Effect of subjective sleepiness on critical skills performance in National Guard medical personnel during extensive disaster training

Denise SMART¹, Lois JAMES¹, Tamara ODOM-MARYON¹ and Stephanie ROWAN²

¹Washington State University–College of Nursing, Spokane, WA, USA
²University of Texas Health Science Center at San Antonio, San Antonio, Texas, USA

Abstract

Aim: This project investigates the effect of subjective sleepiness on critical skills performance of National Guard (NG) medical personnel during highly realistic and immersive disaster-training exercises.

Methods: This study used a longitudinal, observational design in a sample (n=77) of NG medical personnel over three high-intensity disaster-training exercises. During three separate 3- to 5-day disaster response training exercises, participants completed a critical skills test four times per day. At the same time, their subjective sleepiness was measured using the Karolinska Sleepiness Scale (KSS). The effect of subjective sleepiness on: (1) reaction time to complete critical skills questions; and (2) accuracy in completing critical skills questions were analyzed using generalized linear mixed models.

Results: With changing disaster response duties and assignments, 91% of participants responded to 2,228 critical skills questions over two to three different disaster exercises. When levels of sleepiness (KSS self-reported scores) were rated as ‘Extremely alert’, response accuracy averaged 91%; however, when self-reported sleepiness was scored as “Extremely sleepy”, accuracy dropped to 60%.

Conclusions: Increasing sleepiness correlated significantly with decreasing accuracy on critical skills questions. Participant professional licensure was also significantly correlated with accuracy of responses to critical skills questions. Military deployments and disaster-training exercises designed to replicate real-world events require prolonged periods of sleep insufficiency or sleep interruptions. Adjustments to shift work for nurses, attention to pre-deployment screening and education, and a change in cultural attitudes may be needed to make a positive difference in morale, mission completion, and casualty reductions.

Key words: disaster response, National Guard, operational performance, sleep patterns

INTRODUCTION

The National Guard (NG) plays a vitally important role in the United States military strategic plan. The National Guard is represented by both Army and Air National Guard personnel who are trained within specialty career fields for both active duty deployments and state mission-specific needs. As of 2018–19, there are over 340,000 Army National Guard and 107,700 Air National Guard representing all 50 states (US Department of Defense, 2018). Approximately 11% (40,194) of the National Guard are Active Guard/Reserve, which means they serve in a full-time capacity versus the one-weekend-a-month, traditional National Guard role (Military Leadership & Diversity Commission, 2010).

Traditional NG personnel, which includes registered nurses and other medical professionals, are unique among the armed forces in that members typically hold full-time civilian jobs and must transition between civilian and military duties for scheduled training days, state
responses to emergencies or disasters, and active duty service (Thomas et al., 2010). The range of duties required of National Guard personnel include responding to natural disasters (e.g., hurricanes, mudslides, earthquakes, disease outbreaks), as well as homeland security threats. They are the primary military branch for responding to any attack — natural or human-caused — on home soil, and as such their safety, health, and wellness are vital for the country’s interests (Mahon, 1983; Thomas et al., 2010).

While responding to state emergencies or disasters (natural or man-made), National Guard personnel can be exposed to conditions such as extreme heat or cold, mass casualties, and acute sleepiness and fatigue brought about by extended sleep restriction (Kline et al., 2010). National Guard medical personnel have a unique role during disasters, which has large societal relevance. When fully mobilized as a medical team, they are trained to triage civilian and military patients in the field or disaster region and administer medical care until patients are transported to local facilities. Decision-making becomes challenged when confronting extreme environmental conditions, 24/7 operations and sleep fragmentation or fatigue. The effects of these conditions on National Guard personnel are not well understood, as they have traditionally been an understudied group among the military population.

Ample evidence exists that acute reductions in sleep time result in negative health and safety consequences (Grandner et al., 2010; Jackson et al., 2013). Health consequences include molecular, immune, and neural changes leading to increased risk of disease and serious long-term health effects (Luyster et al., 2012). Safety consequences from fatigue include increased risk for errors, accidents, injuries, and reduced job performance (James et al., 2017; Trinkoff et al., 2011). One of the most salient examples of fatigue-related errors is drowsy driving. The U.S. National Transportation Safety Board determines that approximately 40% of all truck driver accidents are caused by fatigue (Caruso, 2014). Research from the American Automobile Association (AAA) has found that up to 20% of drivers have fallen asleep at the wheel (AAA Foundation for Traffic Safety, 2015).

Given the known risks of inadequate sleep, researchers have investigated risks for specific populations. One industry of interest has been health care, where shift work is common. The risks associated with shift work are two-fold: (1) long work hours result in fatigue and sleepiness (Smith-Miller et al., 2014; Caruso, 2014; Caruso & Waters, 2008; James & Vila, 2015); and (2) working during the evening or night results in circadian misalign-
exercises using actigraphy data (Smart et al., 2018).

METHODS

Study design

This study used a longitudinal, observational design to investigate the effect of subjective sleepiness on critical skills performance in a sample of National Guard medical personnel as they responded in up to three high-intensity disaster-training exercises. These disaster response training exercises are designed to simulate real-world disaster conditions, and as such, sleep is restricted (James et al., 2019). These exercises provided the ideal opportunity for monitoring the effect of sleepiness on critical skills within an accessible field setting. The exercises selected for inclusion in the study were similar in number of days and mission between the two Air National Guard (ANG) units located on US Air Force Bases (AFBs). Two exercises completed at both AFBs were 3 and 5 days in length. An additional 4-day exercise was completed at only one of the bases. University Institutional Review Board and Department of Defense (Air Force Human Research Protection Office) reviewed this 2-year study prior to any participant recruitment.

Participants

A total of 77 (out of a possible 90, if all positions were filled and all service members were present during recruitment phases) ANG participants were recruited from two separate AFBs, one in the Northwest (n=40) and the other in the South (n=37), during the monthly “Drill” weekend routinely held the month prior to each AFB’s disaster response training exercise. Participants were monitored prior to three different disaster response training exercises at the Southern AFB and prior to two different disaster-training exercises at the Northwestern AFB.

Materials

Demographics, military history and fatigue (level of sleepiness) collected for this study were selected from the RAND Corporation 2015 study survey instrument (Troxel et al., 2015). The original RAND 2015 survey instrument is a compilation of the Pittsburgh Sleep Quality Index (PSQI – 19 items), PSQI (Bed Partner/ Roommate items – 5), Insomnia Severity Index (7 items) sleep duration question, sleep disturbance questions, Karolinska Sleepiness Scale (KSS), Berlin Questionnaire (modified from 10 items to 1 item to assess the fatigue domain), Sleep Hygiene Index (13 items), Fred Hutchinson Cancer Research Center Caffeine Questionnaire, Post Traumatic Stress Disorder (PTSD) Checklist (17 items), researcher-generated shift-work items and RAND Deployment Life study questions.

Demographic and military history

Demographic data collected at the time of enrollment included sex, age, ethnicity, marital status, number of children living at home, professional licensure, employment history, and civilian work patterns. Self-reported data about military history included rank, deployment history, and combat exposure.

Karolinska Sleepiness Scale

The Karolinska Sleepiness Scale (KSS) is a validated instrument (Kaida et al., 2006) that includes a single item about “How sleepy do you feel right now?”. Participants were asked to rate the level of sleepiness using the KSS (range: 1–9, where 1 = Extremely alert, 3 = Alert, 5 = Neither, 7 = Sleepy, and 9 = Extremely sleepy).

Critical skills

Participants’ ability to respond accurately and in a timely manner to a critical skills question were used as measures of operational performance. Licensed professionals answered skills questions involving medication calculations that could be mentally calculated (not requiring a calculator) and non-licensed service members received skills questions about basic life support (BLS). For this study, military participants’ main priority was mission-focused and task specific. As such, ability to respond to medically related critical skills questions was chosen for the measure of operational performance to minimize disruption to participants’ disaster-training exercises. Commonly used measures such as the Performance Vigilance Task that requires 3–10 min for evaluation (Basner & Dinges, 2012) were unacceptable for use with this study because they would have placed undue burden on participants and disrupted military training missions.

Prior to starting this study, a pilot study was conducted to identify a set of critical skills questions (one set of medication calculation questions and one set of BLS questions) requiring similar average response times and with similar levels of accuracy (i.e., percentage of participants correctly answering the question; Smart et al., 2018). Briefly, skills questions with response times as close to 30 s (±10 s) and accuracy greater than 80% of the questions were selected for use in this study.

Procedures

To assess military participants’ level of sleepiness while participating in the disaster-training exercises, partici-
participants completed the KSS at approximately 08:00, 12:00, 16:00 and 20:00 h each day. At the same time, participants responded to either a medication calculation skills question or a BLS skills question depending on their job role (e.g., licensed medical personnel responded to medication calculation questions, and non-licensed medical personnel responded to BLS questions). KSS and critical skills data were measured at 48 data collection timepoints (four times per day at three disaster response training exercises lasting 3, 4 or 5 days). The critical skills questions were randomly assigned to the 48 different measurement times. In this manner, a unique medication calculation and BLS skills question were assigned to each measurement time avoiding recall bias by participants from having answered the question previously.

Questions were sent to each participant’s cell phone as a text message that included a link to the skills question using Qualtrics Software, Version 2017 of Qualtrics© 2005 (Qualtrics, Provo, UT, USA) a secure online research site for survey development and distribution (https://www.qualtrics.com). Some participants preferred to receive questions via an email containing the link. Prior to starting each training exercise, a test question was sent to each participant to ensure they could access questions during data collection periods. Participants had up to 2 h to respond to the skills questions and rate their level of sleepiness. After 2 h, the link sent via text or email was deactivated. Participants’ availability to answer questions throughout the day depended on access to their cell phone, which may have been restricted based on their job duties.

Study outcomes
Two outcomes were examined: accuracy (percentage of participants answering the question correctly) and response time (s). Response time was calculated as the time required to read the question and multiple choice set of answers, and then submit the response to the question. When examining response time required to perform a task, observations with incorrect responses are excluded from the analyses to avoid bias due to participants immediately guessing (Sternberg, 1969). To standardize responses times among the 48 unique skills questions, participants’ response times were converted to z-scores using the results from the pilot study (by subtracting the corresponding average response time and dividing by the standard deviation for the skills question). We refer to these z-score response times as “normed”.

Statistical analysis
To characterize the participant sample, demographic and military history participant data collected using the RAND demographic tool were summarized as means (standard deviations [SD]) for continuous variables and frequencies (percentages) for categorical variables. Comparisons of participant demographics between the two AFBs, demonstrating that the participant samples from each base were similar with respect to demographics and military history, were reported elsewhere (James et al., 2019; Smart et al., 2018). To characterize the total number of skills questions completed by participants across the disaster-training exercises, number of skills questions available for analysis was summarized as medians and interquartile ranges (IQRs) and separately for the two critical skills outcome variables (accuracy and normed response time).

A generalized linear mixed model (GLMM) was used to examine the association between sleep deprivation and fatigue on operational performance (critical skills). The SAS software, GLIMMIX Procedure (version 9.4; Cary, NC, USA), was used. The GLIMMIX Procedure fits generalized linear mixed models to data with correlations or non-constant variability and where the response is not necessarily normally distributed. Accuracy (1=correct, 0=incorrect) and normed response time (in seconds) were specified as the dependent variable using the logit or identity link function (respectively). For each outcome, unadjusted models were tested including each of the following predictor variables (specified as fixed effects) separately (i.e., univariable models): sex (1=female, 0=male), age (years), group (1=no professional licensure, 0=professional licensure), ethnicity (1=Hispanic, 0=non-Hispanic) and base (1=Northern, 0=Southern) and KSS (scale from 1 to 9). Terms representing the interaction between KSS and demographic variables were also explored. To examine the statistical significance of KSS adjusted for the other predictor variables, a multiple variable model including all predictor variables with a significance level of 0.05 or less was run for both outcome variables separately. Model fit criteria (Bayesian Information Criterion) were used to assess the best fitting models, including the significance of the interaction terms.

Correlations within an individual over repeated measures of time (exercise, day, time within day) and among individuals within each of the military bases were controlled using a random intercept model and an unstructured covariance matrix. The need for a random intercept and slope model was also examined; however, the slope component was uninformative. Results from the GLMM analyses reported include the unadjusted regression coefficients and corresponding standard error (β values).
(SE), F-statistic (df1, df2) and the P-value along with the odds ratio (OR) and corresponding 95% confidence interval (CI=[lower bound, upper bound]) for each predictor variable. For the outcome accuracy, adjusted odds ratios from a multiple variable model are included. All significance testing was done at the 0.05 level (two-sided).

The sample size for this study was based on assuming enrollment of a minimum of 50 participants who would complete, on average, a total of 10 critical skills assessments providing 550 observations with paired measurements of fatigue and critical skills assessments for the final analysis. A minimum sample size of 50 was based on detecting differences in each outcome (percent accuracy and normed response time) between participants reporting a level of sleepiness of 1 compared to a level of sleepiness of 9 (Dupont & Plummer, 1998). For accuracy, differences in accuracy of 80% versus 50% could be detected with 95% power and $\alpha=0.05$. For normed response time, differences in times of 1.0 versus 1.5 s could be detected with 95% power and $\alpha=0.05$.

Among the 77 National Guard participants who enrolled in the study, the average age of participants was 36.0 years ($SD=9.1$), 60% were male and the majority (87%) identified their race as White or Caucasian (Table 1). Slightly less than one-third (29%) of participants were of Hispanic ethnicity. Approximately 65% of participants were either married or living with a partner. Two-thirds of the participants had no professional licensure (medics or Emergency Medical Technicians [EMT]). The remaining one-third with professional licensures included medical doctors ($n=2$), physician assistants ($n=6$), registered nurses ($n=15$) and pharmacists ($n=3$). Eighty-seven percent of the participants were employed full-time outside of the National Guard, with 26% ($n=26$) working variable shifts during their full-time employment. Just over one-third (36%) of the participants held a military rank of officer, and over half of the participants had been deployed three or more times, with 37% having combat exposure.

**RESULTS**

Job duties and assignments allowed 70 (91%) of the participants to complete some skills questions while participating in the disaster-training exercises. Among these 70 service members, 16 (23%) attended all three disaster-training exercises, 35 (50%) attended two disaster-training exercises and 19 (27%) attended only one of the disaster-training exercises. Overall, participants were able to respond to 39% ($n=858$) of the 2,228 skills questions that were sent out during the three disaster-training exercises (Table 2). The median number of skills questions answered by participants was 10, with 25% answering four or less and 25% answering 16 or more skills questions. The median number of skills questions answered correctly by participants was seven, with 25% answering three or less correctly and 25% answering 14 or more skills questions correctly.
Based on the univariable models examined, significant associations between accuracy and KSS, licensure, base, and age were found (Table 3). The unadjusted odds of answering a question correctly decreased with increasing levels of sleepiness (OR = 0.83, 95% CI = [0.77, 0.89]; p < 0.0001). The percentage of participants correctly answering skills questions dropped from 91% when describing their level of sleepiness as “Extremely alert” (KSS=1) to 60% when reporting their level of sleepiness as “Extremely sleepy” (KSS=9; Fig. 1). The unadjusted odds of answering a question correctly was 1.14-fold greater for every 5-year increase in age (95% CI = [1.02, 1.28], p = 0.023). The odds of answering a question correctly was lower for participants with no professional licensure (OR = 0.40, 95% CI = [0.27, 0.60]; p < 0.0001) and for participants from the Northwestern base (OR = 0.59, 95% CI = [0.38, 0.89]; p = 0.015).

Models testing the influence of the interaction between
KSS and each of the demographic variables on accuracy were non-significant. Using a multiple variable model, all significant predictors of accuracy (KSS, age, licensure and base) were tested in one model. After adjusting for KSS and licensure, military base affiliation and age were no longer associated with accuracy. Increasing level of sleepiness (adjusted OR=0.84, 95% CI=[0.78, 0.90]; \(p<0.0001\)) and no licensure (adjusted OR=0.42, 95% CI=[0.26, 0.66]; \(p=0.0006\)) were associated with decreased accuracy.

**Response time**

No associations between normed response time and KSS, licensure, base, sex, ethnicity, and age were found (Table 3). Average normed response times fluctuated from a low of 0.15 seconds (\(SD=1.49\)) when participants reported having felt “Extremely sleepy” (KSS=9) to a high of 1.62 seconds (\(SD=4.95\)) when they reported a KSS score of 4, corresponding to having felt somewhere between alert (KSS=3) to neither alert nor sleepy (KSS=5; Fig. 2). When participants responded as feeling “Extremely alert” (KSS=1), the average normed response time was 0.78 seconds (\(SD=2.71\)).

**DISCUSSION**

The US Army recognizes four domains related to operationally relevant capabilities that are most effected by poor sleep and increased fatigue: (1) acquiring, assigning priorities, allocating and using resources; (2) anticipating and solving problems; (3) managing and exploiting change; and (4) acting decisively under pressure (Headquarters, U.S. Department of the Army, 2009; Wesenten & Balkin, 2013). These four domains reflect the operationally relevant capabilities of National Guard medical personnel during both training exercises.
and real-world deployments. Consistent with these domains, this study found that increasing self-reported fatigue was associated with decreased critical performance measured by accuracy.

Additionally, this study found an association between lower accuracy and non-licensed participants. Nationally certified Emergency Medical Technicians (EMT-Basic) and other non-licensed participants may have full- or part-time non-medical civilian occupations. As such, their recall of BLS knowledge-type questions may have been weak. Additionally, non-licensed participants may have rapidly responded to BLS questions without correctly answering in an effort to resume their military-designated disaster-team roles.

Our study is unique in that we examined a specific skill set for licensed and non-licensed military medical personnel. When participants reported “Extremely sleepy” on the KSS (KSS=9), accuracy in answering critical skills questions dropped to 60 percent. This finding is similar to two studies: Scott et al. (2006) who noted the increase in self-reported medication errors when nurses worked longer shifts or extended hours where frequent episodes of drowsiness and sleepiness were reported, and more recent findings by Wolf et al. (2017), where a large sample (n=1506) of emergency department nurses were given medication calculation questions (timed cognitive performance test) and rated their sleepiness. Wolf et al. (2017) found a significant relationship between accuracy and amount of time to complete questions and increased self-reported sleepiness. Additionally, Olsen et al. (2013) reported a correlation between increased sleepiness and a decrease in the ability to anticipate tactical and moral problems (decision-making), which in turn negatively affects military operations and planning. Jackson et al. (2013) argues that cognitive performance encumbers a number of processes, such as working memory, motor actions, and as such, sleep deprivation effects are task specific. While deconstructing cognitive performance in order to examine the effects of sleep deprivation is important science (Jackson et al., 2013), day-to-day effects of sleepiness has real world implications including individual and patient safety, and overall mission capability. Our study adds to the body of research that uses skill-specific evaluation methods to replicate tasks used by health-care professionals and to studies on nurses who report fatigue or poor sleep quality (Kunert et al., 2007; Scott et al., 2006; Patterson et al., 2010; Smith-Miller et al., 2014; Wolf et al., 2017).

A growing concern is acknowledgement of sleep problems in the military in general and by National Guard personnel more specifically. Sleep problems may be underreported based on perception that a sleep impairment diagnosis may affect military career and continued service (Lincoln et al., 2018). As suggested by Lincoln, Moore and Ames (2018), National Guard service members may use unhealthy strategies to manage their sleep problems, which may contribute to other health and behavioral issues. Proper sleep health screening for insomnia and sleep apnea, in addition to sleep hygiene education, are all low-cost public health measures with strong implications to improving force readiness. A fatigue management reduction program recommended by the American College of Occupational and Environmental Medicine (ACOEM) (Lerman et al., 2012) with support for policies on sleep health could be used to mitigate the growing concern of sleep impairment.

LIMITATIONS

The study’s examination of military medical personnel participating in disaster response training exercises over an extended time period presented challenges. Participants’ main priority was mission-focused and task-specific. The ability to respond to skills questions sent via Qualtrics over a 12- to 14-h workday was limited by their ability to “pause” in their tasks allowing them to access their cell phone and was reliant on having cell service in their physical locations. Although it is possible that switching between disaster exercise tasks and the medication calculation or BLS tasks required for our study may have been taxing, this is not outside of the norm for combat medics or nurses, who must be prepared to switch tasks multiple times depending on mission requirements. Furthermore, both medication calculation and BLS are tasks that are directly relevant to combat medic or nurse duties.

A large component of disaster response training is donning full-body Tyvek suits with gloves, boots and face masks, and then performing simulated patient triage, search and extraction and patient-care functions. Thus, participants’ availability to respond to the critical skills questions may have been limited in this field—operational research study.

We did not collect a measure of physical activity to correlate with KSS and critical skills responses. The actigraphy devices worn by participants for the quantitative component of this study, cited elsewhere in this paper, did not include a physical activity tracking feature and it was not feasible for participants to wear two different devices (research-provided actigraphy devices and a personal smart device). Future studies could
include use of a smart device that tracks activity.

CONCLUSION

Military deployments and disaster-training exercises designed to replicate real-world events often require prolonged periods of sleep insufficiency or sleep interruptions. Operational demands can significantly affect safety of all concerned and compromise mission objectives. Military leadership has dual responsibilities to the accomplishment of mission objectives and service member health and safety. Adjustments to shift work, attention to pre-deployment screening and education, and a change in cultural attitudes that sleep is a luxury and not a biologic requirement, can make a positive difference in morale, mission completion, and casualty reductions.

FUNDING

This study was funded through a TriService Nursing Research Program Grant number: HU0001-16-1-TS01 (N16-501). The work reported herein was performed under United States Air Force Surgeon General General-approved Protocol Study No. FSG20160012H.

CONFLICTS OF INTEREST

The authors have no conflicts of interest to declare.

DISCLAIMER

The information of content and conclusions do not necessarily represent the official position or policy of, nor should any official endorsement be inferred by, the TriService Nursing Research program, Uniformed Services University of the Health Sciences, the Department of Defense, or the US Government.

REFERENCES


Kunert, K., King, M. L., & Kolkhorst, F. W. (2007). Fatigue and
sleep quality in nurses. *Journal of Psychosocial Nursing & Mental Health Services, 45*(8), 30–37. https://doi.org/10.3928/02793695-20070801-07


Copyright © 2020 Disaster Nursing Global Leader Degree Program