care teams attend the patient. This is the basis of the UK and Israeli Defence Force doctrine for physician-led, MEDEVAC crew composition. Additionally, international civil trauma system models for physician-led, prehospital critical care teams exist in Europe (Austria, France, Germany, Italy, Spain, Switzerland, and United Kingdom) and in Australasia (Australia & New Zealand).

In the military setting, there is clearly a spectrum of medical response appropriate to the mechanism and pattern of injury. For example, the hemodynamically unstable victim of an isolated torso gunshot wound mandates rapid evacuation to the nearest capable medical facility for surgical control of noncompressible hemorrhage and does not require specialist prehospital care—this response favors “scoop and run.” Conversely, the comatose victim of an isolated head injury will benefit from early RSI and hemodynamic stabilization, with direct transfer to a neurosurgical facility—this response favors “stay and play.” Combining these divergent philosophies with short on-scene times and rapid provision of physician-led critical care, the UK MERT model has evolved to “stay and play while running as fast as possible in the right direction.”

Unfortunately, combat injury patterns are rarely clear-cut. Victims of dismounted IED blast may have a complex combination of neurological injury, extremity amputation, and hemorrhage, as well as intrathoracic, intra-abdominal, and/or pelvic injury. Intuitively, it might be assumed that these patients would benefit from immediate evacuation to the nearest medical facility, and this assumption has seemingly influenced current U.S. MEDEVAC practice. However, the available evidence argues that severe polytrauma patients have improved outcomes from early en-route critical care and direct transfer to Role 3 or the civilian equivalent.

Secondary Transfer of Combat Casualties

The underpinning principle of secondary transfer is the need to move the trauma patient to a higher echelon of medical care. This commonly occurs after a patient has been stabilized at a Role 2 facility and requires transfer to Role 3 for more definitive care. Importantly, the level of care provided en-route must be equivalent to or higher than the level of care delivered by the referring facility and should be accomplished by a medical crew experienced in both critical care and transport medicine. This occurs routinely during

intertheater tactical evacuation between Role 3 and Role 4 but is frequently lacking during intratheater transfer. We argue that the transfer crew should, ideally, include an experienced physician trained in anesthesiology, emergency medicine, or intensive care medicine with further training in transport medicine. This is an area of great contention, but we feel it demands further consideration in the combat setting to avoid unnecessary morbidity and mortality in the transfer of critically injured polytrauma patients.

As Role 3 receiving clinicians, we have observed several unfortunate scenarios which may portend worse outcome when secondary transfer is performed by inexperienced evacuation teams. It is well known that hypoxia and hypotension worsen traumatic brain injury and that lack of proper sedation and neuromuscular blockade may also aggravate intracranial pressure. Loss of airway control and/or accidental extubation en-route may be catastrophic, and each of these critical events correlates with poorer patient outcome. They may occur secondary to medical judgment or skill-based errors or as a result of the medical team's unfamiliarity with the airframe or medical equipment. These challenges are further compounded by the frequent need to fly a tactical profile, often at night, in a high threat environment.

Human Factors and Critical Care Transport

The occurrence of adverse prehospital events in civilian literature has been correlated with the skill mix of medical teams, their familiarity with the transport platform, and their ability to integrate with other aircraft crew. Such “Human Factors” and “Crew Resource Management” issues are familiar in aviation and deficiencies are known to frequently contribute to aviation mishap causality.

We feel that these facts support the existence of dedicated critical care transport or medical retrieval teams that are fully practiced in the transport environment, rather than being composed of ad hoc teams assembled for the occasional secondary transfer. This philosophy is encapsulated in guidance issued by the American College of Emergency Physicians, the Australasian College of Emergency Physicians, and the Intensive Care Society of Great Britain, and such medical teams are the norm in Australasia and Europe. MEDEVAC crews require additional education and regular training to ensure currency in procedures and situations they may encounter in flight. The analogy to aviation Crew Resource Management is obvious, where a lack of training, familiarity, or currency would present an unacceptable risk to flight safety.

We feel the aforementioned risks could be mitigated by the incorporation of dedicated, specially trained, critical care MEDEVAC teams into an integrated, theater-wide military trauma system. These teams could be based at Role 3 field hospitals or forward deployed as operations and tactics directed, offering short evacuation timelines and early provision of critical care while en-route to the most appropriate medical facilities. Incorporation of such “flying mobile trauma bays” could even conceivably reduce staffing requirements at existing Role 2s, by allowing the most valuable medical and surgical resources to remain at the Role 3. We believe

this model has the potential to improve patient care and safety while reducing costs and resource requirements associated with current battlefield trauma management.

CONCLUSIONS

In light of historical evidence, available resources and our own experiences in Helmand Province, Afghanistan, it is our opinion that the following recommendations be further investigated and seriously entertained by operational and medical commanders in OEF-Afghanistan/Op Herrick and future mature military theaters.

- (1) A thorough restructuring of the deployed military trauma system should be undertaken to integrate all echelons of medical care across the battlefield, taking into account the fixed and mobile prehospital capabilities of all participating nations and services. Current best evidence regarding the “golden hour” and other lessons learned should be applied to this endeavor and defended when met with operational resistance.
- (2) Dedicated, mobile prehospital critical care teams should be developed and employed for primary and secondary retrieval of the most seriously wounded combat casualties. Ideally, these teams should be physician-led and modeled after the highly experienced and successful international teams currently in existence. Thorough integration with the theater-wide military trauma system is essential.
- (3) Critically wounded polytrauma victims should be retrieved by dedicated prehospital critical care teams and triaged to the highest and/or most appropriate level of medical care available within the region, in keeping with current civilian trauma systems and best evidence.

BACK TO TOP

## Veteran Experiences Related to Participation in Shared Medical Appointments

Military Medicine
Shannon Cohen, PhD; Sarah Hartley, PhD; Jasbir Mavi, MD; Bridgette Vest; Monica Wilson
November 2012

ABSTRACT

Shared medical appointments (SMAs) are an innovative way for multidisciplinary teams to work together to meet veteran needs and encourage adherence to healthy lifestyle recommendations. Objective: The purpose of this study was to explore the experiences of veterans who participated in SMAs. Method: Focus groups were utilized to obtain information about SMA experiences. This method encouraged veterans to expand on and clarify the meaning of their experiences. Audio recordings were transcribed and analyzed by the researchers using N'Vivo software and an exploratory process to obtain consensus about themes. Results: The following themes emerged as a result of the focus group analysis: “empowerment, peer support, awareness, positive provider characteristics, teamwork, benefits, and convenience.” Conclusions: Veterans reported improvement in their overall health and well-being, improved self-management skills, and satisfaction with the SMA format. Veterans reported feeling empowered to improve their health and described a deep connection with their peers and group leaders. The connection they experienced with other veterans in the SMAs was similar to the close-knit relationships held with other members of their military unit.

## INTRODUCTION

The nationwide prevalence of overweight and obese veterans is estimated at 73%. Among veterans who are obese, 59% report a sedentary lifestyle placing them at risk for metabolic disorders. In addition, lifetime smoking rates among veterans are over 60%. Thus, the Veterans Health Administration emphasizes the need for programs that address these risk factors in ways that meet veterans' unique needs. Shared medical appointments (SMAs) are an innovative way that health care teams can work together to meet patient needs and encourage adherence to healthy lifestyle recommendations. This study explored the experiences of veterans who participated in the following SMAs at the Department of Veterans Affairs Medical Center (VAMC) in Salem, Virginia: managing overweight/ obesity for veterans everywhere (MOVE), metabolic assistance group intervention clinic (MAGIC), and smoking cessation.

### Obesity and Physical Activity

Obesity is a complex issue necessitating sustained lifestyle and behavioral change. Obesity is associated with hypertension, diabetes, hyperlipidemia, stroke, cardiovascular disease, gallbladder disease, osteoarthritis, numerous types of malignancies, and higher all-cause mortality. Adults who are obese have increased rates of hospitalizations and outpatient appointments, higher prescription medication cost, decreased quality of life, impaired body image, and social stigmatization in addition to increased risk for the diseases noted above. When patients who are obese lose 10% of their body weight, life expectancy increases up to 7 months, and the cost of prescription medication and medical care is reduced. Overall fitness and activity level also affects mortality; physical activity helps control weight and improves cardiovascular fitness, overall mood, self-image, and perception of self-efficacy.

Obesity Among Active Military and Veterans

Veterans are at high risk for obesity with only 27% considered a normal weight, defined as a body mass index (BMI) less than 25. Exposure to extreme stress during combat, physically demanding tasks and training, and eating behaviors learned during military service impact veteran weight as they age and become less active. During training and combat, military personnel ate large amounts of high calorie foods laden with fat and carbohydrates as a survival mechanism and often began smoking to combat hunger. Prisoners of war and soldiers engaged in combat have reported that safe food and water may be scarce affecting long-term eating behaviors including: food cravings, binge eating, nighttime eating, hiding foods, and preference for comfort food or sweets. McNulty also reported use of laxatives, diuretics, and diet pills; vomiting, and fasting before body fat measurement and fitness evaluations among male and female active military.

In a 5-year longitudinal study of 1,865 veterans, over 38% of the sample was obese with high rates of cardiovascular disease and diabetes. The annual BMI mean for the sample remained close to 29.2 despite weight loss advice and intervention, and the highest BMI values increased from 57.3 to 67.6. The increase in BMI experienced in middle-aged veterans is particularly disturbing given the emphasis on fitness and body fat standards in the military; a part of this steady weight gain may in fact be attributable to those same standards.

Smoking Cessation

Tobacco use leads to the deaths of 4 million people each year and is predicted to cause 10 million deaths worldwide annually by 2030. In addition to the risk of cancer and lung disease, smokers double their risk of a fatal cardiac event and have higher rates of diabetes. Smoking cessation reduces the risk of mortality from cardiovascular disease more than any other intervention and is second only to hyperlipidemia as a major risk factor for myocardial infarction. Thus, health care providers are challenged to monitor for risk factors, assist patients with smoking cessation, and prevent tobacco use among younger patients.

Shared Medical Appointments

SMAs use multidisciplinary patient centric methods which are effective in improving outcomes associated with chronic disease. In an SMA, a health care provider (e.g., physician) may be joined by a psychologist who specializes in health behavior change, a registered dietitian, and a nurse. Teams may also include social workers, physical therapists, kinesiotherapists, or other members. The patient who struggles with adherence to health recommendations and requires intensive behavioral management may be the ideal candidate for an SMA. Studies on SMAs report increased patient access to care, decreased unnecessary emergency room visits and outpatient appointments, increased provider productivity, higher levels of patient and health care provider satisfaction,

enhanced chronic disease self-management, and improved adherence to health care recommendations. In this study, patients reported feeling validated and supported and realized that they were not alone in their concerns. They enjoyed sharing their strategies on coping with a chronic illness and felt they were a part of the care giving process.

for a list of the advantages of participation in SMAs over traditional medical appointments and the additional reading section for more information.

MOVE, MAGIC, and Smoking Cessation Programs

MOVE, MAGIC, and smoking cessation are three SMAs offered to veterans at the VAMC in Salem, Virginia. These collaborative programs include experts in primary care, health behavior change and mental health, nutrition, exercise, and smoking cessation. The main focus of the MOVE program is nutrition, weight loss, and increasing physical activity. The MAGIC program focuses on diabetes, hypertension, weight control, and hyperlipidemia management. These programs incorporate motivational interviewing techniques and address depression, anxiety, stress management, and coping strategies. The content of these programs overlap and complement each other.

## METHODS

### Purpose of the Study

The purpose of this study was to explore patient satisfaction and perception of value related to participation in the following programs: MOVE, MAGIC, and smoking cessation.

### Sample, Methods, and Data Analysis

The Veterans Health Administration provides medical care for over 5.5 million veterans nationwide. Veterans are older with more health problems and fewer financial resources than the civilian population. Approximately 112,500 veterans living in a 26-county area of southwestern Virginia receive outpatient medical care at the Salem, Virginia facility. Seventeen people participated in the focus groups between September 2011 and January 2012 out of the 145 veterans who were contacted. The mean age of the sample (*N* = 17) was 62 with a range of 39 to 85. Ninety four percent of the sample (16) was male. Ethnicity of the sample was closely divided between Caucasians (9, 53%) and African Americans (8, 47%). The majority of this purposive sample was unemployed or retired (12, 70.6%). Sampling continued until all of the researchers agreed that saturation had been met and no new insights would be

identified. Inclusion criteria for this study included current enrollment, English speaking, adequate ability to hear, and under age 89 with documentation of participation in SMAs.

De-identified demographic data and information pertaining to class attendance was extracted from the VAMC electronic medical record. Five focus group sessions were held at the facility and lasted approximately 1 to 1.5 hours each. There was no monetary compensation for participation in this study. Participants complied with instructions not to use their real names during the focus group sessions and informed consent was obtained. Notes, memos, and transcribed audiotapes of the focus group sessions identified participants by coded number. Participant responses were coded and analyzed to observe consistent themes. Audio recordings were transcribed and analyzed by the four researchers using N'Vivo software and an exploratory process to obtain consensus regarding themes. The inter-rater reliability was found to be $K = 0.68$ ($p < .0.001$).

Focus groups were utilized to obtain information related to moderator questions but also included the unique interaction between group members who encouraged others to tell stories and examples and clarified the meaning of their experiences. Interaction between group members allowed deeper exploration of complex topics than individual interviews or surveys and allowed participants to consider their thoughts and experiences within the context of others' views. The open-ended focus group questions were designed to encourage conversation and additional details and clarification from participants.

## RESULTS

The following themes emerged as a result of the focus group analysis: “empowerment, peer support, awareness, positive provider characteristics, teamwork, benefits, and convenience.” In addition, participants had suggestions for improvement of the SMAs. Quotes were chosen that researchers felt best described each theme. Participants reported feeling empowered to improve their health and described a deep connection with their peers and group leaders. The connection they experienced with other patients in the SMAs was similar to the close-knit relationships held with other members of their military unit.

### Empowerment

A majority of participants identified empowerment as being a positive factor in the SMAs. They learned self-management skills, gained knowledge about their health, and increased their self-discipline. Participants stated that the SMAs increased their motivation to make behavioral changes, their self-awareness of how their behaviors affect their health, and their determination to take control over their health. This sense of empowerment acted as a catalyst to behavior change. One participant reported that “my success was due to going to the meetings” and another stated “it's knowledge you can take home and share.” The SMAs also increased

participants' self-confidence in their ability to change. This was exemplified in statements such as “[the SMA] gave me the knowledge to know what I should do,” and one participant reported “I can do this.”

Peer Support

Another theme that arose during the focus groups was peer support, which is a key factor in promoting health and well-being, and increasing adherence to recommended health behaviors and self-management. Learning what other veterans had experienced and “tips” on chronic disease self-management provided a much needed perspective for many. Veterans stated that they felt accountability to the group, peer pressure to reach their goals, and assistance, support, and encouragement from their peers. Experiencing a connection with veterans with similar health conditions, socialization, and an “esprit de corp” was a benefit of participating in the SMA. Statements reflecting this group cohesion and support include

“You feel that you are the only one and then you share and find someone else.”

“The program gave me a real connection.”

“Misery loves company.”

“I felt obligated to weigh less each week.”

“There used to be a guy here… he lost like 40 some pounds and we really clapped for that fella cause he really worked hard for that. And it gave us something to try for.”

“[The group] boosted my determination in a way. Because after, once you start going you meet these other Vets and you know they give you support and you give them support. It's much easier to go back.”

Awareness

Researchers debated the meaning of the theme of awareness as participants described this in two very different ways. Participants reported awareness as increased self-monitoring as a result of their involvement in SMAs with self-management as the larger construct, which was ultimately placed within the theme of empowerment. Participants reported feeling empowered to participate in self-care, a greater sense of self-awareness and knowledge of the behaviors affecting their health, and a desire to share knowledge gained with others. In this study, awareness was defined as “getting the word out”; advertising the merits of the SMAs with other veterans and staff.

Positive Provider Characteristics

In the focus groups for all three SMAs, veterans identified characteristics of the SMA providers that they found beneficial in receiving quality health care. They felt the providers were experts in their field, as well as personally invested in their care. Many indicated that they felt a deep concern and care from the SMA staff members, discussing that staff were “always encouraging,” “caring and compassion[ate],” and “really concerned about me.” Participants also commented that providers were “knowledgeable,” “specialists,” and “they know what they are doing.” They reported that providers were supportive advocates who knew them as a person and encouraged them to share in the decision-making process.

Teamwork

A theme of teamwork arose within all focus groups. Patients felt that they were part of a dynamic, cooperative partnership with their health care team. They appreciated collaborating with the staff on goal setting and shared decision making and they felt that lines of communication were always open. Veterans also commented on the teamwork they witness among the SMA health care providers: “combined together they come up with a plan that works for you,” and they “work together as a team.”

Veterans who participated in the focus groups reported that they received individualized care from the team. Patients felt that providers worked with them to meet their individualized needs. They reported that “they work with you,” “you are not just a number,” and they “really met my individual needs.”

Benefits

Another theme that came from the focus groups was the wealth of personal benefits that patients received from attending the SMA. Most veterans identified health improvements from participating in the SMAs, including quitting smoking, reduced blood pressure and glycated hemoglobin levels, smaller waist circumference, increased physical stamina, and weight loss. One Veteran stated that the SMA “made a big difference in my health and well-being.” In addition to health care benefits, a few veterans identified financial benefits to attending the SMA. The group reported saving money by quitting cigarettes and limiting fast-food intake. One Veteran stated that he believed others attended SMAs in order to receive travel pay.

Convenience

The focus group questions specifically addressed access and convenience. Participants reported that groups were, overall, convenient for them; however, they did indicate that access, location, and wait times could be improved. One member commented

that they experienced an “occasional long wait, but it was worth it.” They suggested that the groups be offered “at other times, more often” and “get the meetings in the evenings, have more meetings so that people can attend.”

### Suggestions

Additional suggestions made by the participants included increasing advertising about SMAs and making adjustments to group formats. Suggestions included “involve veterans in teaching,” being “more hands-on,” and “e-mail or call [to] follow-up.” Several veterans enthusiastically volunteered to attend future group clinics to share their chronic disease self-management experience. They also volunteered to call other veterans to help them with smoking cessation.

## DISCUSSION

Veterans reported many benefits to participating in SMAs and were eager to share their experiences and opinions on these programs. Focus groups allowed for a deeper discussion of the needs of our veterans. Veterans reported increased knowledge of their health conditions, which led to feeling empowered and in control of their health. This is consistent with other research studies that have found SMAs improve knowledge when education is provided in a group setting and the increased knowledge and confidence gained leads participants to make health behavior changes.

Veterans reported making behavior changes including tobacco cessation, portion control, increased activity, and making healthier dietary choices as a result of participating in SMAs. Sharing experiences and socializing with peers has been found to be a benefit of SMAs in the general patient population. This previous finding was supported by the current study, in which patients identified peer support as being a characteristic of the SMA that they found helpful. Multiple quantitative research studies have found that SMAs significantly improve health outcomes in patients who participate. In this qualitative research, patients reported that they experienced improved health, including lower blood pressure, reduced glycated hemoglobin levels, and weight loss among other benefits.

Despite the fact that many of the focus group participants had never met, they were very vocal and interacted easily. They allowed others to speak freely without interruption and respected other members' opinions. The researchers were surprised that the men expressed a lot of emotion and appreciation for the group and leaders. Participants reported “ribbing and teasing each other about things we should not be eating,” which added additional accountability to the group. More accountability among the MOVE weight-management group was noted as the group has more consistent members.

### Strengths and Limitations

There were several strengths of this focus group study. Focus groups provide an opportunity for researchers to interact with participants and pose follow-up clarifying questions. Additionally, focus groups allow participants to share their experiences and become engaged with each other, fostering deeper discussion of their needs. In addition, by keeping the questions open, participants were allowed to bring up topics and SMA elements that the researchers may not have considered before this study.

A balanced view was sought from the three types of SMAs; however, “group think” (agreeing with the most popular opinion in the group) was a risk in using focus groups as a method. Participation was voluntary, which could have biased our sample of participants. Veterans who were satisfied with the SMAs may have been more willing to attend a focus group to discuss their experiences and may have given answers that they expected researchers wanted to hear. Last, this small sample may not be generalizable to a larger population or to a civilian patient population.

### Future Research

The findings from this study provided valuable insights about the experiences of veterans participating in SMAs. As SMAs become more numerous and begin to be used with other patient populations, it is important to continue quantitative and qualitative research into participants' experiences and outcomes. Future research could be done on SMAs that use different formats with aggregate data utilized to identify best practices in executing quality SMAs. Future focus group research designed to examine the experiences of health care providers in these groups would be beneficial to ensure provider satisfaction and prevent burnout. In addition, the type of woman who is comfortable being in a mixed gender group would be interesting to explore.

## CONCLUSIONS

SMAs allow health care teams an innovative way to meet Veteran needs and encourage adherence to healthy lifestyle recommendations. Patients reported high levels of satisfaction with SMAs as well as positive health outcomes. The following themes emerged as a result of the focus group analysis: “empowerment, peer support, awareness, positive provider characteristics, teamwork, benefits, and convenience.” As one Veteran stated “you always have a choice. It is all about changing your behavior and with the support of the group you can do it.”

BACK TO TOP

# Is the ASVAB ST Composite Score a Reliable Predictor of First-Attempt Graduation for the U.S. Army Operating Room Specialist Course?

Military Medicine
MAJ Joel Grant, MS AGR; Capt Angel L. Vargas, USAF MSC; MAJ Robert A. Holcek, AN USA; MAJ Carolyn H. Watson, AN USA; Jessica A. Grant, MA1; MAJ Forest S. Kim, MS USA
November 2012

ABSTRACT

The U.S. Army Operating Room Specialist (68D) Course provides first class medical technician training to U.S. Army enlisted soldiers of the Army Medical Command. With a failure rate of approximately 12% over a 2-year period, this study was commissioned to determine whether the Armed Services Vocational Aptitude Battery (ASVAB) skilled technical (ST) Score served as a reliable predictor for successful first-attempt completion of the 68D course. A sample size of 373 was analyzed via a multivariate binary logistic regression model with 6 distinct independent variables. This study found that the ASVAB ST score, gender, and rank were predictors to first-attempt successful completion of the 68D training program. Specifically, students with an ST score 10 points higher than their peers were 5 times more likely to graduate. In addition, females were 2.5 times more likely to succeed than males and Army Privates (E2) were 3.2 times more likely than Privates (E1). Specialists, Corporals (E4), Sergeants (E5), and Staff Sergeants (E6) combined, were 34 times more likely to succeed than E1s. Although further study may be warranted, increasing the minimum ST score requirement in the admission guidelines and/or specific preventive assistance for lower-ranked students may decrease the first-attempt failure rate.

## INTRODUCTION

The Army Medical Department Center and School (AMEDDC&S) manages over 250 medical and medically related administrative training programs for commissioned officers, warrant officers, enlisted soldiers from all military branches of service, civilians, and international military forces (J.B. Lavender, personal communication, March 16, 2011). The AMEDDC&S course offering includes the Operating Room Specialist (68D) Course. Upon graduating from the course, students earn the enlisted job title, or military occupational specialty (MOS), of 68D. The demanding 68D course provides students with the fundamental knowledge and principles of the operating room and sterilization process. The 68D course is open to most soldiers and consists of two phases. The first phase

is didactic, which lasts for 9 weeks, and it is located at Fort Sam Houston, Texas. The second phase consists of clinical on-the-job-training, which lasts for 10 weeks and may take place at one of the 14 locations.

In 2009 and 2010, classes in the 68D course suffered a first-attempt failure rate, which ranged from 4% to 24% (M.E. Underwood, personal communication, March 10, 2011). The terms “first-attempt completion” and “success” are used interchangeably. Students who failed the program were placed into one of three categories: recycled, reclassified, or separated. Recycled students are those placed in a subsequent 68D course to reattempt completion. Reclassified students are those who, upon failure, are recommended for transfer or reclassification to a different MOS-producing course, based on the needs of the Army. Separated students are those removed from the course because of academic or nonacademic issues resulting in separation from the service.

To enter the 68D course, certain criteria must be met. One criterion is completion of the Armed Service Vocational Aptitude Battery (ASVAB). The ASVAB is the entrance exam used by all services for the selection and classification of enlisted applicants. Its lineage can be traced back to 1917. Today, more than one million people take it annually.

The Army selects and then trains and develops enlisted soldiers based on their potential. Considerable resources are spent on training each recruit. The ASVAB is intended to provide a measure of future success within MOS-producing courses. The ASVAB score as a determinant for eligibility for acceptance into a specific MOS is one of several important criteria to review when analyzing attrition rates.

The ASVAB produces nine subtest scores across multiple skill areas. These scores provide a direct assessment of an examinee's aptitude on specific tests. These sub (skill) test scores are subsequently combined into various line (composite) scores with alternate measurement scales. Although the total number of composite scores and the corresponding measurement scales vary among the five military branches, the Army calculates ten distinct composite scores. These scores are designed to best determine if a soldier will be successful in a respective military career field. For this study, the “skilled technical” composite (ST score) was of particular importance. The ST score is currently utilized as one of the determinants of eligibility for the 68D course. Students are currently required to achieve an ST score of 91 or higher to be eligible for the 68D course. The maximum possible ST score has historically been as high as 160 (D. Mills, personal communication, June 19, 2012). The ST score subtest components are General Science (GS), Mathematical Knowledge (MK), Mechanical Comprehension (MC), and Verbal (VE).

## LITERATURE REVIEW

In a 2002 study, it was found that the ASVAB General Technical composite (GT score) was a reliable predictor of course completion for the Army's Pharmacy Specialist Course (68Q). This course is often referred to as the 68Q course and was previously known as the 91Q course. Furthermore, this study found that the ST score was not a reliable predictor for the 68Q course. However, significance was found with respect to the total course score.

Research on the ASVAB's subtests and subsequent line scores has been inconsistent. Carretta and King found the ASVAB to be “a strong predictor of training performance.” The ASVAB's ability to accurately describe cognitive ability has been reproduced in numerous studies. Porter et al found that the Air Force (AF) ASVAB Mechanical composite (M) score was a better predictor of success over the AF ASVAB General Maintenance composite (GM) score for airmen attending the AF's Medical Service Apprentice (MSA) program. The MSA program trains AF medics and is comparable to the Army's “combat medic” (68W) program. This finding was later reinforced by a similar study involving the AF's Radiography Course, which trains X-ray technicians. Here, too, the M score was found to be a predictor of success, as was the Electrical composite (EI) score, which held true when evaluated together and when evaluated independently. Jordon and Curtis additionally found increased M scores correlated to higher student final exam scores for the AF's Vehicle Maintenance program.

In 2008, a study by Fulton et al reported three major findings regarding ST score and student attrition in the Army's 68W course. First, this study found that “most importantly, the student's ST score was the single most significant forecast variable for determining success or failure.” Furthermore, only 59% of the students who attempted the 68W course, with an ST score less than 94, successfully completed the course. Second, this study found that only 44% of students who were reclassified into the 68W course ultimately graduated an MOS-producing program. Finally, rank was found to be a determinant of whether or not a student would pass or fail the course. The findings resulted in recommendations to increase the minimum ST score because it would likely reduce the number of soldiers who would fail the course. Additionally, results indicated that increasing rank at admission would improve graduation rates.

As a result of the literature review, this study considers the question: Is the ASVAB ST composite score a reliable predictor of first-attempt graduation from the 68D Course? The null hypothesis is that the ST score is not associated with first-attempt graduation from the 68D course. The alternate hypothesis is that the ST score is associated with first-attempt graduation from the 68D course.

## CONCEPTUAL MODEL

Before conducting the analysis, a variety of factors were examined that, as shown in the literature, have the potential to affect a student's successful completion of the 68D course. Consistent with previous literature and in correlation with categorizations often utilized within the military, these variables.

Individual factors are defined as those factors specific to the individual's demographics, lifestyle choices, characteristics, or status attributed at the individual's level. Demographics include race, gender, age, and so forth. Lifestyle choices include, but are not limited to, personal study habits and work experience. Status attributes include, but are not limited to, rank, high school grade point average, post–high school education, socioeconomic status, and ASVAB test scores. Many of these singular factors have a potential effect on an individual's motivation, which can influence one's ability to complete the course successfully.

Systemic factors refer to those factors found within or as a result of the 68D course itself. One example includes the instructor's teaching style. These factors are both subjectively qualitative and quantitative in nature.

Nonsystemic factors include all other influences not specifically tied to the program or individual. An example of this includes military law violations, which are known as Uniform Code of Military Justice violations. It should be noted, however, that nonsystemic and systemic factors can, in fact, influence individual factors in a positive or negative manner and vice versa.

In summation, any variable found within this model for individual, systemic, and/or nonsystemic factors, may potentially influence an individual student's successful completion of the 68D course on his or her first attempt.

## EMPIRICAL MODEL

Limitations in data availability prevented inclusion of all individual, systemic, or nonsystemic factors. As explained in the conceptual model and a result of several limitations, only six variables were analyzed in this study. Those variables were ASVAB ST score, ASVAB GT score, gender, service component status, nonprior service status, and Army enlisted rank. Although ST and GT scores were found to be reliable predictors in earlier studies, outside the general mechanic score for the AF, other ASVAB test score components were not found to be significant and were, therefore, excluded. In addition, a separate bivariate analysis of the remaining composite scores indicated individual skill overlap or heavy correlation, most likely attributable to the subtest duplication found in many of the composites. Nonprior service status indicates whether the student previously served in the U.S. armed forces. Service component status is separated into two categories, active duty and reserve duty. Active duty soldiers serve as full-time soldiers upon completion of their training. Reserve duty soldiers are composed of Army Reserves and Army National Guard soldiers

that serve on a part-time basis approximately 30 days per year. No National Guard soldiers were included in this study's data sample.

The research study utilized was cross-sectional in nature and utilized a secondary data source. The dependent variable was first-attempt course success or failure. The variables were examined through a multivariate binary logistic regression analysis with ST score as the primary independent variable and gender, rank, service component status, nonprior service status, and GT score as the independent control variables. The research design included 1-group post-test only; therefore, threats to validity included data history and interaction of selection and maturation. The unit of analysis was the individual student.

The initial data set contained 434 individual enrollment samples from seven distinct classes held in 2009 and 2010. Accounting for repeat attempters, there were 406 distinct students. To further delineate the inclusion criteria of the data set, students were categorized into 3 categories: passed, failed: academically dismissed (ACD), and failed: non-academically dismissed (non-ACD).

Exclusion criteria included recycled secondary entries, non-ACD students, and sets containing missing variables. First, secondary recycled entries accounted for 18 individuals. Next, 5 non-ACD entries were removed. Finally, incomplete or missing data accounted for 28 students who were removed. This resulted in a final sample size of 373 individuals.

The seven variables used in this study were delineated as follows: four were binary (graduation status, gender, service component, and prior service); two were continuous variables (independent variable ST and GT), and one was categorical (rank). Categorical ranks were converted into dummy binary variables. Categories were defined as: Private (E-1); Private (E-2); Private First Class (E-3). Because of the limited number of each, Specialist (E-4), Sergeant (E-5), and Staff Sergeant (E-6) were combined into one category.

## RESULTS

Descriptive Statistics (Continuous Variables) provide the descriptive statistics for this analysis. Of the 373 students analyzed, 327, or 87.7%, successfully completed the 68D course on their first attempt. In addition, 316, or 84.7%, of students were nonprior service. Gender results slightly favored males at 196, or 52.5%, to females at 177, or 47.6%. Reserve duty students outnumbered active duty students 227 to 146 or 60.9% to 39.1%, respectively. Finally, there were 88 E1 (Privates) who composed 23.6% of the sample, 95 E2 (Privates) who composed 25.5%, and 120 E3 (Privates First Class) who composed 32.2%. There were 58 E4 (Specialists/Corporals) who composed 15% of the sample and 12 E5 (Sergeants) and E6 (Staff Sergeants) who composed 3.2% of the sample. The combined E4, E5, and E6 ranks totaled 70, or 18.7%, of the sample. For the purpose of this study, the term rank for both the actual rank (e.g., Private) and for their pay grade (e.g., E1) is used interchangeably. All three rank categories were then compared to E1

Privates, this study's rank reference category. The failure rate for each independent variable varied significantly. Specifically, 25% of all E1s, 16% of active duty, 15% of males, and 14% of nonprior service failed this course. The range for ST and GT were 87 to 141 and 82 to 139, respectively.

Based on the Hosmer and Lemeshow Test, the significance of this regression model was 0.812, which indicated it was not a poor fit. With respect to shared variance, or $R^2$, the overall model was 14.6%. In other words, one-seventh of the variance found in successful first-attempt course completion was accounted for by the ST ASVAB Score and other control variables found in this model.

Binary Logistic Regression Outputs, the independent continuous variable ST and the independent control binary variables of gender, rank E2, and rank E4/E5/E6 were significant variables that served as predictors for successful completion of this program. The results imply that the ST score ($p < 0.0001$) was an accurate predictor, and for every 10 unit increase in ST score, the odds of first-attempt course completion increased by five. Next, females were 2.52 times more likely to pass than males ($p = 0.01$). With respect to rank, when compared to study reference category rank E1, rank E2s were 3.28 times more likely to pass ($p = 0.009$). In addition, E4s and above, who are generally more experienced, were 34 times more likely to pass than E1s ($p = 0.006$). Rank E3, approaching significance at $p = 0.057$, were 2.18 times more likely to pass on the first attempt than E1s.

## DISCUSSION

The results of this study confirmed that the ASVAB ST score is a significant predictor of success in the 68D course. GT score, however, was not found to be significant. These findings contradict Meadows et al study from 2002, which found the ASVAB GT score to be the predictor and ST not to be a reliable predictor of success. However, the findings are consistent with Fulton et al. Therefore, it is probable that raising the ASVAB ST threshold would lead to increased pass rates. In addition, since higher-ranked students (E2, E4, E5, and E6) performed better than E1s, special attention should be paid to the higher-risk students. Higher-ranked students most likely performed better because of experience, maturity, study habits, and/or internal and external motivational factors. This adaptation to the military way of life promotes increased motivation among these students. Last, female students scored significantly higher ($p < 0.01$) than their male counterparts, which was consistent with the findings of Meadows et al. This is also consistent with gender-related school performance research.

## LIMITATIONS

The primary goal of this study was to inquire whether or not ST score was a predictor of first-attempt success for 68D students. Although the study results answered this question, numerous qualitative and quantitative factors may have been overlooked.

First, exit interviews were not examined and student admission statistics were not comprehensive. Subsequently, data on other individual factors, such as previous education level, socioeconomic factors, race, and so forth were not readily available for analysis. With 46 failures over a 2-year period, these interviews could have easily been accomplished and would have been quite valuable to this study. Next, instructors may have a significant impact on student performance and this study was unable to account for variation in systemic factors such as instructor turnover, teaching abilities, and styles.

Finally, and perhaps most importantly, the remaining individual and systemic factors not included from the conceptual model were not taken into account during the analysis. For example, an analysis of rank would have been more useful if age was also provided. An 18-year-old E1 versus a 24-year-old E1 would most likely have a maturity differential that could have provided additional significance to the model.

## RECOMMENDATIONS

Further research that will minimize the effect of the aforementioned limitations is recommended in order to provide a more accurate understanding of why students fail the 68D course on their first attempt. In addition, four additional steps would greatly benefit future studies as well as the 68D course itself.

First, comprehensive and in-depth qualitative exit interviews should be completed with all students that fail. The authors believe this would provide crucial information and insight for reducing course failure rates, as well as creating data points for several items found in the conceptual model. In addition, in-depth qualitative self-reporting surveys could also provide benefits that may preempt any potential failures.

Second, the program should add a precourse and postcourse assessment exam. These assessments could serve as a tool for future analyses by determining where a student's general knowledge stood initially and how it improved, or did not improve, by the end of the course.

Next, E1 was found to be the grade with the most first-attempt failures. The program could potentially prevent failures by providing key mentorship to this specific group. With an obvious correlation, the cohort or staff could address any key warning indicators early on and prevent potential failure. In addition, the program could raise admission criteria to E2 or above.

Finally, increasing the minimum ST score requirement could potentially decrease failures. However, before that recommendation is implemented, more study and analysis of other ASVAB skills and composite scores may be required to provide more useful information. Despite the fact that this study found the ST score to be a significant predictor of course success, there may be a better predictor available.

BACK TO TOP

## Stretch and Wrap Style Tourniquet Effectiveness With Minimal Training

Military Medicine
Piper L. Wall, DVM, PhD; John D. Welander, MD; Amarpreet Singh, BA; Richard A. Sidwell, MD; Charisse M. Buising, PhD
November 2012

ABSTRACT

The objective was to determine if proper application of the Stretch, Wrap, and Tuck Tourniquet (SWAT-T) would stop arterial flow and would occur with minimal training. Methods: Fifteen undergraduates watched a 19 second video three times, practiced twice, and applied the tourniquet to volunteers at 10 locations: 3 above the elbow or knee and 2 below. Results: Successful occlusion (60 second Doppler signal elimination) was more frequent than proper stretch (96 versus 75), more frequent on arms than legs (59 versus 37), and achieved before completed application (16 ± 8 versus 33 ± 8 seconds; each $p < 0.05$). Proper stretch (correct alteration of shapes printed on the tourniquet) was more frequent on legs than arms (30 versus 45; $p < 0.05$). Applications were rated Easy (101), Challenging (37), Difficult (12) with discomfort None (53), Little (62), Moderate (34), Severe (1). The 8 appliers with <70% proper stretch rates received 10 minutes additional training and then retested at mid upper arm, mid-thigh, and below knee (24 applications) for improved proper stretch and occlusion (5 versus 18 and 10 versus 20; $p < 0.01$). Conclusions: Proper application of the SWAT-T is easy and can stop extremity arterial flow but requires some training for many appliers.

INTRODUCTION

Tourniquets that stop injured extremity arterial flow are lifesaving devices in combat-related situations[1] and may be lifesaving in some civilian situations. Beyond stopping arterial flow, military tourniquets must be small and light enough to carry where they are rapidly accessible on every soldier and must be easy to apply correctly.

The tourniquets deployed with U.S. servicepersons since late 2004 are a windlass/stick and strap design, the Combat Application Tourniquet (CAT; Composite Resources, Rock Hill, South Carolina). At 3.8-cm wide, this tourniquet may require pressures >300 mmHg to stop thigh arterial flow (tourniquet width/limb circumference 0.048–0.083). A wider (10.4 cm) stretch and wrap style tourniquet, the Stretch, Wrap, and Tuck Tourniquet (SWAT-T; TEMS Solutions, Abingdon, Virginia) became commercially available in 2008. Published data exist concerning the lab and field effectiveness of the CAT; however, such data are currently lacking for the newer SWAT-T.

Like the CAT, the SWAT-T meets military size and cost criteria to be a tourniquet of interest and is intended for use in tactical, limited supply scenarios or first responder casualty care by soldiers, military medics, police, etc. Individuals in some of these groups receive limited medical training, and “simple to apply” and “use with little to no training” are listed by the military as ideal emergency tourniquet traits. Marketing for the SWAT-T claims effectiveness (rapid control of extremity bleeding) and “Its ease of application is one of its greatest benefits; individuals can effectively apply it in seconds with little to no prior training.” (SWAT-Tourniquet, http://swattourniquet.com) We found both claims interesting; so, our hypotheses for this study were that (1) proper application of the SWAT-T would stop arterial flow through each extremity location and (2) college undergraduates would be able to properly apply this tourniquet at each location with very minimal training.

## METHODS

This prospective study was approved by the Drake University Institutional Review Board. The purchased SWAT-T tourniquets were 95 g, 10.4 cm wide, 11.7 cm circumference rolled up, 150 cm unrolled, only 1 part, and US $8.50.

### Subjects

Fifteen volunteer tourniquet appliers each met the following criteria: undergraduate student taking a course involving physiology-related research with no prior tourniquet-related training. Appliers could be tourniquet recipients, but not before being an applier. Each applier was the same sex as the respective recipient. The 15 volunteer recipients each reported meeting the following criteria: no clotting or circulation abnormalities, no blood pressure problems, and no pain syndromes or peripheral neuropathies.

### Protocol

For a very minimal amount of SWAT-T application training, the applier group watched 19 seconds of a thigh application video 3 times (seconds 14–33, Thigh Application of SWAT-T, http://www.youtube.com/v/KypMY7Ng6ak via http://swattourniquet.com/videos.html)

with the audio portion silenced. The first viewing was silent. The second viewing included brief comments to note (1) how the tourniquet secures to itself after the first circumferential wrap, (2) the geometric pattern change from diamonds inside oblongs to squares inside circles when the tourniquet is properly stretched, and (3) the end tucking to secure the tourniquet (these comments were the only auditory instruction given). The third viewing was silent. Appliers were then allowed one application to the training individual's upper arm and one to the training individual's thigh.

Two days later, appliers and recipients (sitting in T-shirts and shorts) completed the study protocol:

- (1) Applier and recipient characteristics were collected. Heart rate and systolic and diastolic pressure were obtained with a wrist automatic blood pressure monitor (Model MF-77, Mark of Fitness, Shrewsbury, New Jersey). Occlusion pressures were the manometer pressures of the pneumatic blood pressure cuffs (arm or thigh) when the distal arterial Doppler pulse signal became inaudible (wrist radial artery or ankle posterior tibial artery).

- (2) Tourniquet application order was high upper right arm, mid upper left arm, just above right elbow, just below left elbow, mid right forearm, upper right thigh, mid left thigh, just above right knee, just below left knee, and mid right calf. Upper thigh application was over the shorts; the rest were on skin.

- (3) Application time-related data were collected: time to loss of the distal audible Doppler signal (radial artery or posterior tibial artery), time to completed application, and time to return of the audible Doppler signal if it occurred before 60 seconds following application completion (60 second sustained loss of the audible signal was labeled Doppler success and considered successful arterial occlusion).

- (4) After each application, a rating was assigned concerning proper tourniquet stretch (proper application). The conversion of diamonds inside oblongs to squares inside circles (shapes printed on tourniquet) was considered proper stretch. The person making the rating had an example picture of properly versus inadequately stretched tourniquet geometric shapes and determined the adequacy of stretch during the process of application.

- (5) Before removal, an ease of application rating (Easy, Challenging, Difficult) and a recipient discomfort rating (None, Little, Moderate, Severe) were obtained. Both relate to the scales used by Swan et al.

- (6) The tourniquet was removed 60 seconds after application completion.

- (7) The tourniquet was applied at the next location 2 minutes later.

Additional work was done to determine whether failures to properly stretch the tourniquet were related predominantly to training and technique or to strength. First, appliers who achieved a <70% proper stretch rate received additional training: 5 additional applications each accompanied by verbal feedback concerning achieving and maintaining adequate stretch while wrapping (approximately 10 minutes additional training time per person). The two expected technique problems were failure to adequately stretch the tourniquet and failure to maintain adequate stretch while passing the tourniquet around the limb. If these were not strength problems, we believed 5 applications with verbal feedback would be sufficient for correction and allowance of a few correct technique practices.

The ability of these volunteers to properly apply the tourniquet was then reassessed. The reassessment was to determine whether these appliers had easily solvable technique issues and was not to check whether the tourniquet could be effective; therefore, only 3 sites were used: mid right upper arm, mid left thigh, and below the right knee. An arm location was chosen because arms tended to be more mobile during application than legs and therefore potentially more difficult for maintaining proper stretch. The upper arm was chosen because upper arm locations were reported as more common use than forearm locations (162 arm versus 13 forearm of 862 tourniquets on 499 patients reported by Kragh et al[1]) and, based on circumference, are more likely to require proper stretch to achieve cessation of blood flow than forearm locations. The thigh was chosen because it is the location with a U.S. military specified minimum desirable success rate. The thigh is potentially the hardest site at which to achieve occlusion because of circumference, but it may have been one of the easier locations at which to achieve proper stretch because of less limb mobility during application than encountered at other locations. Thigh locations were reported as the most common use location by Kragh et al[1] (436 thighs of 862 tourniquets). The below knee location was chosen because lower leg locations were the third most common site reported by Kragh et al[1] (46 reported).

The tourniquet reassessment recipients were from the original recipient group. Reassessment recipients were restricted to those with initially measured systolic pressures <140 mmHg.

The second method used to determine whether proper stretch failures were predominantly technique or strength issues was an assessment of each applier's pulling and pushing strength. Using weight machines, each applier's strength was assessed on a seated row and an overhead press.

Statistical Analysis

Parametric data (pressures, times, etc.) were compared using unpaired *t*-tests, paired *t*-tests, or two-way repeated measures analysis of variance. Contingency tables (Doppler success, proper stretch, etc.) were analyzed using $\chi^2$ or Fisher's exact test.

Graphing and statistical analyses were done using Microsoft Office Excel 2003 (Microsoft, Redmond, Washington) and GraphPad Prism version 5.02 for Windows (GraphPad Software, San Diego, California). Means ± SD are shown.

## RESULTS

There were 150 tourniquet applications. Seven appliers were recipients.

### Minimal Training Only

Tourniquet application was generally rated “Easy” (101 Easy, 37 Challenging, 12 Difficult). Doppler successes were more frequently associated with application ratings of “Easy,” and Doppler failures were more frequently associated with ratings of “Challenging” or “Difficult” (96 Doppler successes: 85% Easy, 13% Challenging, 2% Difficult versus 54 Doppler failures: 35% Easy, 46% Challenging, 19% Difficult; $p < 0.0001$). Proper or inadequate application stretch was not as associated with specific application ratings (75 properly stretched applications: 73% Easy, 20% Challenging, 7% Difficult versus 75 inadequately stretched applications: 61% Easy, 29% Challenging, 9% Difficult; $p = 0.29$).

Tourniquet application generally involved minimal recipient discomfort (53 None, 62 Little, 34 Moderate, 1 Severe). Doppler successes were more frequently associated with discomfort ratings of “Little” to “Moderate,” and Doppler failures were more frequently associated with ratings of “None” (Doppler successes: 24% None, 44% Little, 31% Moderate, 1% Severe versus Doppler failures: 56% None, 37% Little, 7% Moderate, 0% Severe; $p < 0.0001$). Properly stretched applications were more frequently associated with discomfort ratings of “Little” to “Moderate,” and inadequately stretched applications were more frequently associated with ratings of “None” (properly stretched applications: 20% None, 47% Little, 33% Moderate, 0% Severe versus inadequately stretched applications: 51% None, 36% Little, 12% Moderate, 1% Severe; $p = 0.0002$).

Average application times were <40 seconds for all locations (31 ± 6 seconds male, 34 ± 13 seconds female; $p = 0.02$). When it occurred, audible Doppler signal loss happened before completed application (16 ± 8 versus 33 ± 8 seconds; $p < 0.0001$). The time to completion of applications without audible Doppler signal loss was not significantly different from the time to completion of applications with audible Doppler signal loss (32 ± 14 seconds for completion without audible Doppler signal loss). The times to apply did not decrease as the number of applications increased (leg application times were not shorter than arm application times). There also did not appear to be any time effects related to tourniquet application to the left or right side of the body.

Appliers dropped the tourniquet 14 times. The dropped portion unrolled partially to completely. Completing the application with an unrolled tourniquet did not appear to add to the completion time or the difficulty (35 ± 11 seconds; 7 Easy, 6 Challenging, 1 Difficult).

Doppler success was more frequent than proper stretch, and both were more frequent for males ($p < 0.05$ for each; Thirteen of the Doppler failures involved resumption of an audible pulse before 60 seconds; the remaining 41 Doppler failures had no loss of the audible pulse.

Doppler success was more frequent on arms than legs ($p = 0.0003$) Proper stretch was more frequent on legs than arms ($p = 0.02$). Only once was proper stretch achieved on the arm without Doppler success (mid upper arm, male subject). Tourniquet placement high on the thigh had the lowest Doppler success rate (2 of 8 males, 1 of 7 females), and, when Doppler successful, the longest times to loss of the Doppler signal (28 ± 3 seconds males, 34 seconds female). Doppler success was not achieved high on the thigh without proper stretch, but proper stretch was achieved high on the thigh without Doppler success 5 times (2 involved resumption of an audible pulse before 60 seconds). Doppler success was also not achieved at mid thigh without proper stretch, but proper stretch was achieved at mid thigh without Doppler success 3 times (1 involved resumption of an audible pulse before 60 seconds). The relationships between limb occlusion pressures and limb circumferences were that, in general, the larger circumference limbs had higher occlusion pressures. Circumference, however, was not the sole determinant of occlusion pressure; the recipient's systolic blood pressure also played a role.

Additional Training

Two males and 6 females had initial proper stretch rates <70% and received additional training. One male completed his reassessment applications on a different recipient than his first applications (original recipient systolic pressure >140 mmHg). The remainder completed their reassessments on the same recipients as on their first assessments.

The remeasured systolic pressures of the reapplication recipients were not significantly different from their previous values (118 ± 11 versus 119 ± 7 mmHg). The remeasured mid upper arm occlusion pressures were not significantly different from their previous values (114 ± 17 versus 121 ± 21 mmHg). Leg occlusion pressures were not remeasured.

Reassessments occurred at mid right upper arm, mid left thigh, and below right knee. During the 24 reassessment applications, proper stretch and Doppler success rates improved.

Additional training tourniquet applier success. Reassessments by gender then thigh occlusion pressure, first indicates after minimal training (Doppler white, proper stretch gray), second after additional training (Doppler light gray diagonal white, proper stretch dark

gray diagonal black). Improvement: Doppler success $p = 0.007$, proper stretch $p = 0.004$). Mid thigh was the only location with reassessment failures to achieve Doppler success.

Additional training tourniquet location success. Reassessment locations, first indicates after minimal training (Doppler white, proper stretch gray), second after additional training (Doppler light gray diagonal white, proper stretch dark gray diagonal black). One audible pulse return before 60 seconds and three failures to lose the audible pulse). Every properly stretched reassessment application was Doppler successful. Following the additional training, the times to apply were slightly longer but were all <60 seconds (38 ± 11 versus 33 ± 13 seconds; $p = 0.08$). The average times to loss of the Doppler signal when it occurred were the same on reassessment as obtained previously (18 ± 8 versus 18 ± 9 seconds). Recipient discomfort increased (11 None, 9 Little, 4 Moderate, 0 Severe to 2 None, 7 Little, 8 Moderate, 7 Severe; $p = 0.002$), but application was still generally considered easy (17 Easy, 7 Challenging, 0 Difficult). No tourniquet drops occurred during reassessments.

### Strength

Applier seated row and overhead press weights moved are shown in Table I. The weakest applier to achieve a >70% proper stretch rate with no additional training was a female with a 100% proper stretch rate and a combined weight moved of 64 kg (row + press). Only 2 appliers, both females, moved less combined weight than 64 kg. The weakest applier overall had an initial 0% proper stretch rate and a combined weight moved of 34 kg. She achieved a 100% proper stretch rate after additional training.

## DISCUSSION

Properly applied, the SWAT-T stretch and wrap style tourniquet can stop arterial flow through each extremity; however, despite its apparent ease of use, to achieve proper application stretch many appliers will need some technique training. Arterial occlusion is less likely at larger circumference locations such as the thigh without proper application stretch. Fortunately, trainable technique rather than great strength is the key to achieving proper application.

The tested tourniquet has characteristics desired for military use: <230 g, >2 inches wide, easy to apply in <60 seconds, easy to remove and reapply, rugged, <$25/unit, and, most importantly, it can stop extremity arterial flow. The critical location specified by the military is the thigh with an 80% success rate desired. Data from the 10th, 28th, and 31st Combat Support Hospitals in Iraq support the importance of this location, and knowledge of limb circumferences and under-tourniquet pressure distributions supports use of the thigh as the hardest limb location at which to achieve success. Combining the minimal training and postadditional training results, we observed a 77% mid-thigh Doppler success rate when the SWAT-T was properly stretched. Comparing this to lab data with the

CAT, this is lower than the CAT 100% (18 applications) and >95% (20 applications) thigh Doppler success rates reported by Walters et al and Ruterbusch et al but much higher than the CAT 12.5% (6 of 48) thigh Doppler success rate reported by Taylor et al.

Blood pressure and location circumference play major roles in tourniquet success. The 3 mid-thigh Doppler failures despite proper stretch in this study occurred in subjects with elevated systolic pressures (148, 128, and 152 mmHg), larger thigh circumferences (53, 53, and 55 cm), and high pneumatic cuff thigh occlusion pressures (186, 194, and >250 mmHg). Systolic pressures in some of this study's subjects were higher than in reports concerning the CAT and other tourniquets (range 100–130 mmHg, range 103–140 mmHg). Thigh circumferences in this study (40–60 cm) did not reach the top end for U.S. male soldiers (46–79 cm) or for the subjects in the Walters et al study (52–68 cm). (Circumferences were not reported for the Ruterbusch et al or the Taylor et al studies.)

This study had several limitations. First, there was an attempt to have conditions favoring Doppler success: low stress, well-lit indoor environment with no blood, dirt, or sand on the tourniquet and no thick, full length pants or long-sleeved shirts on recipients. This was done to determine if proper application could result in arterial occlusion and if a high rate of proper application could be achieved with very minimal training. Second, there was no visual bleeding feedback, but there was audible Doppler feedback. Although audible Doppler signal loss is a commonly used method for assessing tourniquet-related cessation of blood flow, it is not always an accurate indicator of the absence of blood flow. Third, the rating of proper stretch was subjective and made during the timed application. Fourth, as already mentioned, the range of subject thigh sizes did not include the top end for U.S. soldiers.

On the other hand, the study had several strengths. First, appliers had no previous tourniquet experience and very minimal training before the first applications. These were desirable characteristics because of interest in the militarily desired traits of “simple to apply” and “use with little to no training.” The military desirability of these traits reflects the ideals of emergency tourniquet use: immediate, on-scene application before sufficient blood loss has occurred to result in shock,[1] and the consequent realities of who is applying most of the tourniquets (when recruited, U.S. Army soldiers average 21 to 22 years of age and have an educational background of a high school diploma and sometimes some college credits [Support Army Recruiting http://www.2k.army.mil/faqs.htm#age]; they do not generally have tourniquet training before what they receive in the military). In their study of military tourniquet use, Kragh et al[1] found that “persons with limited training most frequently made the decision to use tourniquets.” Those “persons with limited training” would be the “casualties themselves, lay bystanders,” and “soldiers” who applied an unspecified number of the prehospital tourniquets to 422 of the 499 patients who received tourniquets.

A second strength is that this study clearly delineated between height, gender, and strength versus trainable technique as factors for achieving success. The key difference between the initial assessment and the reassessments was approximately 10 minutes of additional training with feedback.

A third strength is that this study had both male and female appliers with no requirement for applier strength or fitness. Although strong, fit, male appliers are more likely in military settings, the possibility of variable strength, fitness, and gender appliers clearly exists in nonmilitary settings (wilderness first aid, farm implement-related accidents, etc.).

A fourth strength is that this study looked at a variety of locations. The thigh is the most difficult location to occlude blood flow based on circumference, but factors other than circumference can play a role in tourniquet success. One such factor might be the presence of a single long bone with a single major artery (upper arm and thigh) versus two long bones with more major arterial branches (lower arm and lower leg). As in our study, Swan et al found this anatomic difference between upper and lower arm and leg locations not of importance in achieving tourniquet success. In casualty use, Kragh et al also did not observe high tourniquet ineffectiveness on the forearm and lower leg. Another location-related factor is Hunter's canal in the distal thigh with its suggested medial condyle of the femur protection of the superficial femoral artery from compression. Our study showed that Doppler success could be achieved with this tourniquet at the distal thigh location.

Two additional location-related factors come into play because the tested tourniquet needs to be stretched and wrapped: the limb mobility encountered during application and the ability to pass the wrap around the limb without losing tension. The more frequent achievement of proper stretch on the legs than the arms may have been caused by greater recipient arm movements during the tourniquet applications. The maintenance of tension during the wrapping process may have been an issue contributing to inadequate stretch by some of the appliers after only the initial training and was an area addressed during the additional training. In summary, the strengths of this study were the lack of applier experience, the clear delineation of trainable technique as a key to success, the use of appliers of both genders, and the evaluation at a variety of locations.

At all locations, regardless of proper stretch or Doppler success and both before and after the additional training, appliers considered the tested tourniquet easy to apply. Easy application is a claim with most tourniquets that are currently marketed toward military use, and user comments indicating easy to use are common among tourniquets with lab evaluations. Our findings confirm that appliers found the tested stretch and wrap style tourniquet easy to apply, but they also show that an applier's rating of ease of application does not indicate visually correct application or cessation of distal blood flow. A failure of appliers, even with training, to achieve proper application has been noted with both of the two most commonly used military tourniquets (stick and strap CAT and pneumatic Emergency Medical Tourniquet (EMT): incorrect band routing with the CAT and failure to remove slack with both the CAT and EMT). Similar to the need for training to properly apply other styles of emergency tourniquets, our findings also suggest that a significant percentage of adults without prior tourniquet knowledge would not achieve proper application with the tested stretch and wrap style tourniquet without some technique-related training.

Applications with proper stretch tended to be less comfortable than applications with inadequate stretch. However, none of the properly stretched and Doppler successful applications were uncomfortable enough to result in immediate removal. Although we would agree with Swan et al that “pain is irrelevant” with regards to tourniquet discomfort in the face of life threatening hemorrhage, it may not be irrelevant with regards to training with a particular tourniquet. Since this tourniquet is wider than most other field targeted and used designs (CAT, M2, Self-Applied Tourniquet System [SATS] 3.8 cm, Special Operations Forces Tactical Tourniquet [SOFTT] 2.5 cm, London Bridge tourniquet 2.4 cm—all used in Iraq), it might stop limb arterial flows at lower pressures, and it might well be a more comfortable design with which to practice. What pressures are actually exerted when the SWAT-T and other commercial designs are used should be examined.

CONCLUSIONS

The SWAT-T stretch and wrap style tourniquet can easily be properly applied and can stop arterial flow at a variety of extremity locations. Proper application is associated with cessation of arterial flow. Cessation of arterial flow can occur without proper stretch at smaller circumference locations. Proper stretch can occur without cessation of arterial flow at larger circumference locations with higher occlusion pressures. Despite the ease of application, specific training on stretch and wrap techniques is desirable to achieve high rates of proper application stretch. Considering its effectiveness when properly applied, ease of use, small size, light weight, and low cost, this tourniquet might prove useful in its intended arenas of care (tactical, limited supplies, or possibly first responder casualty care) with adequate user training.

BACK TO TOP

***Sleep***

## Maxillomandibular Advancement as Surgical Treatment for Obstructive Sleep Apnea in Active Duty Military Personnel: A Retrospective Cohort

Military Medicine
MAJ Marc M. Serra, DC USA; MAJ David Greenburg, MC USA; CPT Megan Barnwell; COL David Fallah, DC USA; COL Karen Keith, DC USA; LTC Vincent Mysliwiec, MC USA

ABSTRACT

Objective: The objective of our study is to assess the surgical outcomes of active duty military personnel undergoing maxillomandibular advancement (MMA) for the treatment of obstructive sleep apnea. Methods: Pre- and postoperative data on 37 military personnel who underwent MMA were assessed for changes in apnea–hypopnea index (AHI) and minimum oxygen saturation. A surgical success was defined as a reduction of AHI by 50% or a postoperative AHI of <20. Results: 83.7% had an AHI greater than 20 ($n$ = 33; range 7.6–118) with a mean preoperative AHI of 50.5 per hour. The postoperative AHI decreased by 36.3 to a new value of 14.2 ($p < 0.001$). Most service members experienced a postoperative AHI of less than 20 ($n$ = 28; 76%). Sixteen (43%) had a surgical cure (AHI < 5). The number of surgical successes for this study was 81% ($n$ = 30). The mean minimal nocturnal oxyhemoglobin saturation did not significantly change from preoperative 85% (SD = 6.8%) to postoperative 86% (SD = 7%; $p$ = 0.21). Conclusion: MMA represents a viable surgical treatment option for military personnel in whom continuous positive airway pressure is either not tolerated or for those who desire a fully deployable status.

## INTRODUCTION

Obstructive sleep apnea (OSA) is a common disease affecting 2 to 4% of middle aged adults. It is estimated that 3 million males and 1.5 million females in the United States have OSA. The U.S. Military shows a similar prevalence. Analysis of records from the Veteran's Affairs system revealed that approximately 3% of veterans have OSA. The crude incidence rate of diagnoses from 2000 to 2009 showed a nearly six-fold increase in active duty service members diagnosed with OSA. A study performed at Walter Reed Army Medical Center showed that 72.6% ($n$ = 527) of all patients who underwent a formal sleep study in the year 2000 had OSA.

OSA is characterized by the complete or near-complete cessation of airflow during sleep because of the obstruction of the upper airway. Common symptoms of OSA include snoring, choking or gasping during sleep, insomnia, morning headache, and daytime sleepiness. The pathogenesis of OSA is related to transient occlusion of the upper airway during periods of sleep when tonic pharyngeal muscular activity results in collapse of the upper airway. Military personnel with untreated OSA are at risk for accidents related to excessive daytime somnolence (EDS) and, in severe OSA, increased cardiovascular morbidity and mortality. Patients who are suspected of having OSA are referred for a polysomnogram (PSG) to determine their apnea–hypopnea index (AHI). This

represents the number of apneas and hypopneas recorded per hour of the study. The patient's OSA severity is based on a standardized grading system.

Grading of Severity of OSA

The current standard therapy for OSA is continuous positive airway pressure (CPAP); yet, many individuals do not tolerate this form of therapy. Often discouraged by the lifelong nightly requirement for CPAP, approximately 15 to 30% of patients refuse this option. Additionally, 20 to 40% will discontinue CPAP after 3 months because of lack of compliance. Military personnel with OSA are treated with CPAP. However, with CPAP therapy, there are specific requirements according to Chapter 5-14 of AR 40-501, a U.S. Army Soldier can only be deployed when, "… the area of deployment includes the following: availability of a reliable power source, absence of environmental factors that would render electrical equipment inoperable or unreliable, and the availability of power source of replacement supplies." These requirements potentially limit deployability for military personnel diagnosed with OSA. A treatment modality which either cured or minimized the severity of OSA in service members without the inherent requirements of CPAP would be ideal.

Soldiers with an AHI less than 30 per hour and effectively treated with CPAP are deployable to most theaters of operation; however, those with severe OSA are not deployable without an extensive waiver process.[9] In a nondeployed environment, there are not significant concerns about CPAP malfunctions as resupply is readily available. For Soldiers with severe OSA, no clinical criteria establish how long they can safely go without CPAP. A night or two without CPAP has not been shown to result in an adverse outcome, but resupply for CPAP in a deployed environment can take a prolonged period of time. Thus, an AHI cut-off of 60, twice the upper limit of severe disease, is a frequently used reference to send Soldiers to a medical evaluation board or deem them unsafe to deploy to austere environments.

OSA can be caused by an obstruction at one or multiple levels of the pharynx. Various surgeries exist for the treatment of OSA depending on the level of obstruction and severity. The current Stanford protocol addresses the approach to patients with OSA who have failed nonsurgical therapy. This protocol targets the level of obstruction in a "stepwise" approach. Phase I of this protocol targets the retropalatal and retrolingual area. Phase II directs the patient to maxillomandibular advancement (MMA). There are surgeries that specifically address the redundant soft tissues in the parapharyngeal areas that are causing the obstruction. Adenoidectomy, functional septorhinoplasty, turbinoplasty, various modifications of the uvulopalatopharyngoplasty, base of tongue ablation, and hyoid suspensions are surgeries designed to decrease the amount of soft tissue redundancy in the nose, oral pharynx, and hypopharynx. These procedures are helpful in select patients with generally one level of obstruction but are not indicated for patients with severe OSA, which is the type of patient that we are focusing on in this study. Uvulopalatopharyngoplasty has a long-

term surgical response of 40.7% in select patients; thus, the majority of patients will still have residual OSA despite having undergone a surgical intervention.

In contrast to the above surgeries, MMA addresses all levels of obstruction of the pharynx. It has reported success rates of 95, 96, and 98%. Long-term success rates for OSA treated with MMA reported in the literature since 1995 range from 75 to 100%. MMA is reported as a surgical cure for OSA by increasing the volume of the upper airway. The goal of MMA surgery is to move the maxilla and the mandible forward (anteriorly). When the bone and attached soft tissues are advanced, the airway of the nasopharynx, oropharynx, and hypopharynx increase in the anterior–posterior and in the transverse dimensions. A recent study of reconstructed computed tomography cone beam volumetric analysis of the upper airway, pre- and post-MMA, shows dramatic increases in airway volume. Based on the marked changes in the upper airway, where obstruction occurs, and multiple prior studies showing significant improvement in the AHI, MMA has the best chance of curing OSA, resulting in a fully deployable service member.

The aim of our study was to assess surgical outcomes of active duty military personnel undergoing MMA for the treatment of OSA. To our knowledge, MMA and its surgical outcomes have not previously been reported in an active duty population.

## PATIENTS AND METHODS

The study was a retrospective review of all MMA procedures performed at our institution on active duty military personnel between January 1, 2006 and December 31, 2009 for the treatment of OSA. Our study was performed under an approved protocol by the Madigan Healthcare System Institutional Review Board. All patients had their surgery and preoperative and postoperative PSGs performed at Madigan Healthcare System, Tacoma, Washington. Madigan is an Army tertiary referral hospital that covers the Western Regional Medical Command. We identified 37 service members (36 males and 1 female) for inclusion into the study. The mean age was 35 (range 21–50) years old, and the mean AHI was 50.5 per hour shows the distribution of patients among the three classifications of OSA. As our study was not performed in accordance with a prospective protocol, the selection criteria were not standardized.

A general approach to MMA at out institution is to fabricate an acrylic splint achieving a 10-mm advancement, as well as one achieving an 8-mm advancement of the maxilla, should the greater distance be unattainable because of soft tissue restriction. A standard intraoral LeFort I osteotomy is performed with attention to the piriform rims and anterior nasal spine to minimize potential unaesthetic soft tissue changes to nasal tip projection. The acrylic splint is used to set the maxilla, then 2.0-mm titanium plates and

screws are placed bilaterally at the piriform rims and zygomatic buttresses. Allogeneic bone grafting augments and supports any excessive bony gaps after the advancement.

Once the maxilla is set, the mandible is advanced using a standard bilateral sagittal split osteotomy. The distal segment is fixated to the proximal using a minimum of three bicortical 2.0 titanium screws via a transcutaneous technique. Because of the degree of advancement, the final splint usually remains in place with guiding elastics for at least 2 weeks. To address the hypopharyngeal soft tissue obstruction, a sliding horizontal mandibular osteotomy or a genial tubercle advancement osteotomy, of the maximum distance allowed by the thickness of the cortex, is performed and secured with 2.0-mm screws and plates. This is all performed under general anesthesia and then the patient is transferred to the intensive care unit for postoperative observation and pain control.

The outcome variables of postoperative AHI and minimal nocturnal oxyhemoglobin saturation were assessed. Surgical success was defined as a 50% reduction or a postoperative AHI of less than 20 per hour, consistent with previous published comparative studies. PSGs at our institution are scored according to the alternative method of scoring obstructive events set forth by the American Academy of Sleep Medicine.

## RESULTS

The overall success rate for MMA in this study, as defined by our criteria, was 81% ($n$ = 30; range 0–46). 83.7% of service members had a preoperative AHI greater than 20 ($n$ = 33; range 7.6–118). The mean preoperative AHI was 50.5 per hour, and the mean postoperative AHI was 14.2 ($p < 0.001$). The postoperative AHI dropped by 36.3. Most experienced a reduction to at least 20 in their AHI ($n$ = 28; 76%). Sixteen patients (43%) had an AHI of <5 resulting in no residual disease or a “surgical cure.” Two patients had an increase in their postoperative AHI. The mean minimal nocturnal oxyhemoglobin saturation did not significantly change from preoperative 85% (SD = 6.8%) to postoperative 86% (SD = 7%; $p$ = 0.21). Our data did not show a correlation between AHI and minimal nocturnal oxyhemoglobin saturation

## DISCUSSION

OSA is linked to an increased risk of motor vehicle accidents, neurocognitive impairment, and cardiovascular disease. EDS, one of the cardinal manifestations of untreated OSA, results in delayed thought, concentration, and reaction time. When this is combined with a military that demands their personnel to make life or death decisions and maintain a high level of alertness, it is obvious to

understand the need for a comprehensive medical assessment and treatment plan for military personnel with OSA. The increase in military personnel diagnosed with OSA has occurred concomitantly with a rise in medical encounters coded for overweight/obesity. The major risk factor for OSA is obesity. An optimal treatment plan for military personnel would include addressing weight loss in those who are obese as weight loss, as little as 10 pounds, can improve the severity of OSA, and otherwise result in a more fit warrior.

The goal of this study is to suggest a new treatment protocol for military personnel with severe OSA. Current literature and our study show that MMA is a successful treatment for military personnel with severe sleep apnea. This one surgery would not only allow the patient to be useful to the military by being deployable without the need for CPAP, but it would increase the long-term health of its fighting force who are currently suffering from OSA.

OSA has a significant public health impact because of its co-occurrence with other health-related conditions. Hypertension, cardiovascular dysrhythmias, stroke, myocardial infractions, depression, as well as accidents from EDS can result from the prolonged pathophysiology of OSA. MMA would result in a significant reduction in OSA-related health risks and could represent a considerable financial savings on the health care system. The one-time cost of early MMA is potentially less expensive than multiple less predictable operations and hospital stays. In addition, a lifetime of CPAP with its associated costs of repeat sleep studies, registrations, equipment maintenance and replacement, technical support, and compliance counseling may be less cost beneficial than one-time MMA surgery.

The current standard of therapy for OSA is CPAP with compliance defined by wearing CPAP greater than 4 hours nightly on 50% of nights. Although while wearing CPAP, the patient's AHI is usually <5, when noncompliant, the AHI is the same as the preinterventional AHI, and thus untreated OSA. The resultant mean AHI can be much higher than 5 with CPAP therapy and still be considered successful. These implications are magnified in military personnel as their compliance with CPAP can be affected by multiple issues to include nasal congestion, mask issues, airway dryness, lack of reliable electricity, and austere sleeping environments. This suggests that although the MMA postoperative AHI may be higher than 5, it may be a better health benefit since it is consistent every night vs. an AHI of less than five for the 4 hours that the patient wore the device for half of the week.

There is currently a lack of high-level controlled studies in the surgical literature and an absence of standardized criteria to define surgical success for the treatment of OSA with MMA. This has limited the widespread usage of surgery to treat OSA. Yet, in appropriately selected patients, who undergo MMA, their AHI is known and does not have the potential variability which can occur in patients who are not 100% compliant with their CPAP. Nine previous studies with a total of 234 subjects showed an overall reduction in AHI of 87% with MMA. Our data further support the potential of MMA for the definitive treatment of military personnel with

moderate to severe OSA. Our surgical cure rate was slightly lower than the other reported studies. This may be due to a number of factors including a nonselect group of patients and the lack of standardized selection criteria.

MMA is effective treatment of OSA but not without surgical risk. The LeFort procedure can permanently injure the palatine arteries and the infraorbital nerve that gives feeling to the skin of the face and the upper lip. In addition, there is the risk of possible damage to the roots of the teeth in the area of the osteotomy. The greatest risk of a bilateral sagittal split osteotomy for the advancement of the mandible is permanent damage to the inferior alveolar neurovascular bundle, with subsequent loss of feeling to the lower teeth, lip, and chin. There is also the risk of an unfavorable fracture that would require additional plating or maxillomandibular fixation for approximately 6 weeks. Positive and negative esthetic changes may occur with this surgery. A 2000 study revealed patient satisfaction is extremely high after MMA. Furthermore, previous concerns of unfavorable postoperative facial esthetics do not appear to be significant. In contrast to this study, it has been observed that many do not esthetically tolerate such a large advancement of the facial skeleton, and attention to patient selection and preoperative counseling is paramount.

CONCLUSIONS

In our retrospective review of nonselected active duty military personnel undergoing MMA, 81% of patients experienced a significant reduction in their AHI and 43% had a surgical cure with no residual OSA. Larger, multicenter studies utilizing a standardized surgical protocol, with uniform, validated selection criteria and pre- and postoperative assessments are required. MMA has the potential to cure military personnel with moderate to severe OSA and render them fully deployable to all theaters of operation. A protocolized approach to the surgical evaluation, management, and follow-up of military personnel with OSA could result in improved outcomes in future studies.

BACK TO TOP

***Other***

## Medical Costs of War in 2035: Long-Term Care Challenges for Veterans of Iraq and Afghanistan

Military Medicine
James Geiling, MD; Joseph M. Rosen, MD; Ryan D. Edwards, PhD
November 2012

ABSTRACT

War-related medical costs for U.S. veterans of Iraq and Afghanistan may be enormous because of differences between these wars and previous conflicts: (1) Many veterans survive injuries that would have killed them in past wars, and (2) improvised explosive device attacks have caused “polytraumatic” injuries (multiple amputations; brain injury; severe facial trauma or blindness) that require decades of costly rehabilitation. In 2035, today's veterans will be middle-aged, with health issues like those seen in aging Vietnam veterans, complicated by comorbidities of post-traumatic stress disorder, traumatic brain injury, and polytrauma. This article cites emerging knowledge about best practices that have demonstrated cost-effectiveness in mitigating the medical costs of war. We propose that clinicians employ early interventions (trauma care, physical therapy, early post-traumatic stress disorder diagnosis) and preventive health programs (smoking cessation, alcohol-abuse counseling, weight control, stress reduction) to treat primary medical conditions now so that we can avoid treating costly secondary and tertiary complications in 2035. (We should help an amputee reduce his cholesterol and maintain his weight at age 30, rather than treating his heart disease or diabetes at age 50.) Appropriate early interventions for primary illness should preserve veterans' functional status, ensure quality clinical care, and reduce the potentially enormous cost burden of their future health care.

## INTRODUCTION

In 2010, the United States conflict in Afghanistan surpassed Vietnam as the longest American war in history with approximately 2 million U.S. service members deployed to Iraq or Afghanistan. Operation Enduring Freedom (OEF) in Afghanistan has resulted in 1,903 service deaths with 15,516 fighters wounded in action. Operations Iraqi Freedom (OIF) and New Dawn (OND), which ended on December 18, 2011, have resulted in 4,475 deaths and 32,224 service members wounded in action. Through September 2009, the Veterans Health Administration had treated about 510,000 unique veterans from either conflict, and that number is probably closer to 800,000 today. Following the conclusion of the war in Iraq and the reduction of operations in Afghanistan, war-related medical costs for these veterans are likely to grow exponentially over time because of several major differences in these wars compared with previous conflicts. Specifically: (1) Many warriors now survive injuries that would have killed them in past wars. (2) Current improvised explosive device (IED) attacks have caused “polytraumatic” injuries—multiple limb loss; brain injury; severe facial trauma

or blindness—that may require decades of costly rehabilitation. (3) In 2009, the average OEF/OIF service member was 25 to 29 years old[2] and might be expected to live 50 more years. Today's focus of medical, political, and economic attention has been the primary wounds impacting our veterans, such as IED explosive injuries. The strategic theme of this article is that society is not yet considering the medical costs of caring for today's veterans in 2035—a time when they will be middle-aged, with health issues like those now seen in aging Vietnam veterans, exacerbated by comorbidities of post-traumatic stress disorder (PTSD), traumatic brain injury (TBI), and polytrauma. We are concerned about the reality that in 2035, these acute issues will lead to costly long-term medical consequences.

Research is underway to examine the impact of long-term illnesses in military service members following deployment. Unfortunately, these data do not assess the extent of polytrauma, the overlap between injury categories, the persistence of mild TBI and PTSD, or the prevalence of comorbidities or undiagnosed conditions that manifest later. This makes long-term estimates of the costs of treating war wounds highly challenging; but, an emerging consensus among clinicians, government officials, and economists indicates that such costs are likely to be large. In 2008, economists Joseph Stiglitz and Linda Bilmes, in “The Three Trillion Dollar War,” predicted lifetime Veterans Affairs (VA) medical costs for Iraq and Afghanistan veterans to range from $121 to $285 billion in present value. In their 2010 Congressional testimony, they updated their forecast for lifetime VA medical costs (alone) for this cohort to $201 to $348 billion and estimated that the total costs for providing medical care and disability (from VA and the U.S. Social Security Administration) for returning veterans will be $589 billion to $984 billion, depending on the length and intensity of the conflict. In 2011, the Congressional Budget Office estimated that VA medical-care costs for treating OEF/OIF veterans from 2011 to 2020 (excluding disability) could total $40 to $54 billion in inflation-adjusted 2010 dollars.[4-update] A simple extrapolation of the Congressional Budget Office estimates produces a present value of total lifetime medical costs, i.e., discounted over 40 future years, of between $300 and $600 billion if current trends were to continue. Although there is uncertainty in predicting future costs, the estimates continue to rise.

Our goal is to qualitatively list the medical costs of the war on terror and proactively target those costs that we can reduce using medical interventions. The section, Medical Costs of the Iraq and Afghanistan Wars: Sources and Mitigation, and categorize many sources of medical costs of war, along with strategies for prevention and mitigation of such costs. This article cites emerging knowledge about best practices that have demonstrated cost-effectiveness in mitigating the long-term medical costs of war. We propose that clinicians and health care systems employ early interventions (trauma care, physical therapy, and early PTSD diagnosis); preventive health programs (smoking cessation, alcohol-abuse counseling, weight control, stress reduction); and innovations in technology, surgery, and care delivery to treat primary medical conditions today so we can avoid treating costly secondary and tertiary complications in 2035. Although the scope of this article covers the future medical costs for veterans' care, translating these medical activities into social actions and policies, with adequate funding, is key to providing more optimal and economical veterans' health care through 2035 and beyond.

Although some veterans may experience only acute psychological issues, mild PTSD that resolves, or limited physical injuries, many others will experience long-term medical costs—either from the initial illness or trauma, developing chronic medical conditions, or long-term disability. Amputation care, PTSD, and TBI are likely to cause the greatest long-term medical and disability costs. Backlogged VA benefits claims currently number over 800,000, an issue that VA continues to work to resolve. Economist Dr. Linda Bilmes, testifying before the House Veterans Affairs Committee in 2010, said that, "…veterans from recent wars are utilizing VA medical services and applying for disability benefits at much higher rates than previous wars," and that "…the cost of caring for war veterans … peaks in 30 to 40 years or more after a conflict." Current military service members will begin to hit this peak cost around 2035. For an amputation, the secondary and tertiary consequences in middle age might include decreased mobility, weight gain, coronary artery disease, and diabetes mellitus. For PTSD, the secondary and tertiary comorbidities include obesity, depression, substance abuse, smoking, etc. Our goal is to focus limited medical resources on treating the primary issues today, which should help our patients avoid the more costly comorbid health problems in 2035.

The cost of treating primary injuries is like the "tip of the iceberg" compared to the unknown and large costs of treating secondary and tertiary consequences "below." Our suggestion is to prioritize initiatives "above the line" that would help maintain amputees' primary mobility, functionality, and physical activity, and support rehabilitation for PTSD or TBI so that the wounded veteran could work productively and engage in society. We should help an amputee to reduce his cholesterol and maintain his weight at age 30, rather than treating his coronary artery disease or diabetes at age 50. If we treat a veteran's PTSD at age 21, with counseling and lifestyle interventions, we may help her to reduce suicidal thoughts and avoid abusing tobacco or alcohol. This will save us from having to fund her treatment for chronic obstructive pulmonary disease or alcoholic liver disease in 2035.

Trauma and Polytrauma: Prevention and Acute Care

Trauma and polytrauma, especially from IED injuries, constitute a huge component of medical costs for Afghanistan and Iraq war veterans. Our future goal would be to prevent the trauma or reduce the severity of injuries sustained—through better protective equipment, body armor, armored vehicles, etc., and more methods to detect and disarm IEDs. The acute costs of trauma care are higher for current veterans as a cohort since more wounded service members now survive with profound injuries than in any previous U.S. war.[5] Access to care poses challenges when providers themselves are deployed. During recent conflicts, the Army lacked enough medical providers for the active force; this was increasingly true for women soldiers. Medical costs will accrue at various points in the military health care system, from battlefield evacuation through surgery, transportation, and rehabilitation, and will

include job retraining and assistive devices, such as prostheses and wheelchairs. One way to streamline costs is to make the system of evacuation and treatment optimally efficient through better technologies and improved surgical techniques.

### Amputations

War-related military amputations now occur at double the rate seen in previous wars. A 2010 study found that twice as many wounded U.S. soldiers had limb amputations as in either 2009 or 2008, and three times as many had lost more than one limb. The report cited increased foot patrols in Afghanistan in 2010, during which soldiers could step on buried mines, as a likely cause. As of April 2012, there had been 1448 OIF/OEF/OND amputee patients treated in all Military facilities; 436 (30%) with multiple amputations (Dr. Michael J. Carino, Department of the Army Office of the Surgeon General, Public Affairs, personal communication).

Amputations pose the greatest challenges of war, financially, physically, and socially. In the acute setting, amputees require trauma care, with complications of infections, anemia, and heterotopic ossification. In one study of Iraq and Afghanistan amputees, 80% needed physical therapy, occupational therapy, prosthetics, and psychiatry. OIF/OEF amputees reported lower quality of life if their amputation were accompanied by either a combat-related head injury, a greater injury to the nonamputated limb, or a need for assistance with daily activities. Amputees' anatomy will change with age, and obesity is a serious concern. Amputees may experience cardiovascular disease, osteoarthritis, back pain, and phantom limb pain. Their prostheses and wheelchairs may require replacement or upgrades based on newer technologies (such as neurally controlled robotic arms). Given Moore's Law, which predicts ongoing exponential growth in digital devices' capabilities, these veterans' devices may have ongoing high costs forever.

However, improved technologies and treatments may somewhat mitigate the cost of amputations. Newer prostheses that allow amputees to run and remain active may offset weight gain. And although military amputation rates have risen, improvements in care delivery have allowed more amputees to return to duty: 16.5% of amputees in 2001–2006 returned to active duty, whereas the comparable rate in the 1980s was 2.3%. Research on transplant and tissue-engineering technologies to repair veterans' injuries and restore functioning is ongoing at the Armed Forces Institute for Regenerative Medicine—with great promise for a better future for our brave service members.

### Mental Illness, PTSD, and Comorbidities

Symptoms of PTSD have been observed in veterans of every recent war, with the name varying from “shell shock” to battle fatigue, combat neurosis, post-Vietnam syndrome, and PTSD. PTSD can be a key factor by which trauma translates to poor long-term health. Many veterans develop PTSD whether or not they have been physically injured. These veterans have a lower quality of life and more medical problems than those without PTSD. Suicide rates are high in current conflicts, and returning service members face

financial challenges and high divorce rates. PTSD is associated with smoking; substance abuse; depression and anxiety; heart disease; obesity; diabetes; gastrointestinal, dermatologic, and musculoskeletal disorders; chronic fatigue; and increased dementia.

Treatment for both PTSD and its comorbidities is costly. A comprehensive study of PTSD and major depression in OEF/OIF service members estimated the prevalence of each at 14%, with a 4% overlap. In a 2010 study of U.S. service members in Iraq, the prevalence rates of PTSD and depression after combat ranged from 9 to 31%, depending on the level of functional impairment. That study also estimated that the total societal cost per person—including lost earnings and overlap suicide costs—was between $5,900 and $25,800 over 2 years, of which only about 3% represented treatment costs paid by Department of Defense (DoD), VA, or private payers. A follow-up study, with improved data on personnel and treatments, estimated the 2-year costs at about $16,000 per case. If rates of price inflation and prevalence were to remain unchanged, the per-person societal cost could reach nearly $50,000 over 2 years by 2035. Using this figure as a rough average, it might cost $1,250,000 to treat depression and PTSD in one current veteran for 50 years—not counting other health problems.

Society's best hope of containing costs may be to screen and treat PTSD early, along with related physical-health consequences. For example, PTSD and alcohol/drug abuse require concomitant treatment since psychological disease can reduce substance-abuse treatments' effectiveness. Smoking cessation combined with PTSD care is more effective than smoking-cessation treatment alone. Depending on which condition (PTSD or a physical ailment) is first diagnosed, screening and treatment can be bidirectional. When male VA diabetes patients with PTSD and depression were found to be vulnerable to weight/lipid problems, one study recommended screening of all diabetes patients for mental health comorbidities.

Some issues complicate PTSD diagnosis and treatment. Clinicians often do not agree on how to diagnose PTSD or how to distinguish it from mild TBI, depression, or other mental health conditions. Misdiagnosis also occurs in cases in which veterans may suffer from stress or anxiety that is unrelated to combat trauma. When diagnosed, it is difficult to classify PTSD (or TBI) as minor vs. major. Finally, it is critical to have enough providers to serve veterans; in 2007, there were too few mental health providers in the DoD/VA health care systems, and although measures have been taken to increase staffing and reduce waiting times, more help is needed to achieve optimum treatment. However, despite such limitations, patients will benefit from a preventive approach that employs early intervention and treatment for PTSD and comorbid illnesses, and possible decompression following deployment; this should improve outcomes and avoid a “snowballing effect” on long-term medical-care costs.

Traumatic Brain Injury

TBI from bomb blasts, which cannot be prevented by body armor or rapid medical attention, is being called a “signature wound” of the Iraq war. TBI accounts for roughly 22% of casualties in Afghanistan and Iraq, and was found in 59% of patients exposed to blasts

in one study. Mild TBI can be missed on imaging, and veterans may not show brain damage until years after a blast injury. Head injury in young adulthood may correlate with greater risk of Alzheimer's disease in later life. Mild TBI, like PTSD, correlates statistically with increased rates of psychological, physical, and functional problems, and is associated with alcohol-abuse disorders. Among 2525 U.S. Army soldiers deployed to Iraq, those who had experienced mild TBI, based on self-reported prior loss of consciousness (a noted study limitation), were more likely to report poor health, missed workdays, medical visits, and somatic and postconcussive symptoms, compared to soldiers with other injuries. In a cross-sectional cohort study of 278 TBI patients and 3218 normal controls, mild TBI, even years after the injury, correlated with increased headaches, sleep problems, and memory difficulties; it could also prolong recovery from comorbid conditions, including PTSD.

An important but difficult objective is to separate TBI from PTSD, to diagnose when the two conditions coexist, and avoid attributing health problems to mild TBI if associated PTSD and depression may be the primary problem. Patients diagnosed with both disorders are likely to require collaborative, coordinated support across a broad set of providers who possess combined expertise in mental health, neurology, and internal medicine.

As the veterans of Iraq and Afghanistan age, it will be important to refine screening instruments, to evaluate and treat patients proactively for PTSD and TBI, and to employ resiliency programs. One initiative is the U.S. Army's protocol to “educate, train, treat and track soldiers”—intervening, assessing, evaluating, and tracking progress of a service member after a concussion. The U.S. Marine Corps Combat Operational Stress Control program proactively identifies five core ways in which to manage stress: “strengthen Marines; mitigate and remove unnecessary stressors; identify Marines with stress problems; treat and coordinate care, and reintegrate back to unit.” The U.S. Navy has 10 resilience initiatives to help both veterans and families who care for them. A Veterans Health Administration program trains clergy to help the estimated 30 to 40% of veterans from rural areas connect with mental health resources; the Restore Warriors website provides information and tools on how to manage PTSD and TBI and “aims to educate veterans, service members, families and friends on the potentially life-altering after-effects of war, from post-traumatic stress and depression to relationship problems or general anxiety.”

Additional Health Impacts and Costs for Service Members' Families

Repeated deployments in Afghanistan and Iraq, with extended separations, have caused unusual stress for service members' families—before, during, and after deployments. First lady Michele Obama said that it was “important to recognize that the children of people in the military are also making sacrifices and must often move from school-to-school, or deal with prolonged absences of one of their parents.”Anxiety over impending deployment can cause unhealthy behaviors, including tobacco and alcohol abuse, overeating, children's behavioral problems, increased domestic abuse, reckless driving with potential trauma, or high-risk sexual activity. Military spouses may have increased depression or worries about family functioning during deployment. Many servicemen's

wives “are obliged to give up their own careers when they accompany personnel to new postings, especially overseas” (or to care for them when wounded). The result is that “military families lose not only the wife's salary but also the future pension entitlement that her on-going employment would have earned.” Stress impacts one's mental and physical health, in the short and long terms; some relatives may suffer negative psychological consequences long after their loved one's acute issue is resolved. Long-term costs for the family and society may include at-home nursing care, relocation expenses, and lost job productivity for family members who leave work to care for injured veterans; increased administrative overhead for disabled veterans' claims; and costs associated with stress induced by delayed response to medical claims. Systems of care may or may not be burdened by these “hidden” costs, but regardless, they represent harms to an individual's physical and emotional well-being.

Hidden costs are hard to quantify or mitigate. However, military resiliency and support programs exist to help families cope with personal and financial problems, including the Survivor Outreach Services program, serving 133,000; the Military Family Life Consultant program; the Exceptional Family Member Program; the alcohol and substance abuse program; mobilization readiness programs for both service members and their families; and counseling to treat service members' families for deployment-related stress to hopefully prevent it from escalating into trauma, family abuse, or suicide. Hospital programs in Boston, Massachusetts, Grand Rapids, Michigan, and elsewhere also serve veterans and their families whose health care needs cannot be met within the VA system.

In 2011, VA announced a plan to compensate families caring for wounded service members at home, which could reduce nursing-home costs. (The U.S. Department of Labor's Family and Medical Leave Act allows for military caregiver leave, but only on an unpaid, temporary basis.) In 2012, the Obama administration proposed that service members' spouses, children, or parents be allowed to take “up to 12 weeks of leave from work to help a service member deployed on short notice” to “arrange child care…financial and legal arrangements without fear of losing their jobs.” Also, they would get “up to 26 weeks of leave to care for recent veterans who were injured or became ill in the line of duty.” The Military Spouse Employment Partnership is a group of companies that pledge to “recruit, hire and promote military spouses and help them maintain portable careers.” In Virginia, a retreat for recovering service members and their families is planned, where they can vacation, partake in therapy and recreational activities, rest and heal.

Employing early interventions for sick or injured veterans might also reduce some administrative costs of managing their long-term care needs. Additionally, many strategies proposed to mitigate direct medical costs may reduce or offset indirect costs; however, some intangible, nonreimbursable costs will continue to burden military families.

Beyond the scope of this article, other factors can or potentially will influence the future medical costs of today's veterans, but are harder to predict, and not all have defined mitigating actions. These may include, but are not limited to, the length of the conflict; a

“young” military, with higher lifetime medical costs; the increasing American lifespan; changing political climate and policies; TRICARE and other third-party payer policies; and evolving medical technologies.

## DISCUSSION

This article presents some unique contributions and themes that have not previously been covered in other published articles on the costs of war. Although long-term costs have been discussed in the policy realm, the speakers and authors have mainly been economists and policy experts—not medical experts, planners, or clinical providers in the medical domain who can actually put preventive strategies into place. The most recent articles did not typically discuss medical interventions or nuances, and we believe this article offers a singular perspective with suggested preventive medical strategies to reduce the costs of war.

Much of the current discussion on costs of war focuses on the primary injury costs, such as the costs of a leg amputation, prosthesis, or IED injury. Other articles have not focused, as ours does, on the secondary and tertiary consequences of the original illness or disability. Many authors have not fully considered or examined the practical costs of such second and third degree comorbidities that will ultimately result from today's acute issues.

One limitation of this work is that we list costs qualitatively rather than quantitatively. As previously mentioned, long-term estimates of the costs of war are highly challenging; hence, we chose to classify medical costs qualitatively and target cost categories for which clinical interventions exist that have been shown cost-effective. Also, since there are inherent challenges in implementing any preventive strategies to mitigate care costs—especially in an environment of fiscal budgetary constraint, decreased defense spending, and prioritizing limited medical resources—any new health technologies and interventions should be evaluated for their economic outcomes.

Several proposed interventions have been validated and tested to reduce actual care costs. Injury-prevention technologies like improved body armor did reduce deaths in OIF and OEF; if this were not so, we would have a lower ratio of wounded:dead, more like that of Vietnam. Weight control has been shown to reduce or eliminate amputation comorbidities like diabetes, high blood pressure, heart disease, and morbid obesity. Lehnert et al reviewed 41 obesity-prevention interventions for long-term (40+ years) economic effectiveness and found that the greatest cost savings came from interventions that promoted healthy eating and/or physical activity by modifying a target population's environment through regulatory or fiscal measures (e.g., taxes or subsides on certain foods, distribution of educational information, mandatory nutrition labeling, etc.). Saha et al found that lifestyle interventions for preventing diabetes and cardiovascular disease were cost-effective (notably, diabetes screening and prevention programs, childhood obesity

prevention, and community-based programs). Treatments for stopping smoking and drug abuse (which commonly accompany PTSD) have been demonstrated to reduce long-term comorbidities like lung cancer, liver failure, etc. Kahende et al reviewed economic evaluations of tobacco control programs and found that “in almost every case… tobacco control programs and policies are either cost-saving or highly cost-effective.” Popova et al reviewed cost–benefit analyses of alcohol dependence (AD) treatments and also found that “Most of the available treatment options for AD appear to produce marked economic benefits.” Treating PTSD patients and amputees with lifestyle interventions should thus reduce the incidence of secondary and tertiary complications, and the associated long-term costs.

## CONCLUSION

Improvements in technology and health care delivery have significantly increased veterans' survival from grievous war wounds. Although media coverage of the Iraq and Afghanistan conflicts has focused on IED injuries or PTSD, little attention has been given to how such war injuries will impact the long-term health of these young veterans. Some will face life with PTSD or TBI and others will adjust to a lifetime without limbs, eyesight, or even a normal face. Many will develop comorbid conditions that exacerbate an original illness or injury. Veterans with polytraumatic injuries will need long-term rehabilitation, job retraining, and support for daily life activities.

For polytraumatic injuries, improving assistance to young amputees during the critical period in which they are developing coping skills may foster independence and reduce both patient need and societal cost in the longer term. PTSD and TBI patients will need support to return to active service, if possible, and to avoid comorbidities. We recommend addressing veterans' mental-health and neurological needs via a new paradigm for health care delivery—one that includes outreach, education, early detection, lifestyle interventions, and more accurate postdeployment diagnosis and treatment by skilled providers in primary-care clinics—to improve mental and physical health outcomes and contain costs.

Despite resource constraints, VA and other groups are spearheading improvements that support veterans' care. The 2010 IOM report noted that VA has opened more vet centers. Linda Bilmes' 2010 testimony noted that VA has “expanded the Benefits Delivery at Discharge (BDD) program and Quick Start, … liberalized the PTSD stressor definition, increased … benefits and outreach, provided five years of free healthcare … and is … restoring medical care to 500,000 moderate-income “Category 8” veterans. VA has also hired more medical and claims personnel, [and] invested heavily in IT upgrades to the claims processing system.” She then noted that, “All of these factors contribute to the rising cost estimates.”

Our suggestion for mitigating costs is to work “above the line” and prioritize initiatives that would help maintain primary mobility, functionality and physical activity for amputees, and support rehabilitation for PTSD or TBI so that the wounded veteran could work productively and engage in society. We should help an amputee to reduce his cholesterol and maintain his weight at age 30 to 40, rather than treating his coronary artery disease or diabetes at age 50. If we treat a veteran's PTSD at age 21, with counseling and lifestyle interventions, we may help her to reduce suicidal thoughts and avoid the abuse of tobacco or alcohol. This will save us from having to fund her treatment for chronic obstructive pulmonary disease or alcoholic liver disease in 2035.

Looking forward, we should use outcomes research and real-time epidemiology to track our veterans' progress through the health care system, and guide our therapies and preventative strategies. For example, a 30-year-old soldier who had a profound injury in 2001 is now over 40—we should ask, are lifestyle interventions working? If so, which ones? The correct answers for today may be different than the answers in 2035—when improved outcomes, techniques, and technologies will guide our planning and funding of veterans' health care.

This article's goal has been to highlight the medical costs of war—specifically, the long-term secondary and tertiary costs of care that may result from the primary injuries of our current young veterans—and to suggest prevention and treatment measures that may mitigate such costs for these wounded warriors, their families, and support systems such as VA and TRICARE. We hope that an open discussion might ensue in the military and medical communities regarding how to strategically care for veterans of the current wars in a sustainable manner through 2035 and beyond. Support for veterans' long-term care will require the same diligence, perseverance, research, innovative thinking, and funding that we currently apply to the acute medical issues facing our servicemen and women. Strategizing preventive measures now may improve clinical and social outcomes for our wounded warriors and mitigate the potentially enormous cost burden of their future health care.

BACK TO TOP

## Is Dengue and Malaria Co-infection More Severe Than Single Infections?

Malaria Today
Loïc Epelboin, Matthieu Hanf, Philippe Dussart, Sihem Ouar-Epelboin, Félix Djossou, Mathieu Nacher, Bernard Carme
7 Nov 2012

Abstract

Background
Dengue and malaria are two major arthropod-borne infections in tropical areas, but dual infections were only described for the first time in 2005. Reports of these concomitant infections are scarce and there is no evidence of more severe clinical and biological pictures than single infections.

Methods
To compare co-infections to dengue alone and malaria alone, a retrospective matched-pair study was conducted between 2004 and 2010 among patients admitted in the emergency department of Cayenne hospital, French Guiana.

Results
104 dengue and malaria co-infection cases were identified during the study period and 208 individuals were matched in two comparison groups: dengue alone and malaria alone. In bivariate analysis, co-infection clinical picture was more severe than separated infections, in particular using the severe malaria WHO criteria. In multivariate analysis, independent factors associated with co-infection versus dengue were: masculine gender, CRP level > 50 mg/L, thrombocytopaenia < 50 10 9/L, and low haematocrit <36% and independent factors significantly associated with co-infections versus malaria were red cells transfusion, low haematocrit < 36%, thrombocytopaenia < 50 10 9/L and low Plasmodium parasitic load < 0.001%.

Conclusions
In the present study, dengue and malaria co-infection clinical picture seems to be more severe than single infections in French Guiana, with a greater risk of deep thrombocytopaenia and anaemia.

## Background

Dengue fever and malaria are the most common arthropod-borne diseases in humans and represent major public health problems. Dengue virus (family Flaviridae, genus Flavivirus) and Plasmodium parasites are widespread in American and Asian tropical regions and their endemic areas overlap extensively. Nevertheless, reports of malaria and dengue dual infection are scarce. Since the first case reported in 2005, [1] only case-reports and two descriptive studies have been published. They have been reported with Plasmodium falciparum and/or Plasmodium vivax in India and Pakistan, [2–5] Southeast Asia, [6,7] French Guiana [8] and Brazil. [9] This phenomenon seems to be uncommon. In a study performed in Thailand among 194 patients with dengue, no co-infection with malaria was found, [10] but in French Guiana, a retrospective study performed in 2004–2005 on 1,723 consecutive febrile emergency patients found 17 co-infections, including six acute concurrent infections (e.g. 1% of dengue and 4% of malaria cases). [8] The influence of co-infections on severity is not straightforward, therefore, the aim of this study was to differentiate clinical and biological picture of co-infections from infections alone and determine whether patients infected by both malaria and dengue (MD) were more severe than either infection alone (respectively M and D).

## Methods

Study Location

French Guiana is a French Overseas territory located on the north-eastern coast of South America. About 90% of its surface of 84,000 km 2 is Amazonian rain forest; the remaining 10% in the north is a coastal plain where 90% of the 215,000 inhabitants live and Cayenne and surroundings contain almost 50% of the population in 2009. [11] Malaria and dengue fever (DF) represent two major public health concerns in French Guiana. Malaria is endemic and the annual number of cases ranges from 3,200 to 4,700. [12] Until 2006, P. vivax represented 50% of annual cases. The current proportion of P. vivax malaria is 75%, as in the rest of the Americas. [12–14] Since the first cases of DF were reported in French Guiana in 1943, an increase in the number of DF cases and DF outbreaks and the emergence of dengue hemorrhagic fever (DHF) have been observed. [15] All four dengue virus serotypes circulate in French Guiana. The last two mains epidemics occurred in 2006 and 2009, and dengue is currently endemic. Until 2005, dengue outbreaks were exclusively described on the coast. Since 2006, outbreaks of DF have been reported in interior villages where malaria is endemic. [16]

Study Population

A matched retrospective study was conducted comparing patients infected with concurrent malaria and dengue to patients with either infection alone. The study population included all patients admitted in the emergency department of Cayenne hospital, between June 2004 and February 2010. The diagnosis of dengue and malaria co-infection was made on the basis of concomitant biological diagnosis of dengue and malaria within seven days in patients with a compatible clinical picture. Two control groups were constituted: the group M with positive biological diagnosis for malaria and negative for dengue, according to the criteria defined in the next paragraph, and the contrary for the group D. Control cases were matched on the date of biological diagnosis of infection.

Case definitions were based on compatible clinical history and biological diagnosis. Malaria diagnosis relied on the identification of haematozoa on a thin blood film and/or on a thick blood film stained with Giemsa (group MD and M). The screening sensitivity was ≈ 6 plasmodia/μL (1/1,000 leukocytes). The asexual parasite load (PL) was classified in five classes: class 5: >1.25%; class 4: 0.125 to 1.25%; class 3 : 0.0125% to 0.125%; class 2: 0.00125 to 0.0125%; and class 1: ≤0.00125. Malaria rapid diagnosis tests were not systematically performed on the study period. Due to the evolution of the techniques between 2004 and 2010, the laboratory diagnosis of dengue relied on different methods. Direct diagnosis was based on virus isolation, genome detection by Reverse Transcriptase-Polymerase Chain Reaction (RT-PCR) or NS1 antigen detection introduced in 2006 in French Guiana. Indirect diagnosis was based on detection of specific anti-dengue IgM and/or IgA antibodies in patients' sera. [17] When NS1 antigen detection was available, RT-PCR, which allows serotype identification, was not systematically performed.

Concerning dengue definition (groups MD and D), cases were separated in two groups: "confirmed acute dengue cases" (CADC) were defined by direct biological diagnosis (NS1 antigen and/or RT-PCR and/or virus isolation), IgM seroconversion (early serum sample negative for IgM but convalescent sample positive) or IgA antibodies detection. "Likely dengue cases" (LDC) were defined by

IgM antibodies detection. Indeed, IgM appear between the 3 rd and 5 th day of fever but can persist for over three months and IgA appear concomitantly with IgM but does not persist longer than five to six weeks. [18,19] There was no discrimination between dengue primary infection and secondary infection, e.g., further infection(s) by dengue of a different serotype.

Covariates Included, Data Collection and Statistical Analysis

Patients' data, including socio-epidemiologic data, previous medical history, clinical symptoms, and biological results, were obtained from the computerized medical charts. Data were analysed using R version 2.10.0® and the Epicalc® package.

The continuous variables of interest were categorized following the laboratory cut-off values, or published values. They generally were dichotomized because of the small sample size.

Two analyses were performed comparing separately MD to D and MD to M. For categorical variables, a matched bivariate analysis using the Wald test was performed to identify factors associated with co-infections. Statistical significance was set at $p < 0.05$. Variables with a p-value < 0.2 in bivariate analyses were entered into multivariate model to identify the factors independently associated with dengue-malaria co-infections. As bivariate analysis were made in an exploratory way and used as a selection criterion for inclusion in the final multivariate model ($p < 0.20$), we did not judge as a necessity to adjust bivariate p-values for multiple comparisons. Thus, bivariate p-values near to 0.05 must be relativized. To obtain more powerful models, and because of the missing data inherent to retrospective studies, variables obtained from anamnesis and clinical examination and variables with more than 5% of missing data were excluded from the model. Thus, conditional multivariate backward stepwise logistic regression estimated the adjusted odds ratio (OR) and the confidence intervals linked to co-infections.

Ethical Considerations

The retrospective use of anonymous patient files on the site of patient care is authorized by the French National Commission on Informatics and Liberties. All the data collected retrospectively were anonymized in a standardized case report form and in the database.

Results

Cases Description

Between June 2004 and February 2010, 104 patients satisfied the criteria for MD (Figure 1). Consequently, 208 patients were matched in each comparison group. Among the 104 MD patients, 75 (72.1%) were men and 11 (10.6%) were 15 years-old or under. The mean age was 33.8 years (range: 6 months to 83 years). Forty-one (39.4%) were considered as CADC and 63 (60.6%) as LDC versus 150 (72.1%) and 58 (27.9%; $p < 0.001$) in the group D, respectively ( Table 1 ). The dengue virus serotype could be identified

for only 10 (9.6%) of co-infected patients (DENV-1: 3 (30%); DENV-2: 2 (20%) and DENV-3: 5 (50%)) and 91 in the control group (DENV-1: 25 (27.5%); DENV-2: 25 (27.5%); DENV-3: 28 (30.8%) and DENV-4: 13 (14.2%)). No significant difference was found in the proportion of P. vivax between the MD group and M group: P. vivax (76.7% vs. 68.1% respectively), P. falciparum (20.4% vs. 28%) and association of P. vivax and P. falciparum (2.9% vs. 3.9%). Species identification was not possible in two patients because they had received anti-malarial treatment after positive rapid diagnostic testing in a health centre. A low PL (class 1 and 2) tended to be more frequent in the MD group than M group (p 0.08) ( Table 2 ).

Comparison of Co-infection With Dengue

Clinical and biological pictures of co-infection cases were different from single infections and bivariate comparisons showed more differences between MD and D than between MD and M ( Table 3 ). MD patients were more often adult men. A quarter of them reported having recently visited the forest (military, forest workers or gold miners) versus 3% in the D group. Patients from MD group resided more frequently far from the coast than D patients and had a history of malaria and recent malaria attacks (<3 months). The duration of fever was longer in MD patients and they were hospitalized more frequently than D patients, but hospitalization was not longer. More patients required a transfusion in the MD group. The clinical presentation in MD patients was generally more severe, with more fever above 40°C, tachycardia, initial hypotension, nausea, vomiting and dehydration than D patients. Furthermore, cases which fulfilled one or more of the WHO clinical and/or biological criteria for severe falciparum malaria [20] were more frequent in the MD group than in the D group ($p < 0.001$). Increased C - reactive protein (CRP), especially > 50 mg/L, was significantly associated with MD co-infections ( $p <0.001$) relative to dengue alone. Retro-orbital pain, skin rash and ENT symptoms were significantly associated with D. Anaemia (p 0.02), severe thrombocytopaenia ($p <0.001$), and elevated bilirubin ($p <0.001$) were more frequent in MD patients. CADC diagnoses were significantly more frequent in the D group than in the MD group. Anaemia, severe thrombocytopaenia, male gender, high CRP level and LDC diagnosis were significantly associated with co-infection in multivariate analysis ( Table 4 ).

Comparison of Co-infection With Malaria

No significant difference between MD and M was found in terms of gender, age, place of residence and forest-related activities ( Table 2 ). A history of malaria was more frequent in the M group. The fever duration was longer in MD patients but not hospitalization. They received significantly more transfusions (p 0.02). Low blood pressure, signs of shock, pallor were significantly associated with the MD group. Anaemia and severe thrombocytopaenia were also significantly more frequent in the MD group. Cases which fulfilled one or more of the WHO criteria for severe falciparum malaria [20] were more frequent in the MD group than in the M group (p 0.007). Anaemia, severe thrombocytopaenia, low parasitaemia, and a high number of blood transfusions were independently associated with co-infections in multivariate analysis ( Table 4 ).

Discussion

The unexceptional nature of the association of dengue and malaria is confirmed in French Guiana. In regions where these infections are transmitted in close proximity, the classical concept that malaria occurs in rural areas and dengue in urban areas may thus also be contradicted by facts in many countries and simultaneous infections may result from the overlap of the mosquito biotopes. [16]

This study presented some minor biases. There were a higher number of cases based on the IgM detection in the MD group than in the D group. Co-infected patients with LDC were tested separately, so the association between MD and thrombocytopaenia, anaemia and frequent transfusions persisted in bivariate analysis, but not in multivariate analysis. However, when testing separately all patients with CADC diagnosis, no significant difference was found in bivariate and multivariate analysis which is probably due to the loss of power. The separation in two groups, LDC and CADC is arbitrary since the decision to perform the direct diagnosis relied on non verifiable information provided by patients on fever duration, a relatively unreliable answer given the frequent linguistic difficulties in FG. Studying confirmed acute cases alone was questionable, and the authors decided to study together likely and confirmed acute cases because it allowed a larger sample. Furthermore, associating likely cases would have minimized differences between the group MD and malaria alone because "real" associations would have been mixed with isolated malaria cases. It appears that co-infected patients consulted significantly later than the other groups, which may explains the predominance of IgM-diagnosed cases. However, almost all previous studies on dengue and malaria co-infections relied on IgM diagnosis. [1,2,4–8,21] Another hypothesis to explain the relatively high number of LDC in the study group is that malaria attack could have been triggered by dengue infection, especially as there is a majority of P. vivax infection, possibly relapses, which are coherent with the high frequency of malaria medical history, especially in the last three months in the MD. [22] As low parasitaemia was significantly more frequent in MD than in M group, another explanation could be the discovery of asymptomatic P. vivax infections in patients living in endemic areas, but this phenomenon has been barely described in the Amazonian region in Amerindian population, which is not the case here. [23] To ensure that low parasitaemia in the co-infection group did not result from asymptomatic infections, we tested MD patients with low parasitaemia versus D separately. The association between MD and thrombocytopaenia, anaemia, masculine gender, elevated CRP level, frequent transfusions persisted in bivariate analysis but not in multivariate analysis probably due to the loss of power. This result suggests that they were true MD co-infections with low parasite burdens.

The present study suggests an increased severity of the simultaneous infection compared to the isolated infections, which has only been hypothesized previously, [1,4] in particular with haematological consequences. However, severe malaria cases as defined by the WHO were more frequent in the MD group than in M and D separately in bivariate analysis. [20] Indeed because of insufficient power, P. vivax and P. falciparum malaria were pooled, which limits the study of P. falciparum severe malaria. Therefore, the biological influence of dengue virus, which affects the endothelium, a major protagonist of severe malaria pathophysiology, on the eventual severity of falciparum malaria, needs to be studied. [24]

Co-infected patients presented deep thrombocytopenia more frequently than patients with single infections. Low platelets are common in dengue and malaria. In febrile patients living or returning from endemic areas, it is a good predictive factor of malaria [25,26] and in case of negative malaria diagnosis it is a good predictive factor of dengue. [25] During malaria attack in adults,

thrombocytopaenia is generally not considered to be a risk factor of haemorrhage and increased mortality. [27] Nevertheless, in non-immunized children with a malaria attack, a platelet count below 100 10 9/L has been demonstrated to be a predictive factor of mortality. [28] Furthermore, a study performed in France on 21,888 cases of imported P. falciparum malaria showed that thrombocytopaenia below 50 10 9/L was associated with an increased risk of mortality. [29] Considering dengue fever, high thrombocytopaenia is a known severity criterion and is linked to a higher mortality. [30] In the present study, severe thrombocytopaenia was not really accompanied with a recrudescence of haemorrhagic signs. No significant difference for the thrombocytopaenia between P. vivax and P. falciparum was observed. During malaria attacks, thrombocytopaenia generally worsens linearly with the increase of PL. This relationship did not clearly appear in patients co-infected with dengue so it is notable that deep thrombocytopaenia apparently occurred even with low parasite loads when associated with dengue virus.

Anaemia was more frequent in patients with dual infection. There was a convergence of indirect signs, such as pallor and transfusion need and elevated total bilirubin, probably in relation to increased haemolysis. Anaemia is a classical symptom of malaria but it is barely described in dengue fever. Indeed, elevated haematocrit is found in the severe dengue fever cases, resulting in plasma leakage syndrome. [30] The study performed on 21,888 cases of imported P. falciparum, showed that haemoglobin < 8 g/dL was an independent predictive factor of mortality. [29]

Conclusions

In the present study, concurrent dengue and malaria infection tends to be more severe than single infections notably for haematologic abnormalities, such as thrombocytopaenia and anaemia, known risk factors of severe dengue fever and/or malaria. However, whether this increased severity results from longer evolution duration or increased virulence or both remains to be determined. The study was retrospective so the results should be interpreted with caution. Whether prospective studies with homogeneous biological diagnosis methods and patient groups would be necessary to confirm the greatest severity of co-infection, the feasibility of such a study is questionable because of the very low prevalence of dual infection. The current evolution of these two mosquito-borne infections suggests that co-infections could become a medical problem. Since the biological and clinical characteristics of dengue and malaria are very similar, all clinicians treating patients in or returning from endemic areas should systematically order examinations for both diagnoses, even if one or the other is positive.

BACK TO TOP

## Electronic Health Records and National Patient-Safety Goals

New England Journal of Medicine
Dean F. Sittig, Ph.D., and Hardeep Singh, M.D., M.P.H.
8 Nov 2012

Electronic health records (EHRs) are essential to improving patient safety.1 Hospitals and health care providers are implementing EHRs rapidly in response to the American Recovery and Reinvestment Act of 2009.2-4 The number of certified EHR vendors in the United States has increased from 605,6 to more than 10007 since mid-2008. Recent evidence has highlighted substantial and often unexpected risks resulting from the use of EHRs and other forms of health information technology.8-12 These concerns are compounded by the extraordinary pace of EHR development and implementation. Thus, the unique safety risks posed by the use of EHRs should be considered alongside the potential benefits of these systems.

At a time when institutions are focused heavily on achieving "meaningful use" requirements, we propose that clearer guidance be provided so that these institutions can align activities related to patient safety with the activities required to support a safe EHR-enabled health care system.13 A set of EHR-specific safety goals, modeled after the Joint Commission's National Patient Safety Goals, may provide organizations with areas of focus for sustained improvements in organizational infrastructure, processes, and culture as they adapt to new technology.

EHR implementation is still highly heterogeneous across health care systems and providers, and this heterogeneity leads to equally variable implications for patient safety. For instance, the priorities for patient safety in an organization in the midst of an EHR rollout differ from those of an organization that has used a fully integrated EHR system for 5 or more years. To account for the variation in the stages of implementation and levels of complexity across clinical practice settings, we propose a three-phase framework for the development of EHR-specific patient-safety goals (e-PSGs). The first phase of the framework, aimed at all EHR users but especially at recent and future adopters, includes goals to mitigate risks that are unique and specific to technology14 (e.g., technology that is unsafe owing to unavailable or malfunctioning hardware or software). The second phase addresses issues created by the failure to use technology appropriately or by misuse of technology.15 The final phase focuses on the use of technology to monitor health care processes and outcomes and identify potential safety issues before they can harm patients.16 This framework can lay the foundation for the development of e-PSGs within the context of EHR-enabled health care.

## Goals

### Phase 1: Address Safety Concerns Unique to EHR Technology

Device failures and both natural and man-made disasters are inevitable. The potential consequences of an EHR failure become of increasing concern as large-scale EHR systems are deployed across multiple facilities within a health care system, often across a wide geographic area. These broadly distributed systems may be tightly coupled and lightning fast, but that also means that a malfunction can rapidly affect not only a single department or institution but possibly an entire community.17 Furthermore, because the operations of such systems are often decentralized and relatively opaque to end users,18 problems evade easy detection and

solution. In a recent example, on April 21, 2010, one third of the hospitals in Rhode Island were forced to postpone elective surgeries and divert non–life-threatening emergencies19 when an erroneous automatic antivirus software update set off a chain of events that caused “uncontrolled [computer] restarts and loss of networking functionality.”20 A potential e-PSG, therefore, should be to reduce the effect of EHR downtime on clinical operations and patient safety. Table 1Table 1Framework for Potential EHR-Related National Patient Safety Goals. lists some of the activities that organizations could undertake to achieve this goal.

Safety can also be compromised as a result of miscommunication between the components of an EHR system. For example, it is not uncommon for data-translation tables (used to encode and decode orders transmitted between disparate systems) to have mismatched data fields.34 These mismatched fields may affect orders by introducing inadvertent changes that are virtually undetectable by the computer or by the people not privy to the original sender's intentions. An example of such an error is an order for 30 mg of oxycodone, sustained release, that is correctly entered in the computer-based provider order entry (CPOE) system but erroneously mapped to 30 mg of oxycodone, immediate release, in the pharmacy management system and incorrectly dispensed. Errors related to the transfer of information between systems may be detected by testing interacting components within the “live” EHR environment. However, this process is resource-intensive and therefore may not be carried out with adequate effort or attention. Therefore, an e-PSG could focus on reducing the miscommunication of data transmitted between different safety-critical components of the EHR. Recent evidence has shown that EHR accessibility and information transfer are two of the most common problems reported in EHR-related safety events.9,11,12

Phase 2: Mitigate Safety Concerns Arising from Failure to Use EHRs Appropriately

One rationale for widespread use of EHRs is that certain patient harms can be prevented when EHRs are used appropriately. For instance, EHRs can facilitate and standardize the transfer of information between providers and help close the communication loop by promptly notifying providers when test results are abnormal. However, these benefits are predicated on the assumption that EHRs will be used correctly and as intended in routine practice.35 For example, if CPOE systems were to be used on some nursing units but not others, clinicians would need to check for orders and test results in multiple locations, increasing the likelihood that some information would be overlooked. Other partial uses of CPOE may leave noncomputerized processes more vulnerable to error. For example, if CPOE is used to order medications but not laboratory tests, there would be no way of ensuring closed-loop electronic communication of test results to the ordering providers, potentially leading to more missed results.36 Another hazard can arise if providers bypass structured data fields in CPOE and instead use EHR-based free-text communication to prescribe or discontinue medications, since free-text orders are not standardized and are vulnerable to miscommunication.37 To reduce these safety concerns, another e-PSG could be to mandate the use of CPOE for all medication orders, laboratory tests, and radiologic tests. Table 1 lists several strategies that may help to achieve this goal.

Second, the implementation and use of complex clinical-decision support (CDS) systems embedded in EHRs are prone to human error and cognitive constraints.38,39 Consequently, decisions related to various aspects of CDS interventions must be evaluated periodically.40 For example, although point-of-care CDS interventions are necessary to achieve the full benefits of EHRs and stages

1 and 2 of the meaningful use payments, outlined by the Centers for Medicare and Medicaid Services (CMS),41 alerts that interrupt the clinician's workflow or thought process must be used judiciously. Many organizations turn on alerts with low specificity, which results in high rates of clinician override.24 Frequent overrides are associated with “alert fatigue,” which can lead clinicians to inadvertently ignore important information. Thus, another potential e-PSG could be to reduce alert fatigue. Alerts with override rates above a certain threshold should be discontinued or modified to increase their specificity.42 Similarly, hard stops (i.e., when users cannot proceed with the desired action) must be used only for the most egregious errors.43 Having such a goal will stimulate a multidisciplinary approach to reducing alerts that involves engaging cognitive scientists, human-factors engineers, and informaticians (i.e., scientists trained to work on the sociotechnical issues of information and communications technologies44,45) to work on these complex issues with clinicians (Table 1).

Third, although there is increased safety associated with integrating free text, dictated reports, radiographic images, and other test results into EHRs (including improved legibility and rapid access),46 many institutions are not currently coding some of the critical data needed to maximize safety. The lack of structured or coded data prevents the system from being able to provide the user with meaningful feedback or interpretation (i.e., an alert regarding the use of lisinopril will not be generated if a patient's history of captopril-related angioedema has not been entered as coded allergen data). Therefore, to realize the full safety benefits of complex CDS tools47 (e.g., checks for drug allergies,48 automatic notification of abnormal test results,28 or reminders related to drug-condition interactions29 [e.g., a warning on the use of isotretinoin in patients who are pregnant]), another e-PSG could focus on ensuring that critical data on medications, allergies, diagnostic test results, and clinical problems are entered as structured or coded data in the EHR.49

Phase 3: Use EHRs to Monitor and Improve Patient Safety

To achieve the goals of many national initiatives to improve patient safety and to facilitate the prevention of safety events, electronic data must be used to help detect, manage, and learn from potential safety events in near real-time. The stakeholders include the Agency for Healthcare Research and Quality (AHRQ), the Joint Commission, and the recently formed Partnership for Patients.50 In the current methods used to measure safety events, there is an overreliance on incident reports, which detect only a small proportion of events.32 In contrast, systems can be programmed to automatically detect easily overlooked and underreported errors of omission, such as patients who are overdue for medication monitoring, patients who lack appropriate surveillance after treatment, and patients who are not provided with follow-up care after receiving abnormal laboratory or radiologic tests results.51 EHR-based trigger approaches can also be used to detect errors of commission related to preventable adverse drug events,52 postoperative complications,53 and misidentification of patients.54 Organizations must leverage EHRs to facilitate rapid detection of common errors (including EHR-related errors), to monitor the occurrence of high-priority safety events, and to more reliably track trends over time. EHRs could also play a role in improving the existing infrastructure of reporting to patient-safety organizations by facilitating the generation of data files describing particular safety events (e.g., using the AHRQ common format version 1.2).55 Thus, an e-PSG could relate to the use of the EHR to monitor, identify, and report potential safety issues and events. This would make detection and reporting more efficient and help shift resources toward investigation and action.

## Application of the Three-Phase e-PSG Framework

Given that only 48% of all eligible hospitals and only 20% of eligible physicians have currently attested to achieving stage 1 of the CMS meaningful use criteria,56 the development and application of e-PSGs could partially address the Institute of Medicine's recent recommendation to create an EHR safety action and surveillance plan.8 The recommendations of such a plan should be tailored to the stage of EHR implementation. Recent adopters of EHRs could focus on the goals presented in phase 1 of our safety framework, making sure that the technology is safe to use, whereas organizations that have already achieved stage 1 meaningful-use criteria and have been using EHRs for several years could aim for goals from all three phases. Measurements related to e-PSGs would allow nationwide tracking and benchmarking of EHR-related safety performance.57 Policymakers and EHR vendors could collaborate on the development and certification of automated methods to measure and report new indicators annually from meaningful use certified EHRs in eligible hospitals. Examples of potential measures for e-PSGs might include EHR uptime rate (e.g., minutes the EHR was available to clinicians divided by number of minutes in a year23), CPOE rate (e.g., number of orders electronically entered divided by the total number of orders during the year23), and alert override rate (e.g., number of point-of-care alerts ignored divided by the total number of point-of-care alerts generated23).

These goals will also need to be reviewed regularly and updated as needed in accordance with national priorities and research on EHR-related patient safety. In addition, many strategies not addressed in this article could be considered as recommendations or good clinical practices and progress in a stepwise fashion to future e-PSGs.

## Summary

To create a coordinated, consistent, national strategy that will address the safety issues posed by EHRs, we propose that a concerted effort be made to improve health care safety in the context of technology use. This effort should address preventable risks that may hamper endeavors to create a safer EHR-enabled health care system. Further discussion and consensus among national agencies (e.g., the Office of the National Coordinator for Health Information Technology [ONC], the AHRQ, the Joint Commission, the Centers for Medicare and Medicaid Services) is clearly necessary for the adoption of future national patient-safety goals specific to EHR use. However, this approach must be given immediate priority considering the rapid pace of EHR adoption and the resulting changes in our nation's health care system. National EHR-related patient-safety goals are needed to address current problems with existing EHR implementations and failures to leverage current EHR capabilities. For instance, the ONC has recently taken several important steps in this direction with release of the revised 2014 EHR certification criteria (e.g., emphasis on user-centered design and application of quality management systems in the EHR design and development process58). Such efforts should be expanded in the future. Goals must be technically feasible, financially prudent, and practically achievable within current constraints and be accompanied by specific guidance on achieving them. Input on these goals must be sought not only from EHR developers and clinical end users but also from cognitive scientists, human-factors engineers, graphic designers, and informaticians with expertise in

patient safety in complex health care environments. Creating unique EHR-related national patient-safety goals will provide new momentum for patient-safety initiatives in an EHR-enabled health system.

The views expressed in this article are those of the authors and do not necessarily represent the views of the Department of Veterans Affairs or any of the funding agencies listed below.

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

Supported by a Strategic Health IT Advanced Research Projects (SHARP) Program contract from the ONC (10510592) (to Dr. Sittig); a career-development award from the National Institutes of Health (K23CA125585) (to Dr. Singh); the Veterans Affairs (VA) National Center for Patient Safety; the Agency for Health Care Research and Quality (R18HS017820); and the Houston VA Health Services Research and Development Center of Excellence (HFP90-020). These sources had no role in the preparation, review, or approval of this article.

We thank Michael Shabot, M.D., Eric Thomas, M.D., M.P.H., and Robert Murphy, M.D., for their comments on an earlier version of this article; and Annie Bradford, Ph.D., for assistance with the editing of an earlier version of the manuscript.

Source Information

From the University of Texas–Memorial Hermann Center for Healthcare Quality and Safety, School of Biomedical Informatics, University of Texas Health Sciences Center (D.F.S.), the Houston VA Health Services Research and Development Center of Excellence, and the Houston VA Patient Safety Center of Inquiry, Michael E. DeBakey VA Medical Center (H.S.), and the Section of Health Services Research, Department of Medicine, Baylor College of Medicine (H.S.) — all in Houston.

BACK TO TOP

## Cholinergic Autonomic Dysfunction in Veterans With Gulf War Illness

Archives of Neurology
Robert W. Haley, MD; Elizabeth Charuvastra†, RN; William E. Shell, MD; David M. Buhner, MD; W. Wesley Marshall, MD; Melanie M. Biggs, PhD; Steve C. Hopkins, BS; Gil I. Wolfe, MD; Steven Vernino, MD, PhD
November 2012

Background

Authors of prior small studies raised the hypothesis that symptoms in veterans of the 1991 Gulf War, such as chronic diarrhea, dizziness, fatigue, and sexual dysfunction, are due to cholinergic autonomic dysfunction.

Objective

To perform a confirmatory test of this prestated hypothesis in a larger, representative sample of Gulf War veterans.

Design

Nested case-control study.

Setting

Clinical and Translational Research Center, University of Texas Southwestern Medical Center, Dallas.

Participants

Representative samples of Gulf War veterans meeting a validated case definition of Gulf War illness with 3 variants (called syndromes 1-3) and a control group, all selected randomly from the US Military Health Survey.

Main Outcome

Measures  Validated domain scales from the Autonomic Symptom Profile questionnaire, the Composite Autonomic Severity Score, and high-frequency heart rate variability from a 24-hour electrocardiogram.

Results

The Autonomic Symptom Profile scales were significantly elevated in all 3 syndrome groups ($P < .001$), primarily due to elevation of the orthostatic intolerance, secretomotor, upper gastrointestinal dysmotility, sleep dysfunction, urinary, and autonomic diarrhea symptom domains. The Composite Autonomic Severity Score was also higher in the 3 syndrome groups ($P = .045$), especially in syndrome 2, primarily due to a significant reduction in sudomotor function as measured by the Quantitative Sudomotor Axon Reflex Test, most significantly in the foot; the score was intermediate in the ankle and upper leg and was nonsignificant in the arm, indicating a peripheral nerve length–related deficit. The normal increase in high-frequency heart rate variability at night was absent or blunted in all 3 syndrome groups ($P < .001$).

Conclusion

Autonomic symptoms are associated with objective, predominantly cholinergic autonomic deficits in the population of Gulf War veterans.

BACK TO TOP

## Epidemiology of Contemporary Seroincident HIV Infection in the Navy and Marine Corps

Military Medicine
CDR David M. Brett-Major, MC USN; Shilpa Hakre, DrPH; CAPT Neal A. Naito, MC USN (Ret.); CAPT Adam Armstrong, MC USN; CAPT Eric A. Bower, MC USN (Ret.); COL Nelson L. Michael, MC USA; Paul T. Scott, MD
November 2012

ABSTRACT

Human Immunodeficiency Virus (HIV) infection continues at a steady rate among U.S. Sailors and Marines. This study provides the first service-specific description of HIV infection demographics. All Sailors and Marines identified as HIV infected between January 2005 and August 2010 were included. The project compared personnel and epidemiologic data, and tested reposed sera in the Department of Defense Serum Repository. This group comprised 410 Sailors and 86 Marines, predominantly men. HIV infected Marines were more likely to be foreign born than their Navy counterparts, 42% versus 10%, $p < 0.001$. Approximately half of the patients had deployed including to the wars in Iraq or Afghanistan. Nearly half of each group was infected by the age of 25. Similar to the U.S. epidemic, Black race was over-represented. Unlike national rates, Hispanic Sailors and Marines were not over-represented. Demographics were distinct for those of specific occupational specialties. Certain ship classes carried lower incidences. Clustering of HIV infection risk occurred around deployment. The Navy and Marine Corps have different patterns of HIV infection, which may merit distinct approaches to prevention. The Navy may have unique targets for prevention efforts to include pipeline training and first assignment as well as particular occupational environments.

BACKGROUND

Human Immunodeficiency Virus (HIV) infection has occurred in the U.S. military since early in the epidemic. Service members identified as HIV infected through compulsory, periodic force-wide screening quickly receive subspecialty, multidisciplinary care and most often are retained in service. Detection of HIV infection and subsequent case management is coordinated through the Navy Bloodborne Infection Management Center (NBIMC). Infected Sailors and Marines are evaluated periodically at designated referral centers and assisted in appropriate duty assignments through the Navy Personnel Command and Headquarters Marine Corps. Mechanisms to ensure case finding and management were constructed to do so while maximizing patient privacy and avoiding stigma. However, these mechanisms have made the routine exploration of disease transmission in the force challenging as systematic compilation of potential exposures has not occurred.

The epidemiology of HIV infection in the Navy and Marine Corps has not been fully described. However, detailed reports of features of infection were published nearly a decade ago with data from a mixed group of Sailors and Marines and more broadly across services from the military's robust HIV natural history study cohort. Despite aggressive case finding and treatment as well as force-wide sexual health strategies, Navy and Marine Corps HIV incidence and prevalence have remained stable over the last several years. In 2008, 2009, and 2010, Navy HIV incidence in the active duty component was 0.35, 0.35, and 0.30, respectively per 1,000 persons tested. Incidence in the reserve component was 0.23, 0.22, and 0.42. During this same period, Marine HIV incidence in the active duty component was 0.13, 0.15, and 0.12 per 1,000 persons tested. Reserve component incidence was 0.27, 0.20, and 0.24. Furthermore, effect on HIV infection in the force of repealing restrictions on professed sexual orientation of service members is not known. Optimizing force HIV screening and prevention is a continued challenge in this setting. Early detection may avert transmission through sexual health counseling and treatment. An improved understanding of the contemporary epidemiology of HIV acquisition among the active duty military force may enhance primary prevention efforts.

We sought to characterize contemporary HIV infection in the Navy and Marine Corps. Our objectives were to provide actionable public health intelligence to include informing education and screening practices, identifying transmission networks, assessing program performance and informing practices and providing military medical research entities with sufficient context to continue their work in promoting disease prevention. We report the descriptive epidemiology from this effort.

## METHODS

This project identified all enlisted Sailors and Marines with a first known positive HIV screening test between January 1, 2005 and August 3, 2010. Officers comprise a very small fraction of patients and are demographically distinct so this group was not included.

Candidates for entering military service are screened for HIV infection. If found positive, they are not brought into military service. Once in the service, Sailors and Marines are tested every 2 years. Additional testing occurs as clinically indicated or as related to predeployment and postdeployment practices. Force HIV screening and timely clinical testing is performed through a single contract managed by the NBIMC, though in the presence of clinical suspicion some providers also seek local just-in-time testing. From January 2005 through April 2009, screening utilized a second generation enzyme-linked immunosorbent assay (ELISA) confirmed through Western blot. In May 2009, testing shifted to the use of a third generation ELISA confirmed through Western blot. Program managers and clinicians also obtain specialized testing including Nucleic Acid Testing at the HIV Diagnostics and Reference Laboratory at the U.S. Military HIV Research Program (MHRP), Walter Reed Army Institute of Research (WRAIR). This is requested most frequently in the context of acute seroconversion illness.

This project was coordinated by the Bureau of Medicine and Surgery, NBIMC, and the MHRP. It was reviewed by the WRAIR Institutional Review Board and affirmed as a public health activity. The Navy Personnel Command and the Headquarters, U.S. Marine Corps (USMC) provided personnel histories from initial accession into service through the date of first known positive HIV screening test. The provided data elements included place of birth and home of record (HOR), age, gender, marital status, level of education, current rating (Navy), rank (USMC), assignment units and locations, and represented patient status at the time of data pull. The Epidemiology Data Center, Navy Marine Corps Public Health Center provided pre-HIV infection clinical histories and deployment data. Identified linkages were reposed in a secure database within the Data Collection and Analysis Center, MHRP.

The distribution of these demographic variables was described. For Navy race data, 10% of Sailors who self-identified as Other were manually verified through clinical database information populated by the Defense Eligibility Enrollment Reporting System. As Hispanic was a race category for Marines but not Sailors, Hispanic ethnicity of Sailors was tallied regardless of their race selection.

Presumed last negative (LN) serum samples and when necessary earlier samples from the Department of Defense Serum Repository, Armed Forces Health Surveillance Center (Silver Spring, Maryland) were tested to verify the true LN test and true first positive test dates through both third generation ELISA and nucleic acid testing. These revised dates were used to construct risk periods for acquisition of HIV infection and describe potentially relevant exposure risks.

The dates of initial positive (IP) HIV screening test, LN HIV screening test, and previous negative HIV screening test were not corrected when used to report differences in testing intervals among the services. However, IP and LN were corrected to account for false negative serum from the Department of Defense Serum Repository when generating risk windows and calculating the interval between deployment end date and IP HIV test. Risk windows were bound by the dates of true LN and true first positive HIV test results on pulled sera—dates when each patient was known to be HIV uninfected until known to be HIV infected.

Several analyses were incorporated for the purpose of hypothesis generation including exploration of occupational codes, duty assignment, and deployment history.

The degree to which different Sailor occupational specialties (ratings) were represented among NBIMC enrollees were normalized employing a 5-year average of annual force strengths provided by the Navy Personnel Command as denominators. Longitudinal rating data on Sailors was assessed against the first positive HIV test date so that if a Sailor had changed ratings, the rating before known HIV infection was selected. Ratings were coded as this portion of the analysis was exploratory, confounding may be present and undue stigma is a risk.

Sailors' assignments on ships during their risk windows were analyzed in order to generate an incidence of subsequently HIV-infected Sailors by ship class. The Navy Ship List was utilized for number of vessels in service by class as well as standard crew complements. Rate ratios were calculated to compare the rate of HIV incidence by ship class.

Potential associations between deployment and HIV positivity were explored by calculating the intervals between deployment start dates and true first positive and LN HIV test dates. Means for time from deployment start to first positive HIV test were compared using a *t*-test.

Characteristics among Sailors and Marines were compared using the $\chi^2$ test at an $\alpha$ of 0.05. Incidence rates were assessed for significance by constructing 95% confidence interval (CI)s. Data sets were managed and analyzed utilizing Statistical Analysis Software version 9.2 (SAS Cary, North Carolina).

## RESULTS

The analysis group was comprised of 410 Sailors and 86 Marines. In general, Sailors and Marines identified within the study period had IP HIV screening tests within force mandated screening windows following their LN test (83% of Marines' IP tests and 78% of Sailors' were tested within 24 months of their LN test. More than 50% of both Sailors and Marines had been deployed with 45% and 14%, respectively, having deployed previously to the wars in Iraq and Afghanistan (Operations Iraqi Freedom or Enduring Freedom). Across all deployments, approximately 1 in 5 Sailors and Marines had deployed more than once.

Sailors and Marines were similar in a variety of demographic variables assessed at date of IP serum HIV ELISA, such as gender and fraction with higher education. Sixteen percent of Sailors self-identified as Hispanic under ethnicity, similar to the number of Marines who identified as Hispanic race. The distribution of length of service (LOS) and age at time of diagnosis also were similar, though of

HIV-infected personnel who were identified when over the age of 25 years, Sailors were older than Marines (more likely to be over 30 years old).

HIV-infected Marines were more likely to be born outside of the continental United States, Alaska, and Hawaii than their Navy counterparts representing 42% rather than 10% of those populations. Among Marines, this included 4 individuals from North America, 3 from Europe, 3 from the Caribbean and Central America, 4 from South America, 15 from Asia and the Pacific Islands, and 7 from Africa. Among Sailors, this included 6 individuals from North America, 3 from Europe, 16 from the Caribbean and Central America, 2 from South America, 14 from Asia and the Pacific Islands. Within the United States, Sailors had twice the rate of being born in the southern states than Marines.

HIV-infected Marines were widely distributed across Marine Occupation Specialties (not shown). However, some Sailor occupational ratings demonstrated increased rates of HIV infection in the study period.

Rating F was the most prevalent rating among incident HIV infections. The demographics of HIV-infected Sailors in rating F differed in several ways from the remainder of infected Sailors, and from rating B, the second most prevalent rating. HIV-infected Sailors in rating F were more likely to be a mixed group of self-reported White (39%) and Black (38%) with fairly well distributed regions of birth and homes of record. Those in rating B, the second most prevalent, more commonly self-reported as Black (76%) from the South (82%). Overall, Sailors in rating F were less likely to have deployment experience, but more likely to have deployment experience in Iraq or Afghanistan—18% for rating F, 12% for rating B, 4% for all other ratings. In rating B, 2 in 5 Sailors were identified within 2 years of service in contrast to 1 in 4 of those in other ratings.

HIV-infected Sailors and Marines were well distributed during their risk window for acquisition of infection across shore assignment locations relative to general size of force concentrations (data not shown). When we assessed those Sailors who were assigned to sea duty during their calculated periods of risk for acquisition of HIV infection, the incidence of subsequent identification of HIV infection was lowest for crews of aircraft carriers and ballistic missile submarines HIV infection incidence among Sailors who served on ships. Comparison of HIV acquisition risk window assignments and per ship class incidence. SSN = fast attack submarine. SSBN = ballistic missile submarine, counts including SSGN = guided missile submarine and variants. FFG = Frigate. DDG = Destroyer. CG = Cruiser. LPD/ LSD = small deck amphibious vessels. LHA/LHD/LCC = large deck amphibious vessels and control ships. CVN = aircraft carriers. *LHA/LHD/LCC had the largest and only statistically significant rate ratio when compared to CVN, RR 1.75, 95% CI 1.09 to 2.80.

Large deck amphibious vessels had the highest incidence of HIV compared to aircraft carriers [rate ratio (RR): 1.75; 95% CI: 1.09–2.80]. Although all other ship types, except ballistic missile submarines, had a higher incidence compared to the incidence on aircraft

carriers, none reached statistical significance [RR: 1.41–1.75; 95% CI: 0.76–2.80]. For shore commands and among Marines, HIV infection density in general matched what might be expected with large concentrations of Sailors and Marines. Large bases had the greatest number of new infections (data not shown).

Among those Sailors and Marines with a history of deployment, how closely the IP HIV test occurred in relation to their last deployment varied by deployment type and service. Any plot below the *x*-axis represents a last true negative HIV test result which precedes the deployment. Any plot to the right of the *y*-axis represents a first true positive which follows the beginning of the deployment. This lower right quadrant encompasses most deploying Sailors and Marines later identified as HIV infected. As time passes from the start of the deployment, the density in this quadrant decreases for both Sailors and Marines, most notably for those who had deployed to Iraq and Afghanistan. 53% of Marines and 57% of Sailors who had deployed to the wars in Iraq or Afghanistan were identified as HIV positive within 1 year of their deployment end date. The mean time from deployment start to first positive HIV test differed between those who deployed to Iraq or Afghanistan as compared to other theaters, 20 versus 29 months, respectively ($p < 0.001$).

## DISCUSSION

Two important ideas emerge from this project: work factors are important in the acquisition of HIV infection in the military and HIV infection in the Navy is different than that in the Marine Corps.

This project revealed that the current HIV population in the Navy and Marine Corps differs from the HIV-infected service members previously described. Nearly half of this project's population self-reports as Black in contrast to less than a third as previously described, and the mean age of this population is 2 to 3 years younger. The large fraction of infection occurring early in service suggests that the early enlistment period, in particular social and behavior patterns during occupational schools leading up to service in a first operational unit, merits closer investigation.

Although our analysis of occupational features was meant to explore hypotheses for further study, evidence for this also exists in disparities in disease count (rather than incidence) by rating and ship class. They are intriguing for the possibility of unique social networking which occurs in formative development in service as well as that contingent upon ship assignment, perhaps because of a mix of crew size, ship culture, operational tempo, and vessel homeport.

Sailors demonstrated different incidences of HIV infection by rating. Lateral transfer of HIV-infected Sailors during the period of interest was rare. The most prevalent rating, F, had not received any known HIV-infected Sailors by transfer. The combination of

young age and short LOS together with clustering of infection in certain ratings implicates periods late in pipeline training and early post-training for HIV infection. They highlight the importance of promoting healthy social networks and sexual health habits during accession and training.

Rates of HIV infection amongst shore activities were not instructive. When ships were assessed, however, while assignment on larger vessels during the acquisition risk window contributed the highest number of later identified HIV infections, rates of subsequent infection were lowest among those who served on aircraft carriers and ballistic missile submarines and highest on large deck amphibious ships. Whether this is a coincidence or service on these vessels convey variable risk because of social or other dynamics is not clear from this data.

Sailors and Marines deploy or take other assignments globally, move about the United States and have unique stressors. They are young and typically working in areas removed from family and other pre-existing support networks. Consequently, work specialization and assignment have a potentially large impact on the personal life of service members. In the deployment environment, unique occupational exposures are present both because of exposure to host nation citizens and potential use of the walking blood bank for exigent treatment of life-threatening trauma.

Our data suggest that deployment must be further explored as a potentially important event predisposing to HIV infection. Predeployment, on deployment, and postdeployment time periods may be important particularly for those who had deployed to Iraq and Afghanistan. Personnel who deployed to Iraq and Afghanistan were identified as HIV infected earlier than their counterparts who deployed to other theaters of operation. However, that difference may have been confounded by more stringent processes regarding the collection of postdeployment serum samples after deployment to Iraq and Afghanistan.

Though anchored with conventional contributing risks such as alcohol consumption and prior casual sex, the idea of contranormative situations leading to disinhibition in nonmilitary settings has been published. In the setting of spring break, this has been further explored to examine whether peers predefine acceptable behavior before going on vacation. These concepts may have analogs with liberty port, perideployment leave and liberty, temporary assigned duty, deployment activities, new accession to service and first assignments. Such relationships could have marked impact on relevance of targeted prevention measures to include administrative controls.

Although served by the same medical department, this report is the first to clarify HIV demographic differences between the Navy and Marine Corps. Sailors and Marines had some similarities but also fundamental differences in their demographics which could impact prevention measures and require service specific interventions. This also was true between different Sailor occupational specialties.

HIV-infected Sailors and Marines identified during the study period had similar distributions of age at IP test, length of service, and self-reported race. However, force differences in incidence as well as differences in region of birth, HOR, and deployment patterns, potential for contribution of occupational specialty and assignment type suggest that patterns of HIV acquisition in the two services differ. Southern birth among Sailors was twice as prevalent as among Marines, and particularly high in rating (occupational specialty) B. Rating wide demographics may confound this observation. Consequently, prevention and screening practices may need to be tailored not only by service, but within the Navy by rating. Military wide, approximately nearly one in three military accessions entered from southern states from 2005 through 2010 (personal communication, Navy Recruiting Command).

Recently reported U.S. incidences of HIV are not age adjusted and more than a third of HIV cases occur in age groups not represented by active duty personnel. This may contribute to why Navy HIV incidence is higher and Marine Corps incidence comparable to that estimated in the total civilian population in the United States (roughly 0.2 per 1,000 persons across 2006–2009), though the Navy employs universal screening and 1 in 5 U.S. civilians who live in high HIV disease burden areas are infected and may not be aware of their HIV infection status. Previous U.S. estimates reported rates of 0.27 to 0.43 per 1,000 persons among age groups of military interest. Those individuals self-identifying as Black, or Hispanic, and men who have sex with men continue to disproportionately represent incident U.S. HIV infections. Self-reported Black Sailors and Marines constitute nearly half of incident HIV infections. In contrast, the fraction of Hispanic Sailors and Marines among those with incident HIV infections is comparable to the force distribution of Hispanic Sailors and Marines overall. The Defense Equal Opportunity Management Institute reported that as of September 2010, active duty Black and Hispanic enlisted service members comprised approximately 20% and 17% of Sailors, 11% and 14% of Marines.

## LIMITATIONS

Control data with non-HIV infected service members was not part of this project. Non-equal opportunity demographic force data was not available. For instance, anecdotally, the Navy recruits most heavily from the South and all services differ in their geographic composition. However, specific data on this was not available to the investigators. Ship assignment incidences employed denominator data which did not incorporate embarked Marines or aircraft components such as carrier air wings, Marine elements, or Navy helicopter detachments. Also, Sailors had exposures other than ship assignment.

## CONCLUSIONS

The Navy and Marine Corps have different patterns of HIV infection which may merit distinct approaches to prevention. The Navy in particular may have unique targets for prevention efforts to include pipeline training and first assignment as well as particular occupational environments. Proactive operational research is required to further elucidate the nature of these acquisition dynamics en route to tailored prevention and surveillance strategies which maintain the health of the fighting force. Influences that predate accession into the services, potential role of occupational rating populations and pipelines, the impact of deployment and the role of occupational settings such as ship assignment should be further explored.

BACK TO TOP

## Association of Warfarin Therapy Duration After Bioprosthetic Aortic Valve Replacement With Risk of Mortality, Thromboembolic Complications, and Bleeding

Journal of the American Medical Association
Charlotte Mérie, MD; Lars Køber, MD, DMSc; Peter Skov Olsen, MD, DMSc; Charlotte Andersson, MD, PhD; Gunnar Gislason, MD, PhD; Jan Skov Jensen, MD, PhD, DMSc; Christian Torp-Pedersen, MD, DMSc
28 Nov 2012

Abstract

Context
The need for anticoagulation after surgical aortic valve replacement (AVR) with biological prostheses is not well examined.

Objective
To perform a nationwide study of the association of warfarin treatment with the risk of thromboembolic complications, bleeding incidents, and cardiovascular deaths after bioprosthetic AVR surgery.

Design, Setting, and Participants
Through a search in the Danish National Patient Registry, 4075 patients were identified who had bioprosthetic AVR surgery performed between January 1, 1997, and December 31, 2009. Concomitant comorbidity and medication were retrieved. Poisson regression models were used to determine risk.

Main Outcome Measures

Incidence rate ratios (IRRs) of strokes, thromboembolic events, cardiovascular deaths, and bleeding incidents by discontinuing warfarin as opposed to continued treatment 30 to 89 days, 90 to 179 days, 180 to 364 days, 365 to 729 days, and at least 730 days after surgery.

Results

The median duration of follow-up was 6.57 person-years. Estimated rates of events per 100 person-years in patients not treated with warfarin compared with those treated with warfarin with comparative absolute risk were 7.00 (95% CI, 4.07-12.06) vs 2.69 (95% CI, 1.49-4.87; adjusted IRR, 2.46; 95% CI, 1.09-5.55) for strokes; 13.07 (95% CI, 8.76-19.50) vs 3.97 (95% CI, 2.43-6.48; adjusted IRR, 2.93; 95% CI, 1.54-5.55) for thromboembolic events; 11.86 (95% CI, 7.81-18.01) vs 5.37 (95% CI, 3.54-8.16; adjusted IRR, 2.32; 95% CI, 1.28-4.22) for bleeding incidents; and 31.74 (95% CI, 24.69-40.79) vs 3.83 (95% CI, 2.35-6.25; adjusted IRR, 7.61; 95% CI, 4.37-13.26) for cardiovascular deaths within 30 to 89 days after surgery; and 6.50 (95% CI, 4.67-9.06) vs 2.08 (95% CI, 0.99-4.36; adjusted IRR, 3.51; 95% CI, 1.54-8.03) for cardiovascular deaths within 90 to 179 days after surgery.

Conclusion

Discontinuation of warfarin treatment within 6 months after bioprosthetic AVR surgery was associated with increased cardiovascular death.

Biological prostheses are preferred to mechanical valves for aortic valve replacement (AVR) surgery in elderly patients older than 65 years because of shorter life expectancy and lack of a need to use anticoagulation treatment in the long term. Especially in these patients, the trade-off between thromboembolic complications due to the valve implant and bleeding events as adverse effects from anticoagulation therapy must be balanced. Nevertheless, appropriate duration of anticoagulation treatment postoperatively is yet to be established because the risk of complications when the treatment is discontinued is unknown. Current guidelines1- 3 of anticoagulation treatment after bioprosthetic AVR surgery recommends 3 months of warfarin treatment. However, this recommendation is primarily based on results from 1 retrospective study4 with limited number of events hampered by the observational design as well as low power like all other studies within this scope.

In this nationwide study, we investigated whether discontinuation of warfarin treatment within prespecified periods after bioprosthetic AVR surgery was associated with increased risk of thromboembolic complications, cardiovascular death, and bleeding incidents during a period of 13 years.

Method

Participants and Study Design

All patients having bioprosthetic AVR with or without coronary artery bypass graft (CABG) surgery performed between January 1, 1997, and December 31, 2009, in Denmark were identified through a search in the Danish National Patient Registry. Patients with prior cardiac surgery or other concomitant surgical procedures were excluded. The Danish Data Protection Agency approved the use of personal data (reference 2007-58-0015, internal reference GEH-2010-001). Retrospective studies do not require ethical approval in Denmark.

Hospital diagnoses, procedures, medication, age, sex, and causes of death were retrieved from the National Danish Registries. Due to a law-imposed demand of reporting these data, a high degree of completion is ensured. The Danish National Patient Registry contains information on all hospital admissions and diagnoses in Denmark since 1978. The diagnoses are classified according to the International Classification of Diseases, 8th Revision (ICD-8) and International Statistical Classification of Diseases, 10th Revision (ICD-10) (ie, patients with stroke were identified by stroke diagnoses recorded in the National Hospital Registry as codes I61, I62, I63, and I64). Previously, these diagnoses have proven valid with positive predictive values of 74% to 97%.5- 6 Similarly, to retrieve information on comorbidity, we identified prespecified discharge diagnoses (for ICD-8 and ICD-10 codes, see eTable 1) in the Danish National Patient Registry dating back 18 years preoperatively. Patients with diabetes were identified through their use of glucose-lowering medication.

Main Outcome Measures

Primary events studied included stroke, thromboembolic complications (ie, ischemic strokes, myocardial infarctions, and peripheral arterial emboli), bleeding incidents (ie, gastrointestinal, intracranial, urinary tract, and airway bleedings7), and cardiovascular death (eTable 2). The term cardiovascular death included all deaths, with a cardiovascular cause registered on the death certificates forming the basis of the Danish National Death Registry.

Medication

Use of medication was obtained from the Danish Registry of Medicinal Product Statistics where all prescriptions dispensed from Danish pharmacies since 1995 are recorded. Registration is complete due to linkage to reimbursement from the state, and medication from other sources is limited.8 The information for each medication coded according to the Anatomical Therapeutic

Chemical (ATC) System included number and strength of tablets. Calculations regarding the daily warfarin dosage were restricted by a minimal dose of warfarin of 0.625 mg and a maximal dose of 17.5 mg. A patient starting warfarin therapy was assumed to start on a default dose of 5 mg. Whenever a new prescription was claimed, the average dose of up to 3 previous prescription periods was calculated. Those prior prescription periods, which could represent a continuous treatment with at least the minimal dose, were included in that calculation. To avoid conditioning on the future, later prescriptions were not included in calculations. Patients were assumed to continue with the dose below this dose that is a multiple of 0.625 mg (one quarter of a tablet). The minimal and maximal dose set limits to this calculation. Discontinuance of treatment was assumed when there were no tablets left according to the calculations. The value of these calculations are dependent on the calculations not being highly sensitive to assumptions. Sensitivity analyses were performed by using a minimal dose of 5 mg and a default dose of 7.5 mg, and by using an automatic lengthening of any treatment period of either 14 or 30 days.

Information obtained on other medication at the time of surgery required that each patient claimed at least 1 prescription on prespecified medication within 3 months before surgery.

Statistical Analyses

The studied cohort originates from a fixed population defined by all individuals in Denmark alive and aged older than 18 years on or after January 1, 1997, and patients can only be lost to follow-up by emigrating. In the current study period, 3 patients (0.07%) emigrated and were censored at time of emigration. We determined the occurrence of events 30 to 89 days, 90 to 179 days, 180 to 364 days, 365 to 729 days, and at least 730 days after surgery. At these time points, analysis time was split and event rates for patients with and without warfarin treatment were calculated as the number of events divided by the sum of person-time. Patients were censored at event, emigration, or end of follow-up. Poisson regression models were used for survival type analysis to calculate incidence rate ratios as an indicator of relative risk (RR). With an exposure ratio of warfarin treatment vs nontreatment of 3.6, a risk of cardiovascular death of 7.5% among patients without warfarin and a type 1 error risk of .05, we obtained a power of 66% to detect a difference in RR of 1.3, 86% power for an RR of 1.4, and 96% for an RR of 1.5. Subgroup analyses were preplanned and intended to challenge the results by removing patients with a variety of profiles.

All P values reported were 2-sided and considered statistically significant if $P < .05$. Statistical analyses were performed by using SAS version 9.2 (SAS Institute) and Stata version 11.1 (StataCorp LP).

Results

A total of 4075 patients were identified who had bioprosthetic AVR surgery performed after excluding patients in warfarin treatment before surgery (n = 684) and patients with a diagnosis of atrial fibrillation within 30 days after surgery (n = 1215). A total of 301 patients were both in warfarin treatment before surgery and had atrial fibrillation within 30 days after surgery. In addition, 8 patients discontinued warfarin treatment within the first 30 days after surgery. This period was omitted from the analyses ensuring all patients had an equal chance to claim a prescription for warfarin. A flow diagram of patients included in our study is shown in Figure 1. Mean age was 74.6 years (range, 18-95 years) and 1670 patients (41%) were women. Baseline characteristics by warfarin treatment are shown in Table 1.

To determine the importance of warfarin in antithrombotic treatment after bioprosthetic AVR surgery, we compared the effects of warfarin, aspirin, the combination of warfarin and aspirin, and no treatment on cardiovascular death after surgery (Figure 2 and eTable 3). Monotherapy with aspirin was used infrequently and rates of cardiovascular death did not differ significantly from rates among patients with no treatment; therefore, to provide comparative estimates for the effect of anticoagulation, the remaining analyses were made between warfarin vs no warfarin treatment independent of aspirin treatment.

Overall, 361 patients (8.9%) experienced a stroke, 615 (15.1%) had a thromboembolic event, and 364 (8.9%) encountered a bleeding incident after the date of surgery. During an observation period of 12 557 person-years (median duration of follow-up, 6.57 person-years), 1156 patients (28.4%) died, with 879 (76.0%) of these deaths being related to cardiovascular disease. Within 30 days after surgery, 109 patients (2.7%) experienced a stroke, 201 patients (5.0%) experienced thromboembolic events, and 75 patients (1.8%) had bleeding episodes. Thirty-day mortality was 6.0% and 1-year mortality was 10.9%.

The incidence rates of stroke, thromboembolism, bleeding incident, and cardiovascular death within prespecified periods after surgery with and without warfarin treatment are shown in Figure 3 (for overall mortality, see eFigure and eTable 4). The first postoperative month was omitted from these analyses ensuring all patients had an equal chance to claim a prescription. Table 2 shows the risk of strokes, thromboembolic events, cardiovascular deaths, and bleeding incidents over time unadjusted and adjusted for age, sex, concomitant CABG surgery, comorbidity, and calendar year.

Results from subgroup and sensitivity analyses are shown in eTable 5. Apart from the association found 180 to 364 days after surgery, changes in the estimated risks of cardiovascular death remained the same, except for the analysis excluding patients with concomitant CABG surgery.

Calculating the number needed to treat/number needed to harm within 90 to 180 days after surgery, we found that for every 23 (95% CI, 14-54) patients not being treated with warfarin, 1 patient died from cardiovascular cause; and for every 74 (95% CI, 27-95) patients being treated with warfarin, 1 patient experienced bleeding complications requiring hospital admission.

To exclude other factors explaining the increased risk of cardiovascular death 3 to 6 months after surgery other than discontinuing warfarin treatment, we reviewed diagnoses from hospital admissions within 1 month before cardiovascular death among patients with and without warfarin treatment dying within 3 to 6 months after surgery. We found no specific pattern in diagnoses explaining this finding. Similarly, we retrieved information on medication use within the first 6 months after surgery to investigate whether drugs prescribed could influence cardiovascular death 3 to 6 months after surgery and found no indication of changes in phamacotherapy.

Comment

This is to our knowledge the largest study examining thromboembolic complications with or without warfarin treatment after bioprosthetic AVR surgery. We demonstrated a clear benefit associated with warfarin during the initial 3 months after surgery, and furthermore our data suggest that extension of warfarin treatment to 6 months postoperatively may reduce the risk of cardiovascular death.

The guidelines regarding anticoagulation treatment after bioprosthetic AVR surgery1- 3 suggesting duration of warfarin treatment of 3 months postoperatively were all primarily based on a single observational study4 from 1995, including 424 patients with an antithrombotic regimen involving warfarin, aspirin, dipyridamole alone, or in combination. This study4 found the rate of thromboembolism decreasing significantly over time after surgery from 41% per year (1-10 days postoperatively) over 3.6% per year (11-90 days postoperatively) to 1.9% per year (>90 days after surgery). Furthermore, the authors found that warfarin reduced the risk of thromboembolic complications significantly. Accordingly, due to high risk of thromboembolic events within 90 days after surgery, the authors concluded that anticoagulation was indicated 3 months postoperatively—a finding which was confirmed by other similar small studies.9- 10

The aforementioned rate of 41% per year represented 5 thromboembolic events, with the rate of 41% being an extrapolation from the first 10 days assuming a constant thromboembolic risk throughout the following year. This assumption could be questioned because the release of microemboli during surgery11 would induce an initial high risk of thromboembolic complications immediately after surgery, which would decline rapidly the following month. This was reflected by the decrease in the thromboembolic rate reported 11 to 90 days postoperatively of 3.6% per year, even though this figure in fact only represented 3 thromboembolic events.

We found similar trends concerning thromboembolic events with initially higher rates decreasing the following year after surgery. However, our results were based on 615 thromboembolic events during the study period, with 40 thromboembolic events occurring within 30 to 89 days after surgery. Furthermore, we studied the effect of warfarin treatment on cardiovascular death, which has not previously been the subject of investigation.

In recent years, the benefit of treatment with anticoagulation after bioprosthetic AVR surgery has been challenged. An observational study12 analyzing data on 1151 patients undergoing bioprosthetic AVR surgery with or without CABG surgery found an incidence of cerebral vascular accidents of 2.4% among patients treated with warfarin compared with 1.9% among patients with no anticoagulation. The authors concluded that warfarin had no protective effect on neurological events but 23% of the patients were treated with warfarin preoperatively and, therefore, could be susceptible to stroke due to atrial fibrillation. A retrospective study13 of 500 patients with pericardial bioprostheses having 48 thromboembolic events found an increased risk of thromboembolism among patients treated with warfarin therapy compared with patients with no treatment, with a risk ratio of 3.0 after a 4-year follow-up period. However, this study population included patients with concomitant surgery on the ascending aorta and mitral annuloplasty, making findings difficult to interpret. Our study focused solely on patients with bioprosthetic AVR surgery and included a much larger population. Nevertheless, we did observe a minor increase in rates of thromboembolic events more than 1 year after surgery with warfarin therapy being associated with a worse outcome.

An observational study14 on 861 patients undergoing isolated bioprosthetic AVR surgery sought to determine whether aspirin or warfarin treatment was protective on the risk of thromboembolic complications during the first 90 days after surgery in an institution where antithrombotic treatment was administered according to the discretion of surgeons. The authors from that study14 concluded that early anticoagulation only seemed to reduce thromboembolic risk in high-risk groups (ie, women, highly symptomatic patients, and patients with small aortic prostheses). However, they did not attempt to address confounding by silent postoperative atrial fibrillation. Aspirin was used infrequently as monotherapy in our patients and consequently our study does not contribute valuable data for the use of aspirin as an alternative to warfarin.

Some guidelines recommend aspirin as an alternative to postsurgical warfarin therapy in patients without risk factors.1- 2,15 However, these recommendations are widely based on consensus opinion of experts, case studies, standard of care studies, or small prospective studies with insufficient power to substantiate findings. No prospective study has yet proved the safety of omitting antithrombotic treatment in patients with bioprosthetic AVR surgery in the first 3 months after surgery. Our study is no exception. No conclusions could be drawn due to paucity in patients receiving no treatment.

However, few observational studies have been centered on whether antiplatelet therapy is a sufficient protection against thromboembolic complications. For instance, 1 study16 analyzing 251 patients undergoing bioprosthetic AVR surgery, where treatment was determined by surgeons adherence with guidelines, compared warfarin treatment with aspirin in protecting against neurological events. The authors found no advantages of early anticoagulation but mean age and mean EUROSCORE were significantly higher in patients treated with warfarin. A prospective study17 assessing the efficacy of ticlopidine on thromboembolism included 235 patients with valve repair or bioprosthetic valve replacement in aortic, mitral, or tricuspid position. The authors concluded that ticlopidine seemed to prevent thromboembolism better than other therapy based on 2 and 4 thromboembolic events, respectively. Due to the relatively rare presentation of thromboembolic events and the devastating effects of this complication, large randomized studies comparing warfarin with antiplatelet therapy remains to be seen.18- 19

In addition, in our study a significant association was found between bleeding incidents and discontinuing warfarin 30 to 89 days postoperatively. A possible explanation may be that initiation of warfarin can be challenging to obtain the recommended international normalized ratio in some patients. Furthermore, it is not known whether these patients discontinued the warfarin treatment due to bleeding incidents or discontinued warfarin treatment and had a bleeding incident afterwards. However, considering the calculated number needed to treat/number needed to harm analysis previously shown, the association of bleeding complications with warfarin use seem to be less worrisome as opposed to a possible benefit from the association between warfarin use and decreased cardiovascular death.

Implications

With no randomized trials to guide the length of warfarin treatment, our results call for a review of guidelines in the field to consider an extension of the treatment to 6 months after surgery, especially in patients with an increased risk of cardiovascular death. Increased follow-up procedures may be instituted to monitor, especially in the oldest patients, bleeding complications, although the increased risk of bleeding found among patients treated with warfarin from 90 to 179 days was nonsignificant.20- 26

Study Limitations

We did not have access to information regarding the international normalized ratio. Thus, we were not able to assess time in therapeutic range, which most likely is more influential on the risk of thromboembolic events, and especially bleeding events, than prescription of warfarin alone.27 Similarly, we were not able to monitor adherence to treatment assignment. The fact that we had no data on the reasons for omitting, discontinuing, or extending warfarin treatment may introduce hidden confounders, including confounding by indication. However, warfarin discontinuation according to our estimation of use occurred when the patient had no

tablets left. Thus, for event driven discontinuation to occur, the patient should run out of tablets simultaneously with the occurrence of an event; in which case, the event could be attributed to warfarin treatment inducing an underestimation of warfarin effect on thromboembolic events and possibly an overestimation of the effect on bleeding events. However, we may have underestimated the risk of bleeding in warfarin treatment due to lack of information on bleeding episodes not resulting in hospitalization.

The main limitation of the databases is the lack of detailed clinical information related to important risk factors (eg, smoking), which could influence outcome. Although this is accounted for by including comorbidity, we cannot exclude the influence of some degree of residual confounding. The prescription register contains only data on prescribed drugs subject to reimbursement, making it vulnerable to changes in reimbursement policies. Drugs sold over-the-counter are not registered, introducing a problem regarding aspirin. However, the financial incentive for chronic patients to have aspirin prescribed and be reimbursed instead of buying the drug over-the-counter presumably reduces the sale of aspirin over-the-counter. Furthermore, reliance on administrative databases induces a risk of time-varying measurement errors of exposure as well as biases from survival effects.

## Conclusion

Our study demonstrates that discontinuing warfarin therapy within the first 3 months after surgery is associated with a significant increase in the risk of stroke, thromboembolic complications, and cardiovascular death. The novelty of our study is the finding that discontinuing warfarin therapy within 90 to 179 days after surgery is associated with a significant increase in the risk of cardiovascular death.

International guidelines on anticoagulation after a bioprosthetic value AVR have been written with limited data on the appropriate duration of warfarin treatment after surgery. Consequently, our study challenges current guidelines on the duration of antithrombotic treatment after AVR surgery with biological valves by presenting results suggesting that these patients will gain from an additional 3 months of warfarin treatment in terms of reduced cardiovascular death without risking a significant increase in bleeding events.

## Author Information

Corresponding Author: Charlotte Mérie, MD, Department of Cardiology, Copenhagen University Hospital Gentofte, Niels Andersens Vej 65, 2900 Hellerup, Denmark (c.merie@mail.dk).

Author Contributions: Dr Mérie had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study concept and design: Mérie, Torp-Pedersen.

Acquisition of data: Mérie, Skov Olsen, Torp-Pedersen.

Analysis and interpretation of data: Mérie, Køber, Skov Olsen, Andersson, Gislason, Skov Jensen, Torp-Pedersen.

Drafting of the manuscript: Mérie.

Critical revision of the manuscript for important intellectual content: Mérie, Køber, Skov Olsen, Andersson, Gislason, Skov Jensen, Torp-Pedersen.

Statistical analysis: Mérie, Køber, Gislason, Torp-Pedersen.

Obtained funding: Torp-Pedersen.

Administrative, technical, or material support: Mérie, Skov Olsen, Andersson, Skov Jensen.

Study supervision: Køber, Skov Olsen, Gislason, Skov Jensen, Torp-Pedersen.

Conflict of Interest Disclosures: All authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Dr Køber reported receiving payment for giving lectures on unrelated issues from Servier as well as unrelated consultancy for Bayer. Dr Andersson reported receiving an independent research grant FSS-11-120873 unrelated to current study with 30 months salary and meeting expenses from the Danish Agency for Science, Technology, and Innovation. Dr Torp-Pedersen reported receiving consultancy fees for antiarrhythmic drugs from Cardiome, sanofi-aventis, Merck, and Bristol-Meyers Squibb. No other authors reported any financial disclosures.

Funding/Support: This work was supported by the Research Fund of the Department of Cardiology at Copenhagen University Hospital Gentofte, Gentofte, Denmark.

Role of the Sponsor: The Research Fund of the Department of Cardiology at Copenhagen University Hospital Gentofte, Denmark, had no role in the design or conduct of the study; in the collection, analysis, and interpretation of the data; or in the preparation, review, or approval of the manuscript.

Disclaimer: The findings and conclusions in this study are those of the authors and not necessarily those of the Department of Cardiology, Copenhagen University Hospital Gentofte; Department of Cardiology, Copenhagen University Hospital Rigshospitalet; or Department of Cardiothoracic Surgery, Copenhagen University Hospital Rigshospitalet, Copenhagen, Denmark.

BACK TO TOP

## Global burden of cancer in 2008: a systematic analysis of disability-adjusted life-years in 12 world regions

Lancet
Dr Isabelle Soerjomataram MD a , Joannie Lortet-Tieulent MSc a, D Maxwell Parkin PhD b, Jacques Ferlay ME a, Colin Mathers PhD c, David Forman PhD a, Freddie Bray PhD
24 Nov 2012

Summary

Background

Country comparisons that consider the effect of fatal and non-fatal disease outcomes are needed for health-care planning. We calculated disability-adjusted life-years (DALYs) to estimate the global burden of cancer in 2008.

Methods

We used population-based data, mostly from cancer registries, for incidence, mortality, life expectancy, disease duration, and age at onset and death, alongside proportions of patients who were treated and living with sequelae or regarded as cured, to calculate years of life lost (YLLs) and years lived with disability (YLDs). We used YLLs and YLDs to derive DALYs for 27 sites of cancers in 184 countries in 12 world regions. Estimates were grouped into four categories based on a country's human development index

(HDI). We applied zero discounting and uniform age weighting, and age-standardised rates to enable cross-country and regional comparisons.

Findings

Worldwide, an estimated 169·3 million years of healthy life were lost because of cancer in 2008. Colorectal, lung, breast, and prostate cancers were the main contributors to total DALYs in most world regions and caused 18—50% of the total cancer burden. We estimated an additional burden of 25% from infection-related cancers (liver, stomach, and cervical) in sub-Saharan Africa, and 27% in eastern Asia. We noted substantial global differences in the cancer profile of DALYs by country and region; however, YLLs were the most important component of DALYs in all countries and for all cancers, and contributed to more than 90% of the total burden. Nonetheless, low-resource settings had consistently higher YLLs (as a proportion of total DALYs) than did high-resource settings.

Interpretation

Age-adjusted DALYs lost from cancer are substantial, irrespective of world region. The consistently larger proportions of YLLs in low HDI than in high HDI countries indicate substantial inequalities in prognosis after diagnosis, related to degree of human development. Therefore, radical improvement in cancer care is needed in low-resource countries.

Funding

Dutch Scientific Society, Erasmus University Rotterdam, and International Agency for research on Cancer.

Introduction

Cancer is a major cause of mortality worldwide, contributing to 7·6 million deaths in 2008.1 In the past few decades, combined successes of cancer prevention, early detection, screening, and treatment have reduced overall mortality in some developed countries because of declines in incidence or mortality from specific cancers, including lung, cervical, breast, and stomach cancer, and leukaemia.2, 3 With increases in cancer survivors in medium-resource to high-resource settings, interest has grown in the improvement of quality of life via reduction of cancer-related sequelae that lead to disability.4—7 By contrast, cancer mortality in many low-income and middle-income countries continues to rise,3 whereas the extent to which survival has improved is variable.8, 9 In these settings, globalisation has doubled the burden of cancer with a residuum of infection-related cancers (particularly in sub-

Saharan Africa) and a rise in incidence of cancers associated with progressively westernised lifestyles (eg, breast, colorectal, and prostate cancer) as levels of human development improve.10 Furthermore, increases in life expectancy have contributed to the rise in the global burden of cancer. Future changes in incidence aside, projections that consider population growth and ageing suggest that new cancer cases worldwide will increase from 12·7 million in 2008 to 21·4 million in 2030. This increase will be more substantial in low-resource and medium-resource countries than in high-resource countries (76% vs 25%).11, 12

Information about fatal and non-fatal cancer-related outcomes is needed to establish priorities in cancer control. Disability-adjusted life-years (DALYs) are a key measure for such purposes because they link the burden of cancer mortality with the degree of illness and disability in patients and long-term survivors.13

We present DALYs and their two components—years of life lost because of premature mortality (YLLs) and years lived with disability (YLDs)—for 27 cancers sites, separately and combined, for 184 countries and 12 world regions in 2008. We took account of global indicators of human development.

## Methods

### Data collection

A detailed description of the data sources and methods of estimation used to obtain the measures for calculation of DALYs have been described previously.14 We used the following country-specific and cancer-specific estimates to compute DALYs: population data (UN Population Division15); incidence and mortality (GLOBOCAN 200812); estimates of the proportion of cured and treated individuals14 (based on incidence to mortality ratios, survival estimates, and treatment data from cancer registries); time intervals of distinct disease phases, including duration of diagnosis and treatment, and time to cure and death (estimated from models with data from cancer registries16, 17 and treatment guidelines); standard life expectancy tables (based on the Princeton modified life table18); and disability weights (Dutch and Victorian [Australia] burden of disease studies19, 20).

### Statistical analyses

DALYs are the sum of YLLs and YLDs.13 We calculated YLLs by multiplying the number of cancer-specific deaths in a given age group by the remaining life expectancy of a standard population for that group. We calculated YLDs by multiplying the number of incident cases at each non-fatal disease phase by the mean duration of time associated with each phase. We then multiplied disability weights by these life-years to account for the severity of each event. Finally, we converted YLLs, YLDs, and DALYs to

country-specific rates (per 100 000 population) by dividing the healthy life-years lost by corresponding population estimates. To enable cross-country and regional comparisons, we age-standardised rates with the world standard population.21

We estimated DALYs for each cancer site by sex, age group (0—14, 15—39, 40—44, 45—49, 50—54, 55—59, 60—64, 65—69, 70—74, and ≥75 years), and country. We included 27 sites of cancer: lip and oral cavity (International Classification of Diseases tenth revision [ICD-10], C00—08), nasopharyx (C11), other pharynx (C09-10, C12—14), oesophagus (C15), stomach (C16), colorectum (C18—21), liver (C22), gallbladder (C23-24), pancreas (C25), larynx (C32), trachea, bronchus, and lung (C33—34), melanoma (C43), Kaposi's sarcoma (C46, estimated in GLOBOCAN 2008 for only sub-Saharan Africa), breast (C50), cervix uteri (C53), corpus uteri (C54), ovary (C56), prostate (C61), testis (C62), kidney (C64—66), bladder (C67), brain and nervous system (C70—72), thyroid (C73), Hodgkin lymphoma (C81), non-Hodgkin lymphoma (C82—85, C96), multiple myeloma (C88, C90), and leukaemia (C91—95). To calculate DALYs for all cancer sites (with exclusion of non-melanomas), we first estimated YLLs and YLDs for the group that included all other cancer sites not listed above (appendix); we then added these estimates to the DALYs for the 27 cancer sites being studied.

We grouped estimates for 184 countries into four categories according to a country's human development index (HDI).22 The HDI is a composite index of three dimensions of human development: a long and healthy life (based on life expectancy at birth), access to knowledge (based on a combination of rates of adult literacy and enrolment in primary to tertiary education), and an adequate standard of living (based on gross domestic product per capita adjusted for purchasing power parity [US$]). We grouped countries into four categories of HDI according to estimates from the 2007 UN Development Programme: very high HDI regions (HDI ≥0·9), high HDI regions (between 0·7 and 0·9), medium HDI (between 0·5 and 0·7), and low HDI (<0·5). Furthermore, we grouped countries into 12 geographical regions: sub-Saharan Africa, including the UN geographical regions15 of eastern, middle, southern, and western Africa; Middle East and north Africa (western Asia and northern Africa); Latin America and the Caribbean (central and southern America and Caribbean); North America; East Asia (eastern Asia, excluding China); southeast Asia; south-central Asia (southern Asia, excluding India); eastern Europe; northern Europe; southern Europe; western Europe; and Oceania (excluding Australia and New Zealand). When reporting the estimates, we tabulated China and India, and Australia and New Zealand, separately.

To validate our results we did various sensitivity analyses, including a comparison of observed versus estimated parameters as inputs in the calculation of DALYs, simplifying the natural history of disease model (two-stage versus three-stage models), and exchanging input parameters—eg, proportion of patients with various disease stages as proxies for the proportion treated and cured. Changes to the parameters had a fairly moderate effect on the DALY rates for most cancer sites, with differences of 0—10%.14 We did analysis with STATA (version 11).

Role of the funding source

The sponsor of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Results

An estimated 169·3 million healthy life-years were lost because of cancer in 2008. In absolute terms, Asia and Europe contributed to 73% of the overall burden of DALYs lost because of cancer (appendix p 3). DALYs in China accounted for 25% of the overall burden and those in India for 11%; these two regions accounted for 67% of the burden in Asia. Lung, liver, breast, stomach, colorectal, cervical, and oesophageal cancers, and leukaemia had the highest proportion of DALYs with a combined contribution of 65% to the total cancer burden (appendix p 3).

The table shows the estimated global age-adjusted DALYs lost from cancer in 2008, after adjustment for population size and age. Men had 6% more DALYs than did women (table). Men in eastern Europe had the largest cancer burden, which was driven by the very high DALY rates for several common causes of cancer incidence and death in this region, including colorectal, lung, pancreatic, and kidney cancers (table). For women, DALY totals were highest in Oceania and sub-Saharan Africa (table). Breast cancer was a major contributor to total DALYs in all regions except eastern Asia, and lung cancer in all except sub-Saharan Africa, where 36% of the total DALYs were attributable to Kaposi's sarcoma, liver and cervical cancer, and non-Hodgkin lymphoma (figure 1).

For human development, we noted a fairly small difference in total DALYs lost between the four HDI subgroups (table). Furthermore, the contribution of YLLs and YLDs to DALY totals for each cancer site differed substantially by HDI (figure 2). We noted generally higher YLD totals in very high HDI countries and lower YLD totals in lower HDI countries (figure 2). For site-specific distribution by HDI level, lung, breast, and colorectal cancers were the main contributors to the high DALY totals in very high HDI countries (figure 2). By contrast, cancers of the cervix and liver, and Kaposi's sarcoma contributed the largest proportion of disease burden in low-HDI and medium-HDI countries (figure 2).

In addition to estimates of DALYs and YLLs and YLDs, we calculated YLLs as a proportion of total DALYs. Although the contribution of YLLs to total DALYs was generally low in very high HDI countries and high in low HDI countries, the extent of this relation varied by site of cancer (appendix pp 7—8). The YLL to DALY proportion was largest in sub-Saharan Africa for all cancer sites combined (96%) and was generally lowest in North America (84%; appendix). We noted larger variations in this proportion for cancers of the

testis (38% in very high HDI vs 96% in low HDI countries) and prostate (52% vs 96%) in men, and for breast cancer in women (63% vs 90%).

Figure 3 shows world maps for age-adjusted DALYs, YLLs, and YLDs per 100 000 population in men, and the appendix (p 4) for women. The overall YLL profile was very similar to that for DALYs because YLLs account for 93% of the estimated DALYs (appendix pp 7—8). Worldwide, men in Hungary, Mongolia, Armenia, Belarus, and Uruguay had the highest DALY proportions (figure 3), whereas for women, proportions were highest in sub-Saharan Africa, with Malawi, Uganda, Zimbabwe, Mali, and Zambia ranking as the top five countries globally (appendix p 4). Cross-country variation in YLD rates was much larger than that noted for overall DALYs or YLLs, with variations of 14 times in men and ten times in women (figure 3, appendix p 4).

Worldwide, lung, liver, and stomach cancers were the main contributors to the DALY burden from cancer in most countries in men, ranking first in 87, 37, and 17 countries, respectively (appendix p 5), and accounting for 23%, 28%, and 20%, respectively, of the total DALYs worldwide. A similar picture was noted for YLL rates with a slightly increased number of countries having lung cancer as the main cause of premature mortality (93 countries). The main cancers contributing to YLDs were somewhat different, with prostate cancer as the main contributor in 93 countries and colorectal cancer in 37 (appendix p 5). Prostate cancer was the main cause of YLDs in the Americas, most European countries, Australia and New Zealand, and much of sub-Saharan Africa, whereas colorectal cancer was the main cause in eastern Europe (including Russia), and southeast and eastern Asia (including Japan).

Breast and cervical cancer were key components of the overall DALY rates in women (figure 4), with breast cancer the leading cause of YLDs in 119 of the 184 countries (figure 4), compared with 49 countries for cervical cancer. However, cervical cancer contributed more to premature death than did breast cancer in 23 countries, mostly in sub-Saharan Africa and central and South America (figure 4).

## Discussion

Worldwide, an estimated 169·3 million healthy life-years were lost because of cancer in 2008, with an estimated variation of five times in rates of DALYS by country. Asia and Europe were the main contributors to the overall burden of DALYs lost because of cancer, and breast, prostate, colorectal, and lung cancers made the largest combined contribution to total DALYs. Of YLLs and YLDs, YLLs were the most important component of DALYs in all world regions; however, the relative contribution to the total from all cancers varied by region, and more so by cancer site. Finally, our study confirmed the continuing epidemiological transition in many developing countries, for which can be noted a double burden of cancers (mainly breast, colorectal, and prostate cancer) associated with increased westernisation, and a reduced, but still important, burden from infection-related cancers (including cervical, liver, and

stomach). This transition has led to a higher average burden of cancer in low-resource and medium-resource regions than in high-resource regions, which is shown in the fairly small differences in the final age-adjusted totals of DALYs, as categorised by level of human development.

Global DALY estimates for all disease outcomes have previously been published.13—28 The results of the continuing global burden of disease 2010 study29 might differ from our results because of changes in methods of DALY calculation or use of different input sources in their estimation. To assess long-term changes, DALY rates (as reported by the global burden of disease assessment in 199013) can be compared with our results for various cancers by region. The size of the estimates for DALYs (or YLLs and YLDs) might differ because of differences in analytical methods13, 14 (eg, in natural disease history, estimates of proportions of patients who were cured or treated, and inclusion of discounting and age weighting). However, changes in the relative contribution of each cancer site to the total burden in time can explain the position of a region or a country in relation to the extent of its epidemiological transition.30 The estimates from 1990, for the proportion of burden attributed to colorectal, breast, and prostate cancer in established market economy countries resemble our results for very high HDI countries (24% in 1990 vs 26% in 2008). The contribution of lung cancer to the total DALYs from cancer is likewise similar (5% vs 7%), as are the proportions of cervical, liver, and stomach cancers in sub-Saharan Africa that contribute to the overall DALYs from cancer (27% vs 29%). However, the proportions of colorectal, prostate, and breast cancers have substantially increased from 11% of the total global burden in 1990, to 20% in 2008. Although in China the burden of cancers related to lifestyle and infection was fairly stable, we noted a substantial increase in the relative burden of lung cancer (12% in 1990 vs 20% in 2008). Our results thus confirm those of previous studies, which show that the epidemiological transition has progressed in many developing countries in the past few decades.10, 11 Because premature death from infectious diseases is declining in many countries, life expectancy has increased and the rise in cancer burden is mainly lifestyle related.

Our finding that YLLs contributed more than 90% of the cancer burden, with the highest levels in low-resource countries, confirms previous reports.13, 28, 31, 32 In Australia, where similar analytical methods were used, YLLs contributed to 78% of the overall DALYs related to cancer in 2001.19 Declines in cancer mortality have been reported in several developed countries and regions, including northern America, Australia, and northern Europe,3, 33 because of a reduction in lung and other smoking-related cancers in men, and breast cancer in women. However, in most countries, cancer mortality is still rising,3 and we would thus expect a large sustained contribution of cancer to overall premature death in many low-income to middle-income populations. Improved access to high-quality treatment has not greatly improved survival for cancers associated with poor prognosis and subsequently large DALY totals (eg, lung, stomach, liver, pancreatic, and liver cancer); this finding emphasises the crucial role of primary prevention to reduce much of the disease burden.

When categorised by levels of human development, variations in DALYs of cancer were fairly small, which shows the importance of cancer as a major contributor to overall disease burden, irrespective of level of resource. However, the fraction of DALYs due to loss of life was consistently larger in the lowest (mainly sub-Saharan Africa) than the highest HDI settings, which signifies the poorer prognosis of patients with cancer in low-resource settings.25 This finding confirms previous reports of large inequalities in worldwide cancer survival—eg, 84% survival from breast cancer in the USA26 compared with 30% in Bhopal, India.34 These differences persist according to cancer stage; for example, patients with localised breast cancer have 17% higher survival when diagnosed in more-developed than in less-developed countries.25 This variation implies that improved survival requires methods for cancer control that incorporate early detection programmes as part of a health infrastructure in which adequate and accessible services are available to all.34, 35

The range of YLL to DALY proportions by country in our study further shows the level of avoidable mortality from cancer after initial diagnosis in different settings. We noted a broad range in this proportion for at least half the 27 cancer sites, including colorectal, thyroid, bladder, and breast cancer, Hodgkin's lymphoma, and cutaneous melanoma. The wide range of YLL to DALY for breast and prostate cancer partly shows cross-country variations in the prevalence of mammographic screening and prostate specific antigen testing. The possibility of lead-time bias and overdiagnosis because of these tests36, 37 should be considered in the interpretation of these results, as should the potential benefits of preventive strategies. The particularly wide range of YLL to DALY proportion noted for testicular cancer might be interpreted in view of prospects to reduce the burden from these cancers by improvement of access to effective cancer diagnosis and treatment.38

Our analysis is comprehensive in its global coverage; it is the first attempt to estimate the global DALYs related to cancer with a unified framework that accounts for the natural history of each of the 27 cancer sites studied (panel). We used substantial input from population-based data, mostly from cancer registries, to increase the validity of the final estimates. Zero discounting of the future years lost gives more weight to years lost presently, whereas non-uniform age weights give more weight to deaths in young and middle-aged individuals. The discounting and age weighting used in our study produced fairly low DALYs; however, these methods did not substantially alter the ranking of cancers by DALYs.14

Panel

Research in context

Systematic review

The International Agency for Research on Cancer has synthesised global estimates of the incidence and mortality for 27 cancers.23 To estimate the proportion of individuals cured and treated, we linked country-specific modelled survival proportions to the human development index,21 and used data from existing regional or global population-based survival studies24—26,27 to model and validate the derived survival proportions. We assessed published work16—20 of duration of disease phases, treatment guidelines, quality of life of patients with sequelae, and disability weightings for each of the 27 cancer types. Relevant information was derived from reports and articles mainly obtained from a series of Medline searches using MeSH terms of above mentioned topics published before Jan 1, 2011. No language restriction was employed These data provided the necessary input for calculation of DALYs. Finally, we searched and reviewed key global cancer and other disease burden papers13,19—20 to compare estimates of DALYs, and to inform the discussion of their interpretation in global health planning.

Interpretation

This study is the first to globally assess DALYs of cancer and their two components, YLLs and YLDs. Our findings emphasise the need for increased efforts in cancer control in low-resource settings, where premature mortality from cancer, as measured by YLD, is highest. The inadequate prevention, early detection, and treatment programmes in low-income to middle-income countries should be reassessed in view of the predicted long-term increases in the future burden of cancer.

Our calculations were based on restricted country-level data. Cause of death assignment in less-developed countries is often incomplete and many deaths are recorded as having unknown or ill-defined causes.39 In such cases, we used methods of reassignment to estimate cause-specific death rates, obtaining information from available data in a standard manner to enable cross-country comparison.39 Estimates of incident cases when data are scarce or unavailable are inherently uncertain; furthermore, under-reporting could be a source of bias in cancer registration in developing countries, and national extrapolation of regional data derived from urbanised settings might result in overestimation of the final estimate.1 Despite the various concerns about data quality and methods of estimation, we believe that the estimates in GLOBOCAN 2008 are the most accurate available.

We noted a larger gap in terms of the measures needed to calculate YLDs than in those for calculation of YLLs. These measures included estimates of the proportions cured, treated, and with sequelae, which are based on several data sources for which the quality and extent of information varies greatly. Modelling of these measures to produce country-specific estimates results in substantial uncertainty, especially in less-developed countries where data are scarce. We modelled the proportion cured using the mortality to incidence ratio, which provides a reasonable approximation of 1 minus survival.40—42 Potential codisability is a possible

source of uncertainty that is specific to YLD estimates.13, 19 For some cancers, such as prostate cancer, survivors might have several sequelae, such as incontinence, impotence, or infertility.19 We assumed that the severity weight for the combination of two or three of the disabilities was the aggregate of each of the disability weights. For example, a randomised trial of prostatectomy versus watchful waiting5 reported similar measures of quality of life in men with early prostate cancer, who had been randomly assigned to prostatectomy and had treatment-related side-effects, and those in the watchful-waiting group. Such biases might lead to overestimated YLDs. However, we made a great effort to select the best available data to estimate DALYs and their components, and thus provide a method to compare the cancer burden between countries on the basis of the effects of early mortality, morbidity, and disability. Detailed studies are needed to assess progress in national cancer control, whereas studies of cancer outcomes are needed to improve the estimation of DALYs.

In the past few decades, national policies and interventions have been implemented to reduce cancer incidence and mortality, and to prevent disability in high-resource settings. Our data support the need for better attention to cancer prevention and treatment programmes in low-resource countries, for which our findings have shown the extent of the cancer burden in terms of YLLs, and have confirmed the poor prognosis after cancer diagnosis.25 Programmes for cancer care that are tailored to the needs of individual low-income countries need to be developed and tested.43—45 Inadequate prevention, early detection, and treatment programmes need reassessment in view of long-term increases in the predicted future burden of cancer in these areas.

## Contributors

IS contributed to data collection, study design, analysis, and wrote the first draft of the paper. JL-T contributed to study design, analysis, and finalising the report. JF contributed to data collection and finalising the report. DF contributed to finalising the report. DMP and CM contributed to the study design and finalising the report. FB contributed to study design, analysis, and drafting of the report. All other authors read and approved the final report.

## Conflicts of interest

DMP, JF, CM, and FB were members of an expert group that has contributed to the global burden of disease 2010 study, but the results reported here are independent of the global burden of disease study. All other authors declare that they have no conflicts of interest.

## Acknowledgments

We thank Rubicon from the Dutch Scientific Society and the EUR Fellowship, Eric Chokunonga (Zimbabwe National Cancer registry), Nadia Dimitrova (Bulgarian National Cancer Registry), Fiona Dwane (National Cancer Registry of Ireland), Ryszard Mezyk (Kielce Regional Cancer Registry), and Omar Nimri (Jordan Cancer Registry). This study would not have been possible without the availability of data provided from worldwide population-based cancer registries, as used in the provision of estimates for GLOBOCAN 2008.

BACK TO TOP