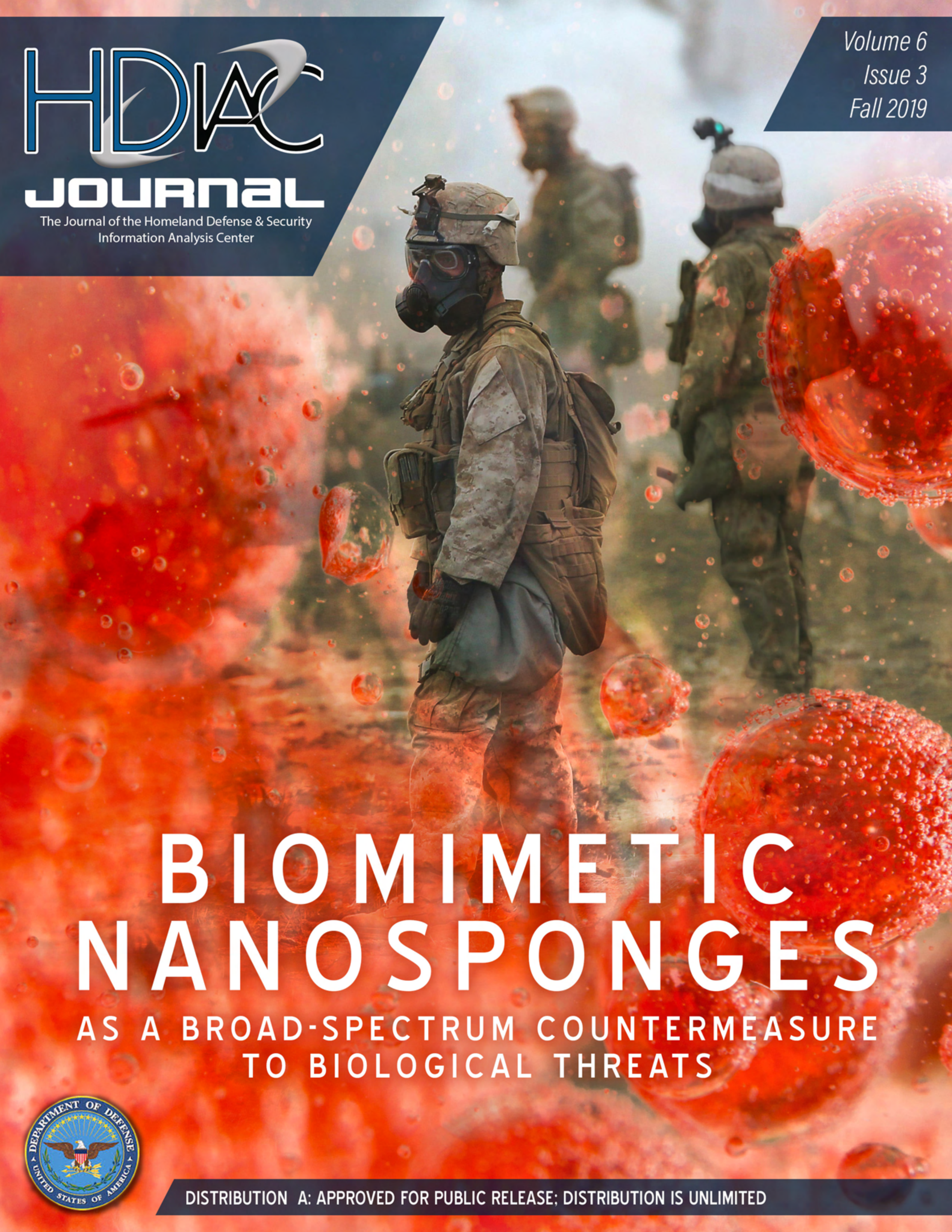


HDIAC JOURNAL
The Journal of the Homeland Defense & Security Information Analysis Center

Volume 6
Issue 3
Fall 2019

BIOMIMETIC NANOSPONGES

AS A BROAD-SPECTRUM COUNTERMEASURE TO BIOLOGICAL THREATS

HDIAC JOURNAL
The Journal of the Homeland Defense & Security Information Analysis Center

Volume 6 • Issue 3 • Fall 2019

Director: Stuart Stough

The Journal of the Homeland Defense & Security Information Analysis Center (HDIAC) is published quarterly by the HDIAC staff. HDIAC is a DoD sponsored Information Analysis Center (IAC) with policy oversight provided by the Under Secretary of Defense for Research and Engineering (USD (R&E)), and administratively managed by the Defense Technical Information Center (DTIC). HDIAC is operated by Information International Associates (IIA) in Oak Ridge, Tenn. Reference herein to any specific commercial products, processes or services by trade name, trademark, manufacturer or otherwise, does not necessarily constitute or imply its endorsement, recommendation or favoring by the United States government or HDIAC. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States government or HDIAC and shall not be used for advertising or product endorsement purposes.

ISSN 2578-0832 (print)
ISSN 2578-0840 (online)

ON THE COVER

Photo illustration created by HDIAC and adapted from a U.S. Marine Corps photo (available for viewing at: https://media.defense.gov/2016/Aug/22/2001612174/-1/-1/0/160818-M-LV138-299.JPG) and Adobe Stock.

Alternative Energy
Biometrics
CBRN Defense
Critical Infrastructure Protection
Cultural Studies
Homeland Defense & Security
Medical
Weapons of Mass Destruction

Table of Contents

Contact

HDIAC Headquarters
104 Union Valley Road
Oak Ridge, TN 37830 • 865-535-0088

Fatena Casey, Technical Office Manager
Joel Hewett, Subject Matter Expert
David McCarville, Social & Multi-Media Editor
Tim Gould, Graphic Designer
Morgan Wells, MIssion Support Specialist

Emese Horvath
Contracting Officer Representative

DoD Information Analysis Centers
8725 John J. Kingman Road
Fort Belvoir, VA 22060
571-448-9753
emese.i.horvath.civ@mail.mil

Follow us on Social Media

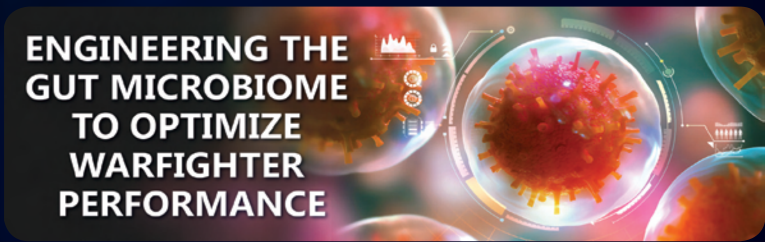
DoD_HDIAC
HDIAC Outreach
DoDHDIAC
HDIAC
dod_hdiac

IN CASE YOU MISSED IT

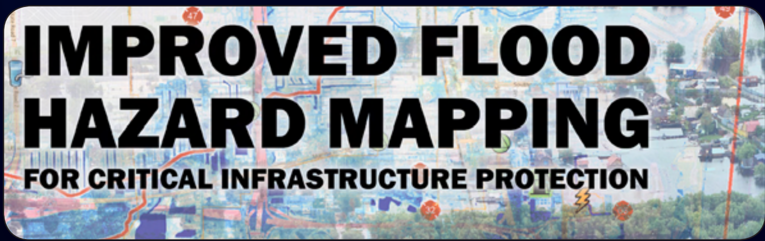
Tech Talk featuring research from HDIAC SMEs Richard LaFountain, Ph.D., and Parker Hyde, Ph.D. (Ohio State University). Watch the video on the HDIAC website in our [Tech Talk Archive.]

Webinar featuring HDIAC SMEs Julia Hirschberg, Ph.D., and Sarah Ita Levitan, Ph.D. (Columbia University). Watch the recording and view the presentation slides on the HDIAC website in our [Webinar Archive.]

Spotlight Article examining technologies for CBRN detection by HDIAC Research Analyst Robert Leedy. Read the Spotlight on the HDIAC website in our [Spotlight Archive.]



Spotlight Article discussing gut health for military personnel by HDIAC SMEs Kyle Giesler, Ph.D., and John Saindon, Ph.D. Read the Spotlight on the HDIAC website in our [Spotlight Archive.]



Spotlight Article exploring flood mapping technologies by HDIAC Research Analyst Vincent Palandro. Read the Spotlight on the HDIAC website in our [Spotlight Archive.]

About the Journal

The HDIAC Journal is a quarterly publication featuring articles by researchers, Subject Matter Experts (SMEs), and other leaders in HDIAC's eight focus areas. Authors come from academia, industry, and government/military institutions to discuss their latest research and innovations. With a network of more than 10,000 subscribers, including military/government program managers and other decision makers, the HDIAC Journal is an excellent platform for researchers to showcase their work. Each issue of the HDIAC Journal is available online, and physcial copies are available at conferences and events. If you are interested in writing for the HDIAC Journal, please contact publications@hdiac.org

HDIAC Products & Services

HDIAC is responsible for aquiring, analyzing, and disseminating relevant scientific and technical information in support of DoD and U.S. government research and development activities. In order to accomplish this, HDIAC provides a variety of products and services:

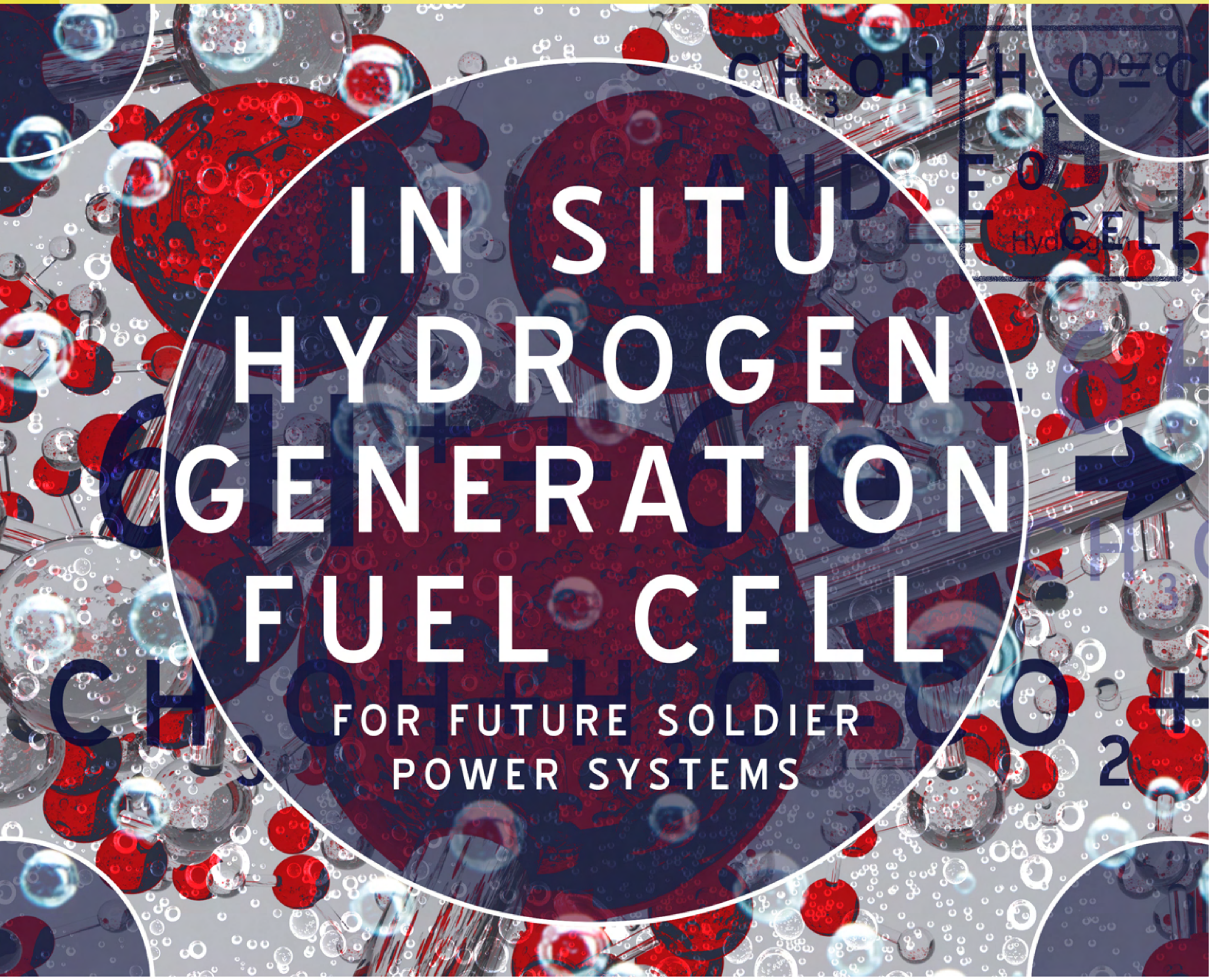
Technical Inquiry Service
HDIAC provides up to four free hours of information services, including literature searches, product/document requests, and analysis within our eight focus areas. This service is available to academia, industry, and other government agencies.

Core Analysis Task
This contract vehicle has a rapid start time and is designed to improve industry productivity and provide affordable, value-added capability to the warfighter. The HDIAC Core Analysis Task (CAT) incentivizes productivity and facilitates technological innovation though collaboration between industry and government.

Spotlight Articles
HDIAC Spotlights are brief, high-level overviews of relevant research and emerging technologies published on the HDIAC website throughout the year.

Webinar Series
HDIAC Webinars are presented by HDIAC SMEs and offer in-depth analysis of relevant research and emerging technologies.

Tech Talks
HDIAC Tech Talks are informational videos featuring DoD-applicable research in out eight focus areas.



IN SITU HYDROGEN GENERATION FUEL CELL

FOR FUTURE SOLDIER POWER SYSTEMS

Rongzhong Jiang
& Dat T. Tran

A highly efficient and compact electric power source is required on the battlefield in order to operate telecommunication devices, night-vision goggles, sensors, portable computers, and unmanned aerial vehicles. While batteries can provide the power needed for short-term missions, they come with several limitations. It is impossible for warfighters to carry a sufficient number of batteries for a mission longer than a few days.

As a result, much research and development attention has turned to fuel cells—an electrochemical device that converts a fuel's chemical energy into electric power—for long-duration applications [1–3]. The most efficient fuel cells use hydrogen as their fuel source, as it has high power, high energy, and zero CO_2 emissions [4].

An innovative, in situ hydrogen generation fuel cell has been developed at the U.S. Army Combat Capabilities Development Command's Army Research Laboratory (ARL) for future warfighter power systems. The fuel cell is composed of two functional parts: one for hydrogen generation, by methanol electrolysis; and the other for electric power generation, through an internal hydrogen/air power plant. No external tank is required to store compressed hydrogen gas, and all hydrogen generated is used instantly by the hydrogen/air fuel cell. The hydrogen gas generated in the first part is directly transported to the second part through a 2mm thick junction electrode for electricity generation.

For use in the in situ fuel cell, our team at ARL developed high-potential cathode catalysts for hydrogen evolution and oxygen reduction, and anode catalysts for alcohol electrooxidation. Experimental results to date have demonstrated that 100g of methanol can be converted to ~16.8g of hydrogen gas (excluding the mass of water). A prototype has been tested by running methyl alcohol. A fuel energy density of approximately 1,300 Wh/kg of fuel was achieved—a level 56% higher than standard methanol-reforming fuel cells (~830 Wh/kg) [5]. This technology will enable the warfighter to power a variety of electric devices with a compact, light-

Image Credit | *Photo illustration created by HDIAC and adapted from Adobe Stock.*

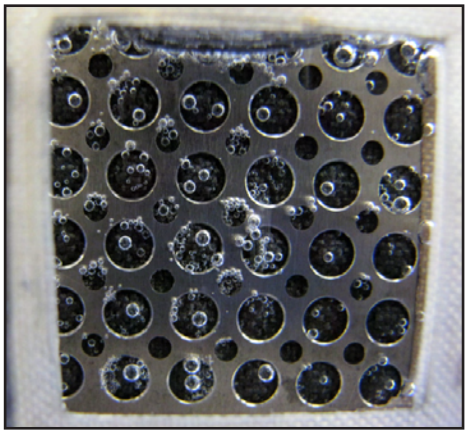


Figure 1. Hydrogen generation at the cathode electrode

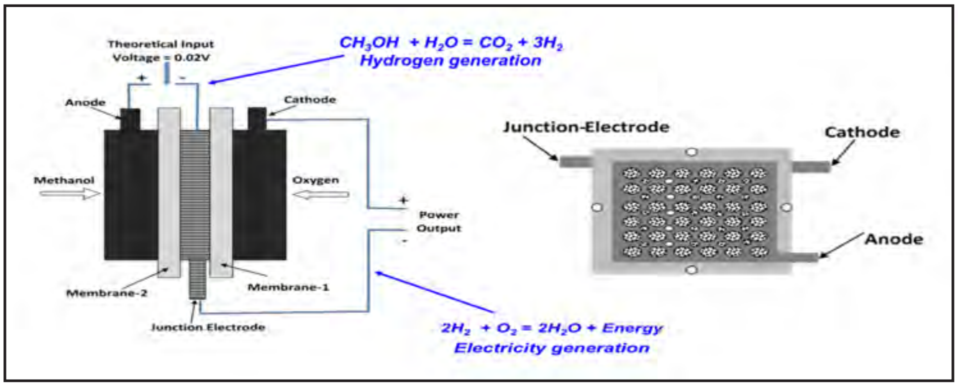


Figure 2. Schematic view of an in situ hydrogen generation fuel cell using methanol to generate hydrogen. (Left) A combination of an electrolyzer and a fuel cell joined by a junction electrode and (Right) experimental components containing three electrodes and two electrolyte membranes.

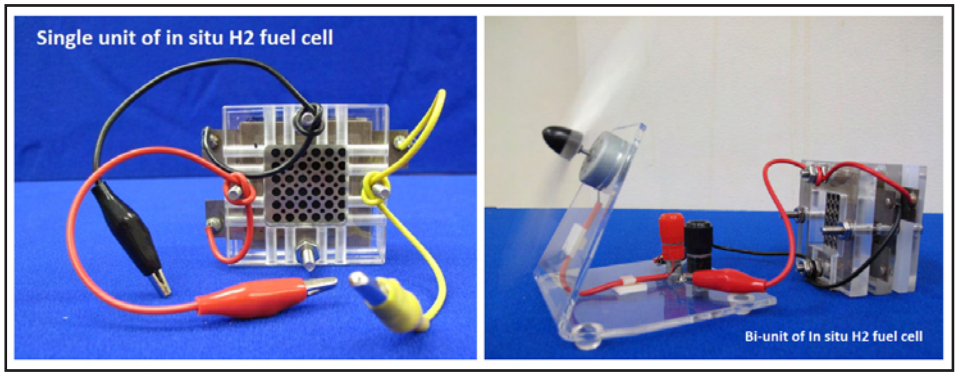


Figure 3. Single unit (Left) and bi-unit (Right) of in situ hydrogen fuel cells using methanol to generate hydrogen

weight, high-power, and high-energy power source capable of powering longer missions than those supported by batteries alone.

Hydrogen-Generation Fuel Cell Technologies

Several onboard hydrogen-generation technologies have been proposed by using hydrogen rich solid or liquid materials to produce hydrogen gas for fuel cells. The U.S. Army evaluated three of these technologies for onboard hydrogen generation for fuel cells.

The first is thermal decomposition of alane (AlH_3), which has a theoretical hydrogen content of 10% (wt%) and an experimental hydrogen yield of 7.7–9.2% (wt%) [6–8]. Although the initial release of hydrogen gas occurs at <100°C from alane, the full hydrogen yield cannot be obtained despite temperatures as high as >250°C. Furthermore, the high temperature decomposition of alane causes pressure buildup and safety concerns [8].

The second candidate for onboard hydrogen generation is hydrolysis of aluminum (Al) metal nanoparticles by the reaction of Al with water, which requires large amounts of water for a water/Al ratio >3/1 [9–11]. The Al method's theoretical hydrogen yield is only 3.7% (wt%), including water. Adding a large quantity of water limits automobile and portable electronics applications.

The third candidate is onboard thermal-reforming methanol to generate hydrogen gas for a fuel cell [12–13], which needs about 30% water for methanol steam reforming at 200–300°C. In order to maintain a high reforming temperature, additional methanol must be burned in the thermal-reforming methanol fuel cell, which further lowers the fuel energy density and results in a density ~830 Wh/kg. In addition, thermally reforming methanol causes a carbon monoxide (CO) release as high as 1,000 ppm at startup, which creates an inhalation safety concern. (These results were obtained from ARL's evaluation of several reforming methanol fuel cell systems manufactured by UltraCell, LLC.)

In a methanol-reforming fuel cell, the process of reforming must be carried out at 250–300°C to generate and store hydrogen gas. The temperature of hydrogen from the reformer is too high for a polymer electrolyte membrane fuel cell (20–100°C), and also too high for a fuel cell using polybenzimidazole as an electrolyte membrane (160–180°C). All methanol-reforming fuel cell systems require a cooling system. These high operating temperatures create issues related to cost, complex system design, slow startup and shutdown cycle speeds, and a shortened lifespan.

After reviewing these onboard-hydrogen generation technologies, we designed, constructed, and evaluated a prototype of an innovative in situ hydrogen generation fuel cell capable of providing electric power continuously. The in situ hydrogen fuel cell works at low temperatures (20–90°C) to *electrochemically* process methanol for hydrogen generation and electric power generation, simultaneously. Unlike other hydrogen generation architectures, the in situ prototype does not need to carry a compressed hydrogen gas tank, as its hydrogen is self-produced via hydrogen-rich alcohols.

Method of Electrochemically Converting Methanol to Hydrogen

Methanol can be electrochemically converted into hydrogen within an electrochemical cell. At the anode:

$$CH_3OH + H_2O = CO_2 + 6H^+ + 6e^-, \text{ and } E^0 = 0.02V. \quad (1)$$

At the cathode:

$$6H^+ + 6e^- \rightarrow 3H_2, \text{ and } E^0 = 0.0V. \quad (2)$$

The overall electrochemical reaction is:

$CH_3OH + H_2O = CO_2 + 3H_2$, and $E^0_{cell} = 0.02V$. (3)

Theoretically, we only need to apply a 20-mV electric voltage to convert methanol to hydrogen gas. That is, we only need to apply a small amount of energy to convert methanol into hydrogen gas by a method of electrolysis, and then use the hydrogen to generate more electrical energy. This is a very promising method for producing hydrogen on demand. Furthermore, methanol has a high hydrogen content of 12% by mass, including water, and 18.75%, excluding water. Because the in situ fuel cell generates water, if the generated water is recycled, the hydrogen yield may approach 18% by mass.

Experimental results demonstrated that 90% hydrogen yield can be obtained electrochemically for hydrogen generation [14]. Therefore, 100g of methanol can be converted up to 16.8g hydrogen gas (excluding the mass of water). Compared to alane, Al-nano particles (or Al-alloy), and thermal-reforming technologies for hydrogen generation, the methanol electrolysis technology has the highest hydrogen yield. In addition to methanol, several other alcohols can be used to generate hydrogen (e.g., iso-propanol alcohol [rubbing alcohol]). Figure 1 shows hydrogen generation at the cathode electrode of an electrochemical apparatus [14] by electrolysis of 4M 2-propanol alcohol.

Design and Assembly of an In Situ Hydrogen Generation Fuel Cell Prototype

In order to realize in situ generation of hydrogen gas, we designed, assembled, and experimentally tested a fuel cell prototype [15]. Figure 2 presents a schematic view of a methanol-fueled in situ hydrogen generation fuel cell. The prototype is composed of two functional parts—one for hydrogen generation by methanol electrolysis and the other for electric power generation through a hydrogen/air fuel cell. The two parts share a junction electrode, which is a cathode for the hydrogen generation electrolyzer, and an anode for hydrogen/air fuel cell. This junction electrode is designed with micropores for gas and water transport. This design minimized the distance of hydrogen gas transfer between the hydrogen generation electrolyzer and the hydrogen/air fuel cell. The hydrogen transfer distance is less than 2mm across a junction electrode. As long as the hydrogen is generated, it is used instantly for

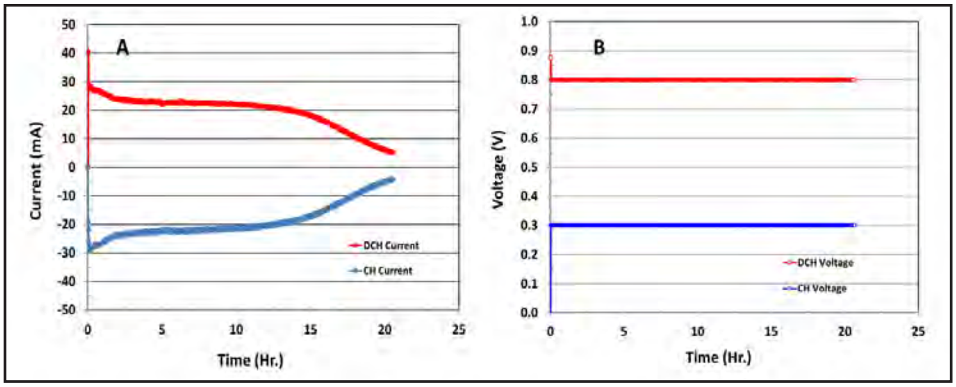


Figure 4. Performance of a prototype of in situ hydrogen generation fuel cell operated with 6-mL, 1.0M methanol at 30°C

Power Source	Energy Density (Fuel Only) (Wh/kg)
In situ hydrogen fuel cell	1300
Thermal-reforming methanol fuel cell	830
Direct-methanol fuel cell	1000

Table 1. Comparison of energy density from three types of fuel cell systems at ARL

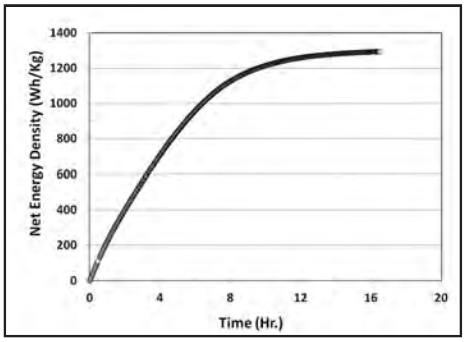


Figure 5. Net energy density obtained from a prototype of in situ hydrogen generation fuel cell operated with 6mL, 1.0M methanol at 30°C

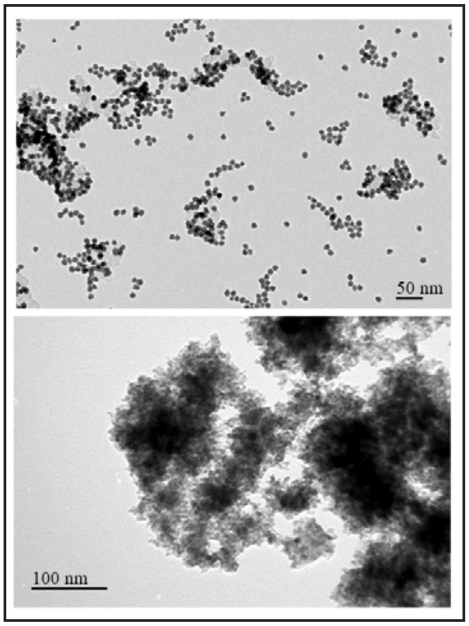


Figure 6. TEM images of cathode catalyst of PtCo (Top) for hydrogen evolution and oxygen reduction and anode catalyst of PdNiP (Bottom) for alcohol electro-oxidation

electric power generation in the part of hydrogen/air fuel cell. By eliminating the need for an external gas tank, the in situ fuel cell reduces the weight and size of a hydrogen fuel cell system.

Evaluation and Results

We built experimental prototypes and examined the feasibility of an in situ hydrogen generation fuel cell (see Figure 2). A single unit of the prototype (see left photo in Figure 3) consists of two membrane electrolyte assemblies (MEAs)—one as an electrolyzer for hydrogen generation and the other as a hydrogen/air fuel cell for electric power generation. A junction electrode was positioned between membrane-1 and membrane-2 (see left diagram in Figure 2). The single unit prototype was used for the experimental proof of the proposed concept's feasibility. The bi-unit of the prototype (see right photo in Figure 3) is a combination of two single units, which was used for demonstrations such as running a small electric fan.

Figure 4 shows experimental results obtained from a single unit of prototype [14]. Here, 6mL 1.0M methanol is filled into the fuel chamber and the temperature is held at 30°C. The input voltage (CH voltage), or the electrolyzer's voltage, is constant at 0.30V, and the output voltage (DCH voltage), or the fuel cell's voltage, responds at 0.80V. The overall net voltage output is 0.50V. In this example, the input current (CH current) is provided by an external power for the electrolyzer, and the output current (DCH current) is generated by a hydrogen/air fuel cell. The experimental result shows that the input current is approximately equal to the output current or ~100% coulomb efficiency from the fuel processing electrolyzer to the fuel cell is obtained. The methanol is first converted into hydrogen in the fuel-processing electrolyzer. Meanwhile, the hydrogen is simultaneously used in the hydrogen/air fuel cell for electric power generation.

Figure 5 shows fuel energy density obtained with a single unit of in situ hydrogen generation fuel cell. This single unit prototype is operated with 6mL 1.0M methanol at 30°C. Here, CH voltage is held at 0.4V, and DCH voltage responds at 0.75V. The energy density obtained from this prototype is 1,300 Wh/kg (fuel only). Table 1 lists energy density obtained from three types of fuel cell systems evaluated at ARL. The in situ hydrogen generation fuel cell gives the highest fuel energy density, which is 56% higher than the thermal-reforming methanol fuel cell.

The Role of Electrode Catalysts

Practically, methanol is not electrooxidized at the anode, nor is hydrogen generated at the cathode by applying only a theoretical potential of 20mV. This is due to overpotentials caused by slow electrode kinetics and the fuel cell's internal resistance. Therefore, we had to develop efficient electrode catalysts for sufficiently lowering the overpotentials of methanol electrolysis. At ARL, we developed high-potential cathode catalysts for hydrogen evolution and oxygen reduction, and anode catalysts for alcohol electrooxidation [15–17]. Figure 6 shows transmission electron microscopic (TEM) images of a PtCo cathode catalyst [15, 18] and PdNiP anode catalyst [16].

Conclusion

The hydrogen-powered fuel cell is a promising power source for the Department of Defense, capable of providing continuous electric power for long-term missions. Our experimental results demonstrate that 100g methanol can be converted into ~16.8g hydrogen gas, excluding the mass of water. These results could help make significant strides in meeting the challenges of supplying the warfighter with ample expeditionary power, by generating on-demand energy unencumbered by hydrogen gas storage tanks.

Acknowledgments

The authors would like to thank Drs. Deryn Chu and Cynthia Lundgren for their discussions and the U.S. Department of the Army and U.S. Army Materiel Command for supporting this work.

References

1. Srinivasan, S. (1989). Fuel cells for extraterrestrial and terrestrial applications. *Journal of the Electrochemical Society, 136*(2). doi:10.1149/1.2096647
2. Bacon, F. (1969). Fuel cells, past, present and future. *Electrochimica Acta, 14*(7), 569–585. doi:10.1016/0013-4686(69)87042-8
3. Perry, M. L., & Fuller, T. F. (2002). A historical perspective of fuel cell technology in the 20th century. *Journal of the Electrochemical Society, 149*(7). doi:10.1149/1.1488651
4. Colella, W., Jacobson, M., & Golden, D. (2005). Switching to a U.S. hydrogen fuel cell vehicle fleet: The resultant change in emissions, energy use, and greenhouse gases. *Journal of Power Sources, 150*, 150–181. doi:10.1016/j.jpowsour.2005.05.092
5. ARL have evaluated several methanol-reforming fuel cell systems from UltraCell, LLC. At optimized operating condition, these fuel cell systems gave power output of 20–25 watts with fuel energy density ~830 wh/kg.
6. Ahluwalia, R., Hua, T., & Peng, J. (2009). Automotive storage of hydrogen in alane. *International Journal of Hydrogen Energy, 34*(18), 7731–7740. doi:10.1016/j.ijhydene.2009.07.013
7. Thampan, T., Atwater, T., Cook, C., Novoa, J., & Sutorik, A. (2016). Hydrogen generation from aluminum hydride for wearable polymer electrolyte membrane fuel cells. *International Journal of Hydrogen Energy, 41*(22), 9402–9409. doi:10.1016/j.ijhydene.2016.04.095
8. Grew, K. N., Brownlee, Z. B., Shukla, K. C., & Chu, D. (2012). Assessment of alane as a hydrogen storage media for portable fuel cell power sources. *Journal of Power Sources, 217*, 417–430. doi:10.1016/j.jpowsour.2012.06.007
9. Huang, X., Gao, T., Pan, X., Wei, D., Lv, C., Qin, L., & Huang, Y. (2013). A review: Feasibility of hydrogen generation from the reaction between aluminum and water for fuel cell applications. *Journal of Power Sources, 229*, 133–140. doi:10.1016/j.jpowsour.2012.12.016
10. Rosenband, V., & Gany, A. (2010). Application of activated aluminum powder for generation of hydrogen from water. *International Journal of Hydrogen Energy, 35*(20), 10898–10904. doi:10.1016/j.ijhydene.2010.07.019
11. Prabu, S., Hsu, S., Lin, J., & Wang, H. (2018). Rapid hydrogen generation from the reaction of aluminum powders and water using synthesized aluminum hydroxide catalysts. *Topics in Catalysis, 61*(15-17), 1633–1640. doi:10.1007/s11244-018-0970-x
12. Iulianelli, A., Ribeirinha, P., Mendes, A., & Basile, A. (2014). Methanol steam reforming for hydrogen generation via conventional and membrane reactors: A review. *Renewable and Sustainable Energy Reviews, 29*, 355–368. doi:10.1016/j.rser.2013.08.032
13. Li, M., Duraiswamy, K., & Knobbe, M. (2012). Adsorption enhanced steam reforming of methanol for hydrogen generation in conjunction with fuel cell: Process design and reactor dynamics. *Chemical Engineering Science, 67*(1), 26–33. doi:10.1016/j.ces.2011.07.024
14. Jiang, R. Z., Tran, D. T., & Chu, D. D. (2015, September 8). Power-free apparatus for hydrogen generation from alcohol. U.S. Patent No. 9,127,370. Washington, DC: U.S. Patent and Trademark Office.
15. Rong, C., Jiang, R., Sarney, W., & Chu, D. (2010). Ultrasound-assisted micro-emulsion for synthesis of Pt and PtCo nano-particles. *Electrochimica Acta, 55*(22), 6872–6878. doi:10.1016/j.electacta.2010.05.078
16. Jiang, R., Rong, C., & Chu, D. (2011). Surface coverage of Pt atoms on PtCo nanoparticles and catalytic kinetics for oxygen reduction. *Electrochimica Acta, 56*(5), 2532–2540. doi:10.1016/j.electacta.2010.11.060
17. Jiang, R., Tran, D. T., Mcclure, J. P., & Chu, D. (2014). A class of (Pd–Ni–P) electrocatalysts for the ethanol oxidation reaction in alkaline media. *ACS Catalysis, 4*(8), 2577–2586. doi:10.1021/cs500462z
18. Jiang, R. Z., Tran, D. T., & Chu, D. D. (2017, January 3). Methods and apparatus of an anode/cathode (A/C) junction fuel cell with solid electrolyte. U.S. Patent No. 9,537,167. Washington, DC: U.S. Patent and Trademark Office.

Rongzhong Jiang, Ph.D.
Senior Research Scientist, CCDC Army Research Laboratory (ARL)

Rongzhong Jiang is a senior research scientist at ARL responsible for developing innovative materials and technologies for fuel cells and electrochemical devices. His research experience includes hydrogen/air fuel cells, direct alcohol fuel cells, modeling and evaluating fuel cell stacks, electrocatalysis, and kinetics. He has authored 105 publications, including 4 book chapters, 66 refereed journal articles, 24 symposium publications, and 11 government publications. In addition, he has received nine U.S. patents, two Army Research and Development (R&D) awards, one Best Paper award from an Army science conference, and four ARL Directors' Research Initiative (DRI) awards. Dr. Jiang obtained his Ph.D. in chemistry at Changchun Institute of Applied Chemistry, Chinese Academy of Sciences, followed by post-docs at California Institute of Technology (Caltech) and Lawrence Berkley National Laboratory.

Dat T. Tran, Ph.D.
Senior Research Chemist, CCDC Army Research Laboratory (ARL)

Dat T. Tran is a senior research chemist at ARL with more than 15 years of experience in the synthesis and characterization of inorganic materials and 10 years of experience in jet fuel analysis. His most recent research is focused on developing adsorbent materials for Army fuel cell applications. His other research interests include catalysts and materials related to fuel cell, battery, and combustion applications. Dr. Tran obtained his Ph.D. in materials chemistry at Binghamton University-NY, followed by a post-doc at Cornell University.

SYNTHETIC OPIOIDS:

A NEW CLASS OF CHEMICAL WEAPONS?

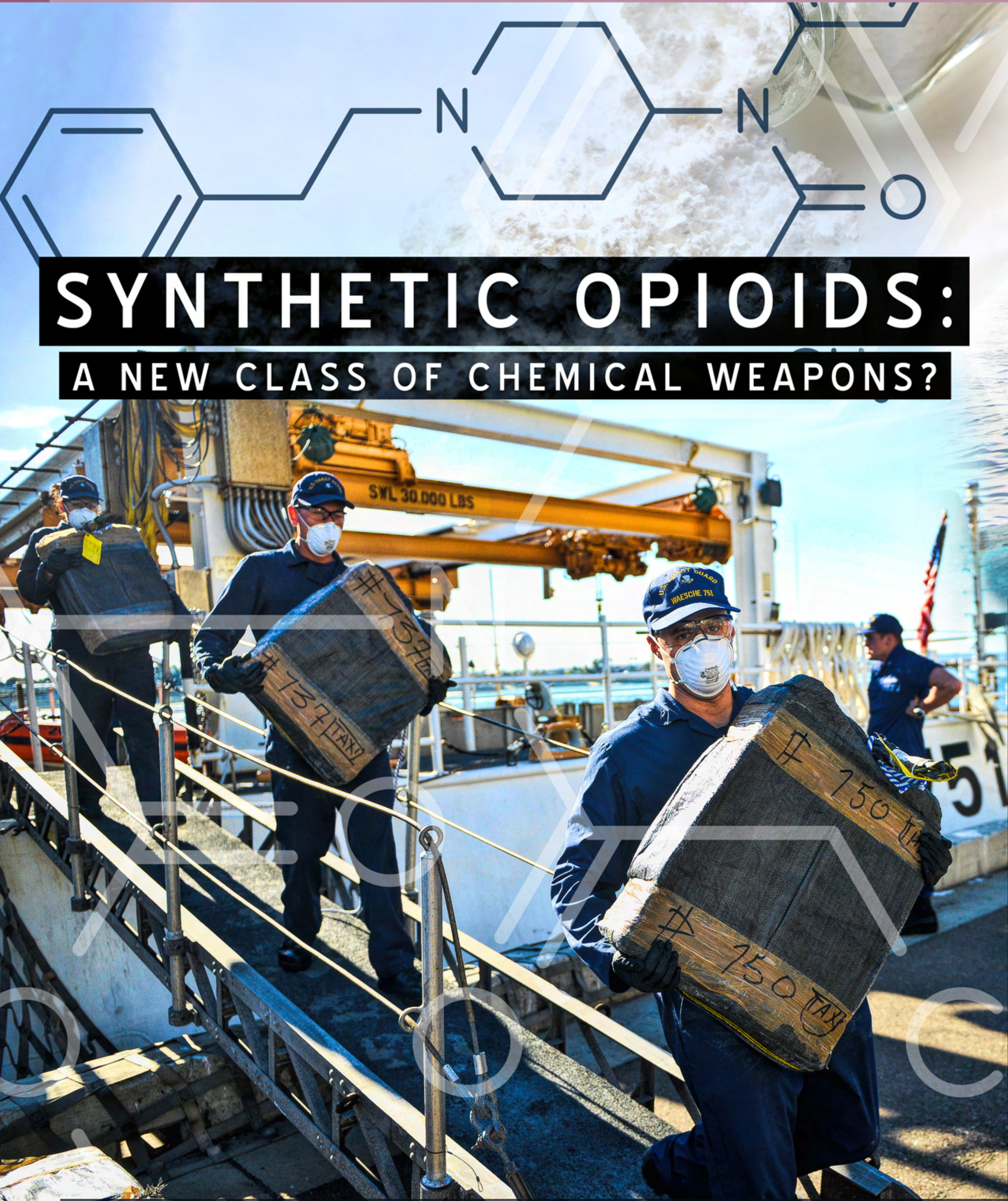


Image Credit *Photo illustration created by HDIAC and adapted from Coast Guard photos by LaNola Stone (available for viewing at: https://coastguard.dodlive.mil/2015/08/in-the-zone-a-crews-determination-over-the-horizon-part-2/) and Luke Pinneo (available for viewing at: https://coastguard.dodlive.mil/2015/07/225-years-of-service-to-nation-drug-interdiction/) and Adobe Stock.*

Kyle E. Giesler
& John M. Saindon

Article II of the Chemical Weapons Convention (CWC) broadly defines a toxic chemical as any compound “that can cause death, temporary incapacitation or harm to humans or animals [1].” Figure 1 shows the five major classes of toxic chemical agents outlined by the CWC. With the exception of riot control agents used by law enforcement and the anti-proliferative properties of certain nitrogen mustards [2], many of these substances have no medical use and their weaponization is banned by the 193 nations committed to the Convention.

However, a potentially new class of chemical weapons is beginning to emerge—synthetic opioids. In 2017, opioids were responsible for more than 47,000 deaths in the U.S. and there is growing concern regarding their use in next-generation chemical weapons [3].

Apprehension is felt at the national and international level with the U.S. Department of Homeland Security (DHS) positioned to designate fentanyl a weapon of mass destruction and China pledging to halt the production of synthetic opioids altogether. Herein, we discuss the therapeutic properties of opioids, their pharmacology and toxicity, and the potential misuse of synthetic opioids in the theater of war.

Medical Use: A Double-Edged Sword

Opioids are highly addictive analgesics widely used in the management of moderate to severe pain. The Centers for Disease Control and Prevention (CDC) estimates that in 2017, 16% of U.S. counties saw enough opioids prescribed for every person to have a prescription [4]. These include a handful of natural, synthetic, and semi-synthetic compounds, such as morphine, oxycodone, codeine, tramadol, fentanyl, and hydrocodone.

The pervasive use of opioids in the U.S. healthcare system distinguishes them from other chemical agents shown in Figure 1. Opioids offer many advantages over other analgesic drugs, namely superior pain relief, high central nervous system (CNS) permeability, a long duration of action, few cardiovascular risks, and a fast onset time.

Despite these therapeutic advantages, opioids produce fatal side effects when administered at the incorrect dose, echoing the onerous toxicology adage “the dose makes the poison.” This is particularly true for highly potent synthetic opioids like fentanyl and its derivatives (shown in Figure 2). A lethal dose of fentanyl is approximately 2 mg—the equivalent of a few grains of salt (see Figure 3) [5].

Opioid Pharmacology and Mechanism of Toxicity

The on-target toxicity of opioids is inherently tied to the complexity of opioid signaling. Agonist opioids like fentanyl produce analgesia upon binding to their cognate receptors located throughout the central and peripheral nervous systems. Of the three classical opioid receptors (μ, κ, δ) it is thought that the activation of the μ-opioid receptor (MOR) in the brain is the primary mechanism of opioid-induced analgesia [6]. However, receptor activation is tightly regulated by β-arrestin2 to control the extent and duration of opioid signaling after the receptor binds to its target [7]. Regulation by β-arrestin2 does not produce analgesia, but rather “arrests” opioid signaling and triggers a series of downstream signaling events that briefly remove the receptor from the cell surface. The temporary disappearance of MORs within the brain creates a negative feedback loop that results in tolerance, requiring more opioids to produce analgesia with the remaining receptors. However, if additional opioids are administered beyond the therapeutic dose, the remaining receptor population also becomes compromised, further dampening MOR signaling throughout the CNS which leads to significant respiratory suppression and subsequent death [8]. Indeed, this is the primary mechanism behind opioid-related overdose fatalities and is the very property that transforms therapeutic opioids into potential chemical weapons.

Weaponization of Synthetic Opioids

The idea of opioid-based chemical weapons dates back to the discovery of fentanyl in 1960. Fentanyl and several fentanyl analogues were the focus of extensive U.S. military research from the 1970s through the 1990s, including inhalation studies on non-human primates to develop novel incapacitating agents "for situations where a quick knock-down agent is needed [9]." These efforts led to the discovery of sufentanil and carfentanil, two of the most potent synthetic analogues known, with respective potencies of 500x and 10,000x that of morphine (see Figure 2). Carfentanil causes severe respiratory depression and death in non-human primates at doses ranging from 2–14 µg/kg and its aerosolization may be more toxic than the nerve agent VX [10]. In 2016, the appearance of carfentanil in illegal markets prompted the U.S. Drug Enforcement Administration to issue an explicit warning stating that the drug is a serious danger to public safety, law enforcement, and healthcare personnel [11]. Carfentanil was responsible for approximately 11% of all opioid-overdose deaths in ten states from June 2016 to July 2017 and has a high potential for weaponization [12].

Most U.S. programs relating to the weaponization of synthetic opioids were terminated shortly before the U.S. joined the CWC in 1993; however, research in other countries continued [9]. In 2002, Russian authorities used an undisclosed mixture of aerosolized synthetic opioids believed to contain carfentanil (dubbed Kolokol-1) to counter forty insurgents during the Moscow theater crisis [13]. All forty insurgents were killed along with 204 unsuspecting hostages after exposure to the Kolokol-1 aerosol, revealing the potential use and limitations associated with synthetic opioids as non-lethal incapacitating agents.

The lethal consequences of Kolokol-1 did not deter Russian authorities from using synthetic opioids to counter other domestic terrorist groups. In 2005, Russian Special Forces descended on the town of Nalchik and subsequently deployed another unknown aerosol agent against militants holding hostages in a local shop. It is believed that the aerosol contained carfentanil and remifentanil, similar to the Kolokol-1 cocktail used in the Moscow incident [14]. Later that year, Russian researchers announced, but did not fully disclose, progress toward the development of additional incapacitating agents at the 3rd European Symposium on Non-Lethal Weapons [15].

It's unclear if these programs remain in operation today or if Russia has a continued interest in weaponizing synthetic opioids. However, growing precedent suggests that Russia is actively trying to obstruct CWC provisions relating to chemical weapon investigations in a likely attempt to shroud their motives, increasing political tensions between Moscow and the West [16].

To the best of our knowledge, there are no additional reports disclosing the use of weaponized opioids in acts of terror or on the battlefield at the time of this writing. Nonetheless, concerns are mounting. In January 2019, Massachusetts senator Ed Markey contacted Secretary of State Mike Pompeo and Acting Secretary of Defense Patrick Shanahan asking if their departments have "devised strategies to deal with the potential weaponization of fentanyl [17]." Fentanyl was responsible for nearly 28,000 overdose deaths in 2017 [18]. In response, DHS recently announced that it was considering classifying fentanyl as a weapon of mass destruction [19]. The announcement seeks to increase international awareness and the severity of repercussions to those harboring and distributing large quantities of fentanyl.

China has been the main supplier of fentanyl in the U.S., Canada, and Mexico, although it's unclear if radical terrorist groups or nations have taken steps to exploit China's supply for nefarious means [20]. Interestingly, China revealed a nation-wide ban on the production of synthetic opioids in April 2019, shortly before DHS announced its position on fentanyl [21]. The cross-talk between China and the Trump Administration represents a clear and coordinated effort to restrict the flow of synthetic opioids into illicit markets and to minimize potential security threats that may arise from the weaponization of China's fentanyl supply.

Antidotes

Naloxone hydrochloride (Narcan®) is the most pervasive antidote on the market and is widely distributed to emergency medical services (EMS) at the forefront of the opioid epidemic. Originally developed in the 1960s, naloxone shares many structural features with morphine (see Figure 4); however, subtle differences in structure give rise to significantly disparate properties. Unlike morphine naloxone is a competitive MOR antagonist

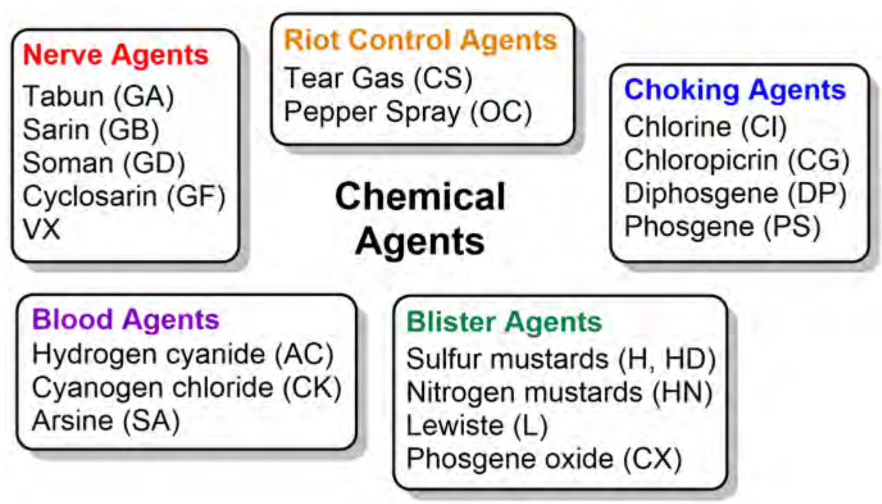


Figure 1. The five major classes of chemical agents identified by the Chemical Weapons Convention with examples listed in each group.

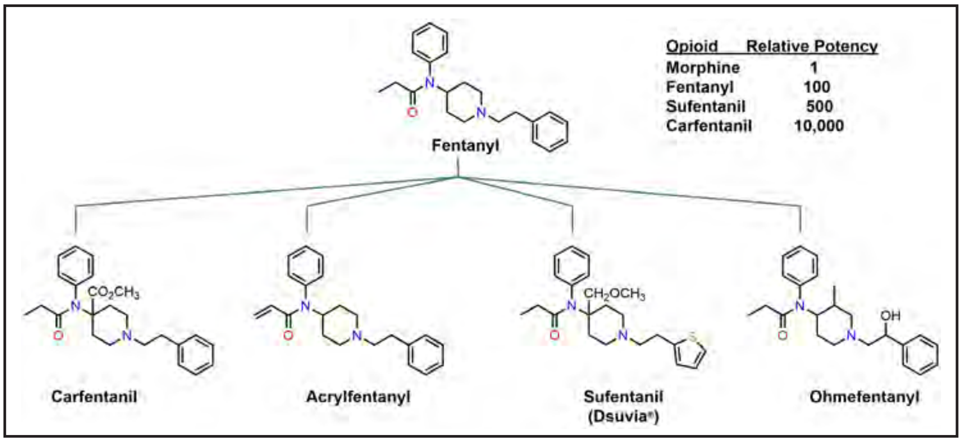


Figure 2. Many synthetic opioids are derivatives of fentanyl and are significantly more potent than morphine.

Figure 3. A lethal dose of fentanyl. Fentanyl is approximately 100-fold more potent than morphine and boasts an LD_{50} of 0.03 mg/kg in monkeys [5]. (Source: Drug Enforcement Administration)

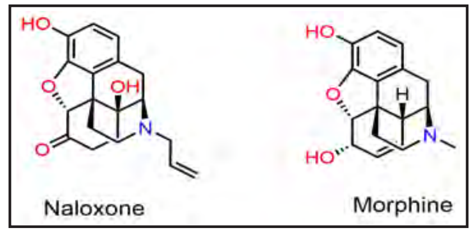


Figure 4. The chemical structures of naloxone and morphine share many structural similarities. However, naloxone is an antidote whereas morphine is an opioid.

that effectively dislodges opioids from the receptor to reverse opioid-induced toxicity, specifically respiratory suppression.

The efficacy of naloxone has been extensively reviewed and depends on several key factors, including the route of administration, dose, and the identity of the opioid that is to be countered [22]. Naloxone has limited oral efficacy and is typically administered as an intravenous, intramuscular, or subcutaneous injection to circumvent its poor oral bioavailability and extensive first-pass metabolism. Complementary intranasal formulations are also available that have a rapid onset time and eliminate the use of needles [23]. In 2019, the U.S. Food and Drug Administration (FDA) approved the first generic formulation of naloxone hydrochloride nasal spray for use in a community setting by individuals without medical training [24]. Increased access to a user-friendly antidote will likely lessen the burden on EMS personnel and place more control in the hands of the public to mitigate the opioid epidemic. The technology may also find significant value in protecting civilians against a potential opioid-based chemical attack.

Next-Generation Medical Countermeasures

Auto-injectors containing 0.4 mg of naloxone hydrochloride (Evzio®) and Narcan® nasal spray (2 mg/mL) sufficiently reverse respiratory suppression from most overdoses involving heroin and morphine; however, these doses are inadequate for potent synthetic opioids [25]. A recent report by Sommerville et al. found that 83% of suspected fentanyl overdoses in Massachusetts from 2014–2016 required more than two doses of Narcan® nasal spray before any resuscitative response was observed [26]. The same study also found that of the fatalities examined, 36% overdosed within minutes after fentanyl ingestion and 90% of these patients were pulseless upon EMS arrival. These findings demonstrate how quickly opioid-induced respiratory suppression progresses to death and is a stirring reminder of the Moscow theater hostage crisis where synthetic opioids proved to be more than non-lethal incapacitating agents. The potency of fentanyl and its derivatives calls into question the proper dose of naloxone that should be administered to rescue patients in the era of synthetic opioids.

Single-use, high potency antidotes that work on a majority of the population are critical in protecting the public from synthetic opioids and mitigating the effects of a domestic chemical attack. In 2016, the FDA approved a new formulation of Evzio® that contains 2mg of naloxone hydrochloride—a dose five times higher than the original device. However, a high potency antidote doesn't necessarily translate to single use. This is especially true for lipophilic opioids that leach out from adipose tissues and agonize the MOR for long periods of time.

To address this concern, researchers at the American Chemical Society National Meeting & Exposition (Spring 2019) announced the development of a next-generation single-dose opioid antidote using naloxone-loaded polylactic acid nanoparticles that slowly release therapeutic concentrations of naloxone over 24 hours [27]. No results are available at this time and it remains to be seen if this technology will pave the way to single-use antidotes against synthetic opioids. In a similar vein, the U.S. Department of Health and Human Services is sponsoring the development of a long-lasting intranasal formulation of naloxone to circumvent the need for multiple doses [28].

Additionally, DHS Science & Technology Directorate (DHS S&T) is working to protect first responders at risk of exposure to harmful levels of fentanils at emergency scenes. DHS S&T partnered with Clear Scientific to create ClearAlert—a wearable sensing device that aims to quickly alert first responders to their level of exposure to fentanils, thus allowing them to rapidly exit the toxic environment and receive medical countermeasures necessary to prevent serious injury [29]. According to DHS S&T, "Experimental results indicated successful and rapid detection of submicrogram quantities of aerosolized carfentanil [29]." Future development of the device will focus on sensor performance and operational life optimization, and field testing and market research will be performed to ensure the final product meets the requirements of the first responder community.

Conclusion

Synthetic opioids represent a unique class of potential chemical weapons with significant medical use. Historical evidence suggests that weaponized opioids are not the preferred agent of choice for radical states or nations; however, Russia's newfound defiance over CWC provisions and previous use of Kolokol-1 makes the future of chemical peace unclear. The Trump administration and members of Congress recognize the weaponization potential of synthetic opioids and have taken steps to potentially classify fentanyl as a weapon of mass destruction [19]. In addition, China's pledge to ban the production of synthetic opioids will significantly limit fentanyl distribution in underground markets and hinder illicit stockpiling by bad actors [20-21].

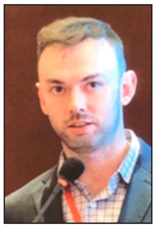
As DHS aims to control the flow of fentanyl overseas, the FDA continues to expedite approvals for novel naloxone formulations to combat the opioid epidemic. New technologies such as user-friendly nasal aspirators and single-use naloxone antidotes circumvent the need for extensive medical training and provide the public with direct access to medical countermeasures that may prove invaluable during a potential chemical attack with synthetic opioids.

References

1. Defense Threat Reduction Agency. (2002). Chemical Weapons Convention. Retrieved from https://www.opcw.org/chemical-weapons-convention/articles/article-ii-definitions-and-criteria
2. Singh, R. K., Kumar, S., Prasad, D. N., & Bhardwaj, T. R. (2018). Therapeutic journery of nitrogen mustard as alkylating anticancer agents: Historic to future perspectives. *European Journal of Medicinal Chemistry, 151*, 401-433. doi:10.1016/j.ejmech.2018.04.001
3. Scholl, L., Seth, P., Kariisa, M., Wilson, N., & Baldwin, G. (2019). Drug and opioid-involved overdose deaths—United States, 2013–2017. *Morbidity and Mortality Weekly Report, 67*(5152), 1419.
4. Centers for Disease Control and Prevention. (2018). *U.S. opioid prescribing rate maps*. Retrieved from https://www.cdc.gov/drugoverdose/maps/rxrate-maps.html
5. Burns, S. M., Cunningham, C. W., & Mercer, S. L. (2018). DARK classics in chemical neuroscience: Fentanyl. *ACS Chemical Neuroscience, 9*(10), 2428-2437. doi:10.1021/acschemneuro.8b00174
6. Pathan, H., & Williams, J. (2012). Basic opioid pharmacology: An update. British journal of pain, 6(1), 11-16. doi:10.1177/2049463712438493
7. Raehal, K. M., & Bohn, L. M. (2014). β-arrestins: Regulatory role and therapeutic potential in opioid and cannabinoid receptor-mediated analgesia. *Handbook of Experimental Pharmacology, 219*, 427-443. doi:10.1007/978-3-642-41199-1_22
8. Boom, M., Niesters, M., Sarton, E., Aarts, L., Smith, T. W., & Dahan, A. (2012). Non-analgesic effects of opioids: Opioid-induced respiratory depression. *Current Pharmaceutical Design, 18*(37), 5994-6004.
9. Pearson, A. (2006). Incapacitating biochemical weapons. *The Nonproliferation Review, 13*(2), 151-188. doi:10.1080/10736700601012029
10. Stanley, T. (2003). Human immobilization: Is the experience in Moscow just the beginning? *European Journal of Anaesthesiology, 20*(6), 427-428.
11. U.S. Drug Enforcement Administration. (2016). *DEA issues carfentanil warning to police and public*. Retrieved from https://www.dea.gov/press-releases/2016/09/22/dea-issues-carfentanil-warning-police-and-public
12. O'Donnell, J., Gladden, R. M., Mattson, C. L., & Kariisa, M. (2018). Notes from the field: Overdose deaths with carfentanil and other fentanyl analogs detected—10 states, July 2016–June 2017. *Morbidity and Mortality Weekly Report, 67*(27), 767.
13. Riches, J. R., Read, R. W., Black, R. M., Cooper, N. J., & Timperley, C. M. (2012). Analysis of clothing and urine from Moscow theatre siege casualties reveals carfentanil and remifentanil use. *Journal of Analytical Toxicology, 36*(9), 647-656. doi:10.1093/jat/bks078
14. Davison, N., & Lewer, N. (2006). *Bradford non-lethal weapons research project (BNLWRP). Research Report No. 8*. Retrieved from University of Bradford, Department of Peace Studies, Centre for Conflict Resolution: http://hdl.handle.net/10454/4000
15. 3rd European Symposium on Non-Lethal Weapons. (2005). Retrieved from http://www.non-lethal-weapons.com/sy03index.html
16. Sanders-Zakre, A. (2019, February). Russia blocks consensus at CWC conference. *Arms Control Today*. Retrieved from https://www.armscontrol.org/act/2019-01/news/russia-blocks-consensus-cwc-conference
17. Edney, A. (2019, January 28). Senator seeks strategy to prevent fentanyl terror attacks. *Bloomberg*. Retrieved from https://www.bloomberg.com/news/articles/2019-01-28/strategy-to-prevent-fentanyl-terror-attacks-sought-by-senator
18. Centers for Disease Control and Prevention. (2018). *Overdose deaths involving synthetic opioids*. Retrieved from https://www.cdc.gov/drugoverdose/data/fentanyl.html
19. Turner, A. (2019, April 16). Homeland security is repordetly considering labeling fentanyl as a weapon of mass destruction. *CNBC*. Retrieved from https://www.cnbc.com/2019/04/16/dhs-considering-labeling-fentanyl-as-weapon-of-mass-destruction-report.html
20. U.S. Drug Enforcement Administration. (2016). *Counterfeit prescription pills containing fentanyls: A global threat*. Retrieved from https://content.govdelivery.com/attachments/USDOJDEA/2016/07/22/file_attachments/590360/fentanyl%2Bpills%2Breport.pdf
21. Myers, S. G., A. (2019, April 1). China bans all types of fentanyl, cutting supply of deadly drug to U.S. and fulfilling pledge to Trump. *The New York Times*. Retrieved from https://www.nytimes.com/2019/04/01/world/asia/china-bans-fentanyl-trump.html
22. Rzasa Lynn, R., & Galinkin, J. L. (2018). Naloxone dosage for opioid reversal: Current evidence and clinical implications. *Therapeutic advances in drug safety, 9*(1), 63-88. doi:10.1177/2042098617744161
23. Robinson, A., & Wermeling, D. P. (2014). Intranasal naloxone administration for treatment of opioid overdose. *American Journal of Health-System Pharmacy, 71*(24), 2129-2135. doi:10.2146/ajhp130798
24. U.S. Food and Drug Administration. (2019). *FDA approves first generic naloxone nasal spray to treat opioid overdose*. Retrieved from https://www.fda.gov/news-events/press-announcements/fda-approves-first-generic-naloxone-nasal-spray-treat-opioid-overdose
25. Moss, R. B., & Carlo, D. J. (2019). Higher doses of naloxone are needed in the synthetic opioid era. *Substance Abuse Treatment, Prevention, and Policy, 14*(1), 6. doi:10.1186/s13011-019-0195-4
26. Somerville, N. J., O'Donnell, J., Gladden, R. M., Zibbell, J. E., Green, T. C., Younkin, M., . . . Walley, A. Y. (2017). Characteristics of fentanyl overdose - Massachusetts, 2014-2016. *MMWR: Morbidity and Mortality Weekly Report, 66*(14), 382-386. doi:10.15585/mmwr.mm6614a2
27. Cottingham, K. (2019). Next-generation single-dose antidotes for opioid overdoses. Retrieved from https://www.acs.org/content/acs/en/pressroom/newsreleases/2019/april/next-generation-single-dose-antidotes-for-opioid-overdoses.html
28. U.S. Department of Health and Human Services. (2018). *HHS sponsors development of intranasal form of long-acting opioid overdose drug*. Retrieved from https://www.hhs.gov/about/news/2018/09/20/hhs-sponsors-development-intranasal-form-long-acting-opioid-overdose-drug.html
29. Department of Homeland Security. (2019) *ClearAlert: Wearable sensing badge for solid aerosol and contact exposure to fentanils*. Retrieved from https://www.dhs.gov/sites/default/files/publications/dhs_clearalert_onepager_final-508.pdf

Kyle E. Giesler, Ph.D.
Postdoctoral Scholar, University of California, Berkeley

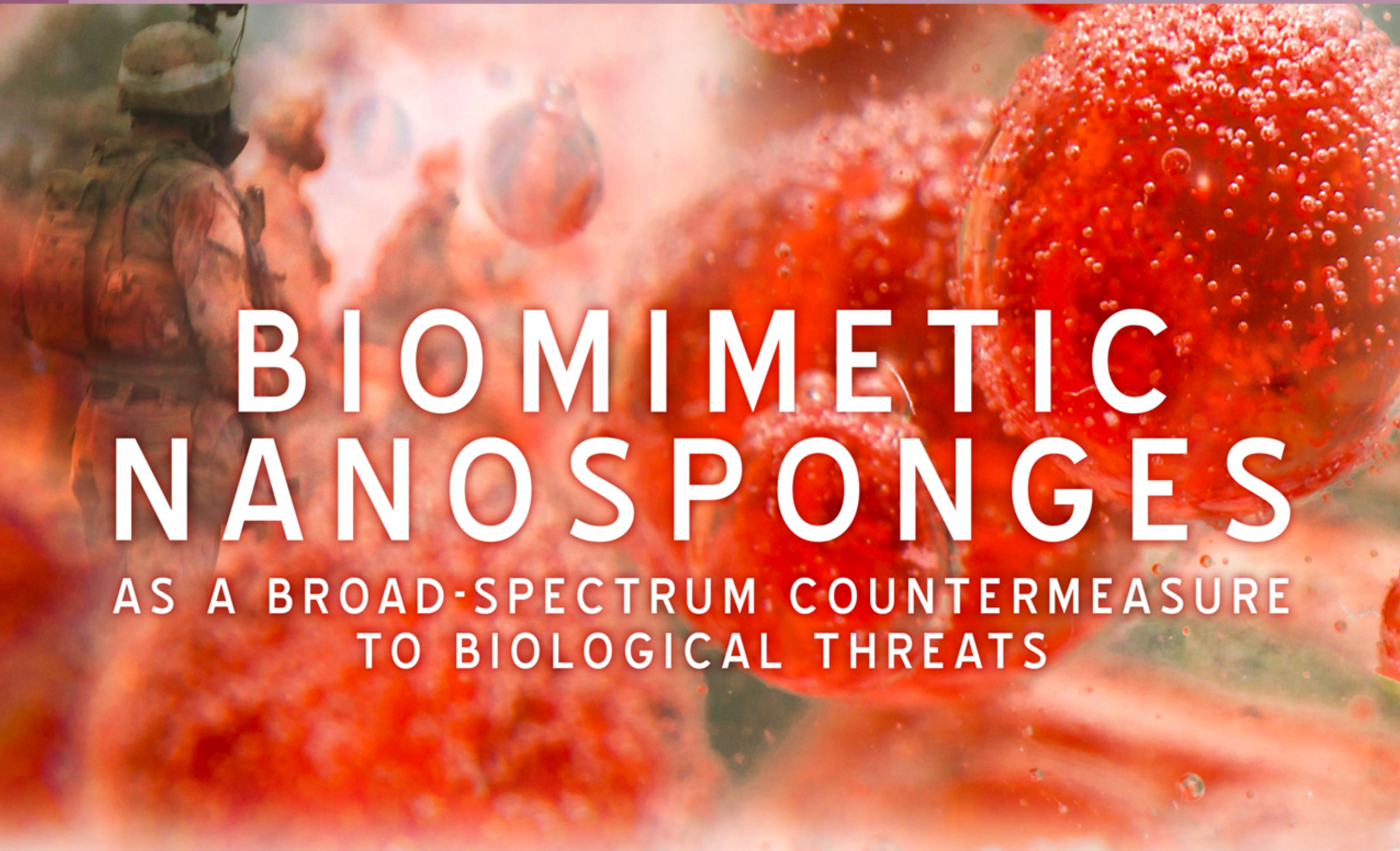
Kyle E. Giesler is a postdoctoral scholar at the University of California, Berkeley investigating non-viral delivery strategies for CRISPR/Cas9 and other emerging biologics. He received his Ph.D. in medicinal chemistry from Emory University under the tutelage of Dennis Liotta where he designed novel small molecules for the treatment of HIV and other chronic viral infections. From 2015 to 2017, he interned with the former Senior Vice President and Chief Patent Counsel at GlaxoSmithKline, Sherry Knowles. His research interests include pharmacology, drug delivery, immunology, virology, stereoelectronic effects, and high-risk drug discovery ventures. He also has a keen interest in medical research relating to military operations, regenerative medicine, medical countermeasures, and CBRN defense strategies.



John M. Saindon, Ph.D.
Senior Health Security & CBRNE/WMD Advisor

John M. Saindon is a senior medical and CBRN/WMD advisor with over 15 years of domestic and international experiences. He received his Ph.D. and a second doctorate degree (DrHSc) in the Health Sciences from Nova Southeastern University. He also has a clinical laboratory specialization in medical technology from George Washington University. Saindon has served in multiple health security and CBRNE/WMD non-proliferation roles while deployed to Africa, Asia, and the Middle East. His research interests are in health security, emerging medical therapies on and off the battlefield, CBRN preparedness, and WMD non-proliferation.

Image Credit | *Photo illustration created by HDIAC adapted from USAF photo by Tech Sgt. Brian Kimball (Available for viewing at http://www.eucom.mil/media-library/article/35920/royal-international-air-tattoo-riat-2017).*



BIOMIMETIC NANOSPONGES

AS A BROAD-SPECTRUM COUNTERMEASURE TO BIOLOGICAL THREATS

Ronnie H. Fang
& Liangfang Zhang

A number of organisms secrete biological toxins that can cause significant harm to the human body. For example, pathogenic bacteria, such as methicillin-resistant *Staphylococcus aureus* (MRSA), utilize a number of different virulence factors to attack their host and enhance colonization [1]. Poisonous animals use their venom, which can contain a mixture of cytolytic and neurotoxic agents, to immobilize prey [2]. Military personnel are potentially at risk of catching a difficult-to-treat infection, being subject to envenomation, or coming under attack from a chemical nerve agent (e.g., sarin, tabun) while deployed—which could significantly compromise health and mission readiness. As such, research and development efforts focused on developing effective and efficient countermeasure therapeutics are of considerable interest to the Department of Defense (DoD).

Most traditional means of toxin neutralization work by a "lock and key" mechanism, whereby a single therapeutic is used to counteract the biological effect of a single target toxin [3]. These can be highly effective in certain scenarios, particularly when the exact nature of the threat is known. However, there are many situations—combat and forward-deployed environments in particular—in which this type of approach is not practical. This includes cases where the threat has not yet been identified, or when there is insufficient time to develop a treatment against an identified threat. And as many organisms secrete multiple toxins at once, some of which microbiologists are still characterizing, the development and fabrication of antidotes and countermeasures can be complicated.

Developing a New Class of Biomimetic Countermeasure Therapeutics

A common feature of most biological toxins, regardless of how they exert their activity, is that they must interact with cells via their plasma membrane. In the case of hemolysins, these toxins dock onto the membrane and physically disrupt integrity or form pores on the surface [4]. Other toxins engage surface receptors and are transported into the cell, whereupon they can affect vital intracellular processes [5]. Taking advantage of these toxin–cell interactions, our research group at the Department of NanoEngineering at the University of California San Diego developed a new class of biomimetic countermeasure therapeutics [6]. "Nanosponges" are generated by coating a layer of cell membrane onto a nanoparticulate substrate. The resulting nanoparticles can then act as nanoscale decoys that are capable of attracting and neutralizing a wide range of toxins, preventing them from destroying healthy cells.

Nanosponges differ from traditional platforms for toxin neutralization by employing a function-driven approach that bypasses the need to identify and characterize individual toxins [7]. Instead of focusing on toxin structure, a greater emphasis is placed on their natural interactions with cells. By faithfully presenting cell membrane material on their surfaces, nanosponges are inherently multi-specific, and a single nanoparticle can protect against a wide range of toxins. This greatly streamlines the development process and makes the platform applicable to a broad array of DoD missions and requirements, where one formulation can be indicated against a number of different CBRN agents, infections, or even venomous injuries. Research into biomimetic toxin-absorbing nanosponges meets

Image Credit | *Photo illustration created by HDIAC and adapted from a U.S. Marine Corps photo (available for viewing at: https://media.defense.gov/2016/Aug/22/2001612174/-1/-1/0/160818-M-LV138-299.JPG) and Adobe Stock.*

the mission of DoD's Chemical and Biological Defense Program (CBDP), which seeks to examine "factors which influence the behavior of chemicals, toxins, and pathogens in relation to the host or target [8]." Future research may also advance biomimetic nanosponge applications to include pre-emptive anti-toxin vaccines for warfighters, to inoculate them in advance of deployment into hostile environments.

Cell Membrane-Coated Nanoparticles

Nanosponges are based on the cell membrane-coated nanoparticle concept (see Figure 1), which is part of a new wave of biomimetic nanotechnologies that have become increasingly popular over the past decade [9, 10]. By taking design inspirations from nature and leveraging natural components, highly functional nanoparticle platforms can be fabricated without having to start from ground zero. Cell membrane-coated nanoparticles are made using a top-down approach that first involves deriving the plasma membrane from cells with a combination of mechanical disruption and differential centrifugation. The purified membrane material is then fused onto the surface of nanoparticulate cores, resulting in a characteristic core–shell structure. The coating of the membrane onto a solid substrate helps with stabilization, preventing the lipid material from fusing with other membranes. At the same time, the core can be used to deliver functional payloads, such as drugs and imaging agents.

Importantly, the coating process has shown to successfully transfer membrane surface markers from the original cells onto the final nanoparticle formulations. For example, CD47 is a marker highly present on red blood cells (RBCs) that acts as a "don't eat me" signal to prevent phagocytic uptake [11]. When RBC membrane is used as a coating

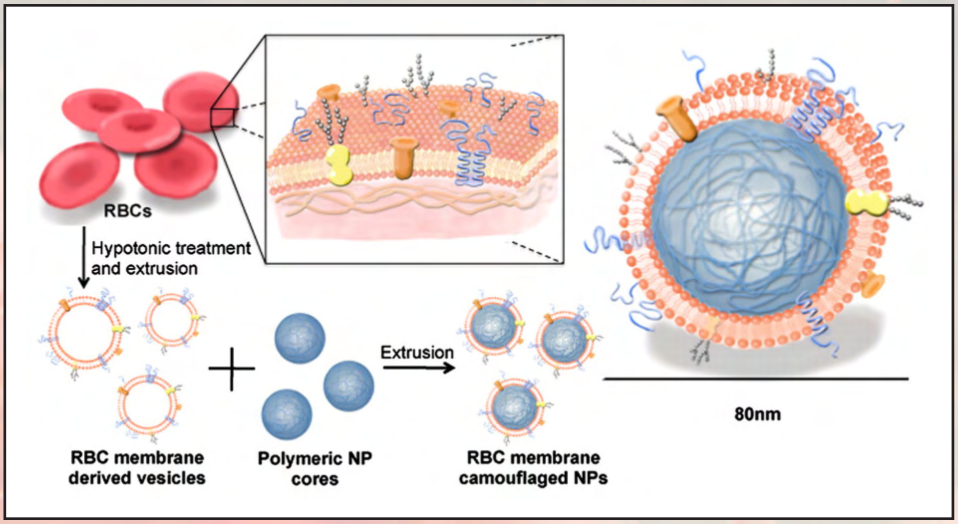


Figure 1. The cell membrane-coated nanoparticle. Cells are employed as a source of membrane material, which contains a wide range of functional ligands. The membrane is then fused onto nanoparticulate cores, resulting in a core–shell, cell-mimicking nanostructure that can be used for a variety of biomedical applications [10]. (Reprinted with Permission)

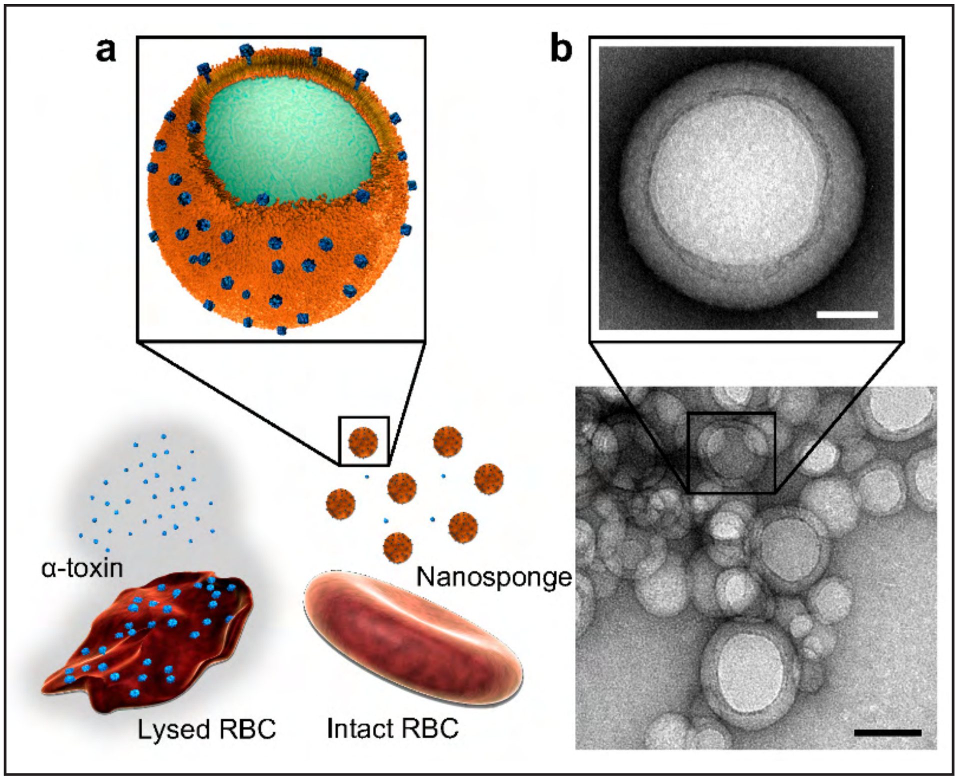


Figure 2. Biomimetic nanosponges for pore-forming toxin neutralization. (a) Under normal circumstances, hemolytic toxins such as staphylococcal α-toxin can attack RBCs and cause lysis. When nanosponges are introduced, these nanodecoys neutralize the toxins and leave the RBCs intact. (b) Transmission electron micrographs reveal the core–shell structure of the nanosponges (top scale bar = 20nm; bottom scale bar = 80nm) [6]. (Reprinted with Permission)

material, the resulting RBC membrane-coated nanoparticles display CD47 in a right-side-out orientation and at the same density as the source cells [12]. Taking advantage of this protein's functionality, the nanoparticles have exhibited reduced uptake by macrophages compared with uncoated particles. Similarly, comprehensive protein analysis of nanoparticles coated with platelet membrane has demonstrated successful functionalization with a panel of platelet surface markers, including those that can be leveraged for immune evasion and targeted delivery [13]. In essence, membrane-coated nanoparticles present themselves as miniaturized versions of the cells from which they are sourced and exhibit many of the same functions as their live counterparts.

As a result of their cell-mimicking properties, membrane-coated nanoparticles have been successfully used for a number of different biomedical applications [9]. This includes the nanodelivery of therapeutic and diagnostic agents, which benefit greatly from the long circulation and enhanced disease site localization afforded by the membrane functionalization. Depending on the source of the membrane material used, formulations capable of targeting cancer, atherosclerotic plaque, and bacteria have all been reported. When loaded with an immunostimulatory agent, it is possible to generate an immune response against the antigenic material on the nanoparticle surface, and this has been leveraged for the design of potent anticancer vaccines [14]. Coating with biological membrane is not limited to nanoparticle substrates however, as nanofibers and planar sensors have also been successfully functionalized [15, 16].

Toxin Neutralization

One unique application of cell membrane-coated nanoparticles is their ability to serve as biomimetic nanosponges for toxin neutralization (see Figure 2) [6]. The nanosponges work by taking advantage of the natural interactions that exist between biological toxins and cellular membrane. This interaction can be nonspecific, as is the case with some small peptide toxins like honeybee melittin, which inserts itself into lipid bilayers and causes membrane disruption [17]. Very often the interaction of toxins with cell membrane is mediated by specific surface receptors [18]. This is the case with *S. aureus* α-hemolysin, which has greater affinity to membranes with a high expression of ADAM10. Streptolysin-O, one of the characteristic toxins for *Streptococcus pyogenes*, uses cholesterol as its receptor; the toxin belongs to a family of cholesterol-dependent cytolysins that all work by similar mechanisms.

Due to differences in surface marker expression, biological toxins are likely to exhibit differential affinities for different cell types. For example, while hemolysins may prefer to attack and lyse RBCs [19], bacterial endotoxins exert their activity through engagement of white blood cells (WBCs) [20]. With the nanosponge technology, this can be easily addressed by fabricating the nanoparticles using membrane derived from cells with the highest affinity for the target of interest. Using next-generation proteomic analysis, it has been demonstrated that nanosponges fabricated using different membrane sources preferentially enrich varying protein subsets from crude bacterial secretions [21].

Not only were the nanosponges able to bind known toxins, but they also demonstrated affinity towards proteins with functions that have yet to be elucidated. A biomimetic "virulomics" approach employing nanosponge enrichment could ultimately be used to identify novel virulence factors, and this may have major implications for the future development of new antibacterial therapies. Research and development into new therapies may provide the warfighter with both pre-deployment antidotes and battlefield countermeasures, increasing force survivability and resilience.

The function-driven approach for toxin neutralization employed by nanosponge technology represents one of its biggest advantages. As suggested by the aforementioned proteomic study, it is not necessary to know the specific identity of a toxin before designing a nanosponge to neutralize it [21]. The only knowledge required is the type of cell most affected by exposure to the toxin. Because this approach does not follow a one-to-one neutralization scheme, it is possible to leverage the numerous receptors found on natural cell membranes to achieve one-to-many multivalency. For example, RBC-based nanosponges have been shown to concurrently bind and neutralize the activity of α-hemolysin, Panton–Valentine leukocidin, and γ-hemolysin from *S. aureus*—all at the same time [22]. With their broad neutralization capabilities, nanosponges can be applied to many types of diseases and conditions, even in some cases where previously no good solution existed.

Detoxification Applications

With the rising incidence of antibiotic resistance, significant efforts have been dedicated to finding novel means of managing bacterial infections. One approach that has shown promise is antivirulence therapy, which seeks to disarm bacteria by neutralizing the toxins used to promote survival and host colonization [23]. Given their ability to neutralize biological toxins broadly, nanosponges have been explored for a number of antibacterial applications. It was demonstrated that, in a murine model of lethal toxin challenge using staphylococcal α-hemolysin, RBC nanosponges can significantly improve survival in both protective and therapeutic scenarios [6]. The benefit of these nanosponges has also been demonstrated using lethal challenge with multi-toxin bacterial secretions [24], as well as during live infection with group A streptococcus (see Figure 3) [25].

In addition to their ability to neutralize bacterial toxins, nanosponges are also adept at binding endogenous proteins with pathological properties. For example, both RBC and platelet-based nanosponges have been used to neutralize autoimmune antibodies in models of autoimmune hemolytic anemia and thrombocytopenia, respectively [26, 27]. WBC nanosponges can be used to bind proinflammatory cytokines and have shown exceptional utility for treating chronic inflammatory conditions such as arthritis [28].

Additionally, with their bacterial toxin-binding capacity, WBC nanosponges can blunt the effect of bacteria-induced sepsis, which is associated with poor outcomes in the clinic and is oftentimes lethal. WBC nanosponges could serve the dual purpose of binding gram-negative endotoxin as well as cytokines, thereby reducing the immune activation that is characteristic of sepsis and enhancing survival in a murine model of the condition [29].

Nanosponges have potential use against pathogens other than bacteria, including parasites and viruses. In one work, it was demonstrated that WBC nanosponges could be used to bind virulence factors from protein secretions derived from the eggs of *Schistosoma mansoni*, one of the most common parasites in humans [21]. More recently, WBC nanosponges derived from CD4-expressing T cells were used to bind and neutralize human immunodeficiency virus (HIV) [30]. It is well-known that HIV evolves in a manner that ultimately outpaces the immune system's capacity to address it, and broadly neutralizing antibodies against the virus are highly sought after. The use of $CD4^+$ nanosponges represents a completely different approach that relies on the natural receptors the virus employs to gain cellular entry, which may ultimately enable the nanoparticles to retain their HIV-binding capabilities over time.

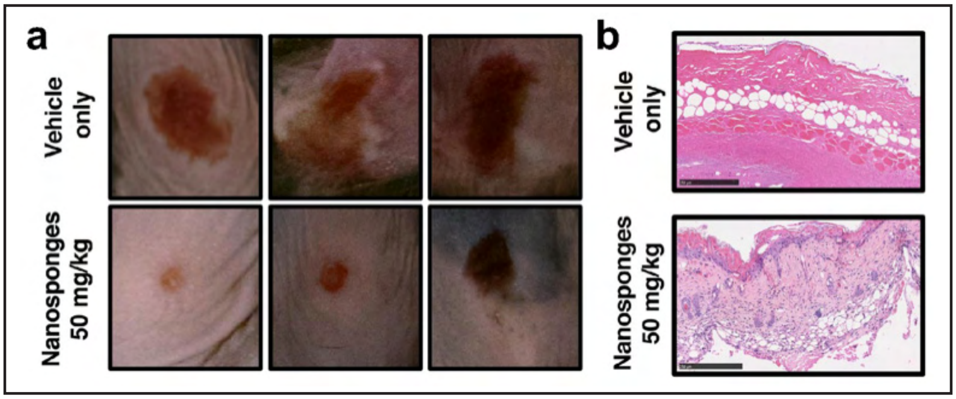


Figure 3. Nanosponges protect against group A streptococcus skin infection. (a) At three days post-infection, mice treated with nanosponges exhibited significantly reduced lesion sizes. (b) Histopathologic analysis of lesion biopsies showed that nanosponge-treated mice had less necrotic tissue injury compared with vehicle-treated mice [25]. (Reprinted with Permission)

Similarly, nanosponges fabricated using mosquito cell membranes have recently been used against Zika virus, greatly improving survival in animal models of infection [31]. This application would be of particular benefit to DoD, since significant numbers of personnel serve in Areas of Responsibility where Zika, Dengue virus, and other arboviruses are common [32].

Although the majority of work using nanosponges has centered on toxins of biological origin, there are also examples of the platform's application to chemical agents. In one example, RBC nanosponges were shown to bind and neutralize dichlorvos, a neurotoxic organophosphate compound that covalently binds and deactivates acetylcholinesterase [33]. Mice treated with the nanosponges were protected from the lethal effects of the compound. Membrane-coated magnetic particles have also been used as tools for affinity-based drug screening and identification [34, 35]. Future research may delineate the efficacy of biomimetic nanosponges against weaponized chemical agent exposure or chemical intoxication.

While the nanosponge platform has mostly been employed for biodetoxification to protect against toxin exposure in therapeutic and prophylactic scenarios, it has also been used in the design of more effective anti-toxin vaccine formulations [36]. Since toxins that are complexed with nanosponges are completely neutralized, they can be delivered back into the body in order to train the immune system against subsequent exposure. This method of using nanoparticle-based toxin detainment for formulating vaccines overcomes many of the challenges associated with traditional toxoid vaccines, which require harsh denaturation techniques that can affect both the immunogenicity and antigenicity of the final formulation. This approach has been used to successfully protect against live MRSA infection, significantly reducing bacterial load in vaccinated mice [37]. The generation of multiantigenic formulations is facile and simply requires coincubation of the nanosponges with bacterial secretions [22]. More recently, active propulsion using micromotor technology has shown to enhance intestinal delivery for eliciting anti-toxin titers in the mucosa [38]. The provision of anti-toxin vaccines to warfighters prior to active-duty service or deployment will inoculate them against toxin exposure, without encumbering them with another medical countermeasure device or medication that must be carried with them into the field.

Conclusion and Outlook

Cell membrane-coated nanosponges are a new class of biomimetic nanoparticle that can be employed in a number of different biodetoxification applications—many of which are of considerable interest to the DoD. Military forces may be supplied with anti-toxin and anti-agent vaccines or injections prior to deployment, or supplied with a broad-spectrum single-point countermeasure available in the event of a suspected CBRN exposure. Future developments in biomimetic nanosponge research may also lead to their use in defending the public and DoD from novel microbiological weapons, a threat posed by rapid advances in bioengineering and synthetic biology [39]. The platform functions by targeting the working mechanism of biological toxins, most of which must interact with cell membrane in some fashion. With this function-driven approach, a single nanosponge can neutralize many toxins at once, including those whose precise function has not yet been elucidated. The technology has major implications for the treatment and prevention of biological threats, particularly some pathogens that have traditionally been hard to manage.

Looking forward, more work can be done along the lines of applying nanosponges towards envenomation or chemical intoxication. Successful translation of the platform will also require efforts to significantly scale-up production to clinically relevant levels. While RBC-based nanosponges have been the most commonly used up to this point, other types of nanosponges, including those based on platelets and WBCs, should excel depending on the specific condition to be treated. Cocktail treatments, or the use of fusion membranes [40], could be considered to further increase the applicability of nanosponges. Overall, future prospects for this biomimetic platform are encouraging, and continued development will enable the realization of its full potential.

References

1. Powers, M. E., & Bubeck Wardenburg, J. (2014). Igniting the fire: *Staphylococcus aureus* virulence factors in the pathogenesis of sepsis. *PLoS Pathogens, 10*(2), e1003871. doi:10.1371/journal.ppat.1003871
2. Andreeva-Kovalevskaya Zh, I., Solonin, A. S., Sineva, E. V., & Ternovsky, V. I. (2008). Pore-forming proteins and adaptation of living organisms to environmental conditions. *Biochemistry (Moscow), 73*(13), 1473–1492. doi:10.1134/S0006297908130087
3. Hua, L., Hilliard, J. J., Shi, Y., Tkaczyk, C., Cheng, L. I., Yu, X., . . . Sellman, B. R. (2014). Assessment of an anti-alpha-toxin monoclonal antibody for prevention and treatment of Staphylococcus aureus-induced pneumonia. *Antimicrobial Agents and Chemotherapy, 58*(2), 1108–1117. doi:10.1128/AAC.02190-13
4. Geny, B., & Popoff, M. R. (2006). Bacterial protein toxins and lipids: Pore formation or toxin entry into cells. *Biology of the Cell, 98*(11), 667–678. doi:10.1042/Bc20050082
5. Lord, J. M., Smith, D. C., & Roberts, L. M. (1999). Toxin entry: How bacterial proteins get into mammalian cells. *Cellular Microbiology, 1*(2), 85–91. doi:10.1046/j.1462-5822.1999.00015.x
6. Hu, C. M., Fang, R. H., Copp, J., Luk, B. T., & Zhang, L. (2013). A biomimetic nanosponge that absorbs pore-forming toxins. *Nature Nanotechnology, 8*(5), 336–340. doi:10.1038/nnano.2013.54
7. Fang, R. H., Luk, B. T., Hu, C. M., & Zhang, L. (2015). Engineered nanoparticles mimicking cell membranes for toxin neutralization. *Advanced Drug Delivery Reviews, 90*, 69–80. doi:10.1016/j.addr.2015.04.001
8. Department of Defense. (2019, March). *Department of Defense Fiscal Year (FY) 2020 Budget Estimates, Chemical and Biological Defense Program, Defense-Wide Justification Book Volume 4 of 5*. Retrieved from https://comptroller.defense.gov/Portals/45/Documents/defbudget/fy2020/budget_justification/pdfs/03_RDT_and_E/RDTE_Vol4_CBDP_RDTE_PB20_Justification_Book.pdf
9. Fang, R. H., Kroll, A. V., Gao, W., & Zhang, L. (2018). Cell membrane coating nanotechnology. *Advanced Materials, 30*(23), 1706759. doi:10.1002/adma.201706759
10. Hu, C. M., Zhang, L., Aryal, S., Cheung, C., Fang, R. H., & Zhang, L. (2011). Erythrocyte membrane-camouflaged polymeric nanoparticles as a biomimetic delivery platform. *Proceedings of the National Academy of Sciences of the United States of America, 108*(27), 10980–10985. doi:10.1073/pnas.1106634108
11. Oldenborg, P. A., Zheleznyak, A., Fang, Y. F., Lagenaur, C. F., Gresham, H. D., & Lindberg, F. P. (2000). Role of CD47 as a marker of self on red blood cells. *Science, 288*(5473), 2051–2054. doi:10.1126/science.288.5473.2051
12. Hu, C. M., Fang, R. H., Luk, B. T., Chen, K. N., Carpenter, C., Gao, W., . . . Zhang, L. (2013). 'Marker-of-self' functionalization of nanoscale particles through a top-down cellular membrane coating approach. *Nanoscale, 5*(7), 2664–2668. doi:10.1039/c3nr00015j
13. Hu, C. M., Fang, R. H., Wang, K. C., Luk, B. T., Thamphiwatana, S., Dehaini, D., . . . Zhang, L. (2015). Nanoparticle biointerfacing by platelet membrane cloaking. *Nature, 526*(7571), 118–121. doi:10.1038/nature15373
14. Kroll, A. V., Jiang, Y., Zhou, J., Holay, M., Fang, R. H., & Zhang, L. (2019). Biomimetic nanoparticle vaccines for cancer therapy. *Advanced Biosystems, 3*(1), 1800219. doi:10.1002/adbi.201800219
15. Chen, W., Zhang, Q., Luk, B. T., Fang, R. H., Liu, Y., Gao, W., & Zhang, L. (2016). Coating nanofiber scaffolds with beta cell membrane to promote cell proliferation and function. *Nanoscale, 8*(19), 10364–10370. doi:10.1039/c6nr00535g
16. Gong, H., Chen, F., Huang, Z., Gu, Y., Zhang, Q., Chen, Y., . . . Zhang, L. (2019). Biomembrane-modified field effect transistors for sensitive and quantitative detection of biological toxins and pathogens. *ACS Nano, 13*(3), 3714–3722. doi:10.1021/acsnano.9b00911
17. van den Bogaart, G., Guzman, J. V., Mika, J. T., & Poolman, B. (2008). On the mechanism of pore formation by melittin. *Journal of Biological Chemistry, 283*(49), 33854–33857. doi:10.1074/jbc.M805171200
18. Los, F. C., Randis, T. M., Aroian, R. V., & Ratner, A. J. (2013). Role of pore-forming toxins in bacterial infectious diseases. *Microbiology and Molecular Biology Reviews, 77*(2), 173–207. doi:10.1128/MMBR.00052-12
19. McCullough, J. (2014). RBCs as targets of infection. *Hematology-American Society of Hematology Education Program, 2014*(1), 404–409. doi:10.1182/asheducation-2014.1.404
20. Galanos, C., & Freudenberg, M. A. (1993). Bacterial endotoxins: Biological properties and mechanisms of action. *Mediators of Inflammation, 2*(7), S11–16. doi:10.1155/S0962935193000687
21. Lapek, J. D., Jr., Fang, R. H., Wei, X., Li, P., Wang, B., Zhang, L., & Gonzalez, D. J. (2017). Biomimetic virulomics for capture and identification of cell-type specific effector proteins. *ACS Nano, 11*(12), 11831–11838. doi:10.1021/acsnano.7b02650
22. Wei, X., Gao, J., Wang, F., Ying, M., Angsantikul, P., Kroll, A. V., . . . Zhang, L. (2017). *In situ* capture of bacterial toxins for antivirulence vaccination. *Advanced Materials, 29*(33), 1701644. doi:10.1002/adma.201701644
23. Dickey, S. W., Cheung, G. Y. C., & Otto, M. (2017). Different drugs for bad bugs: Antivirulence strategies in the age of antibiotic resistance. *Nature Reviews Drug Discovery, 16*(7), 457–471. doi:10.1038/nrd.2017.23
24. Chen, Y., Zhang, Y., Chen, M., Zhuang, J., Fang, R. H., Gao, W., & Zhang, L. (2019). Biomimetic nanosponges suppress *in vivo* lethality induced by the whole secreted proteins of pathogenic bacteria. *Small, 15*(6), 1804994. doi:10.1002/smll.201804994
25. Escajadillo, T., Olson, J., Luk, B. T., Zhang, L., & Nizet, V. (2017). A red blood cell membrane-camouflaged nanoparticle counteracts streptolysin O-mediated virulence phenotypes of invasive group A streptococcus. *Frontiers in Pharmacology, 8*, 477. doi:10.3389/fphar.2017.00477
26. Wei, X., Gao, J., Fang, R. H., Luk, B. T., Kroll, A. V., Dehaini, D., . . . Zhang, L. (2016). Nanoparticles camouflaged in platelet membrane coating as an antibody decoy for the treatment of immune thrombocytopenia. *Biomaterials, 111*, 116–123. doi:10.1016/j.biomaterials.2016.10.003
27. Copp, J. A., Fang, R. H., Luk, B. T., Hu, C. M. J., Gao, W., Zhang, K., & Zhang, L. (2014). Clearance of pathological antibodies using biomimetic nanoparticles. *Proceedings of the National Academy of Sciences of the United States of America, 111*(37), 13481-13486. doi:10.1073/pnas.1412420111
28. Zhang, Q., Dehaini, D., Zhang, Y., Zhou, J., Chen, X., Zhang, L., . . . Zhang, L. (2018). Neutrophil membrane-coated nanoparticles inhibit synovial inflammation and alleviate joint damage in inflammatory arthritis. *Nature Nanotechnology, 13*(12), 1182–1190. doi:10.1038/s41565-018-0254-4
29. Thamphiwatana, S., Angsantikul, P., Escajadillo, T., Zhang, Q. Z., Olson, J., Luk, B. T., . . . Zhang, L. (2017). Macrophage-like nanoparticles concurrently

absorbing endotoxins and proinflammatory cytokines for sepsis management. *Proceedings of the National Academy of Sciences of the United States of America, 114*(43), 11488–11493. doi:10.1073/pnas.1714267114
30. Wei, X., Zhang, G., Ran, D., Krishnan, N., Fang, R. H., Gao, W., . . . Zhang, L. (2018). T-cell-mimicking nanoparticles can neutralize HIV infectivity. *Advanced Materials, 30*(45), 1802233. doi:10.1002/adma.201802233
31. Rao, L., Wang, W., Meng, Q. F., Tian, M., Cai, B., Wang, Y., . . . Wu, J. (2018). A biomimetic nanodecoy traps Zika virus to prevent viral infection and fetal microcephaly development. *Nano Letters, 19*(4), 2215–2222. doi:10.1021/acs.nanolett.8b03913
32. Garg, H., & Joshi, A. (2019). Multivalent, virus-like particle vaccine and reporter virus assay for arboviruses. *Journal of the Homeland Defense & Security Information Analysis Center, 6*(1), 32–37. Retrieved from https://www.hdiac.org/wp-content/uploads/Multivalent-Virus-Like-Particle-Vaccine-and-Reporter-Virus-Assay-for-Arboviruses_V6I1.pdf
33. Pang, Z. Q., Hu, C. M. J., Fang, R. H., Luk, B. T., Gao, W., Wang, F., . . . Zhang, L. (2015). Detoxification of organophosphate poisoning using nanoparticle bioscavengers. *ACS Nano, 9*(6), 6450–6458. doi:10.1021/acsnano.5b02132
34. Bu, Y. S., Hu, Q., Ke, R. F., Sui, Y., Xie, X. Y., & Wang, S. C. (2018). Cell membrane camouflaged magnetic nanoparticles as a biomimetic drug discovery platform. *Chemical Communications, 54*(95), 13427–13430. doi:10.1039/c8cc08530g
35. Sherwood, J., Sowell, J., Beyer, N., Irvin, J., Stephen, C., Antone, A. J., . . . Ciesla, L. M. (2019). Cell-membrane coated iron oxide nanoparticles for isolation and specific identification of drug leads from complex matrices. *Nanoscale, 11*(13), 6352–6359. doi:10.1039/c9nr01292c
36. Angsantikul, P., Fang, R. H., & Zhang, L. (2018). Toxoid vaccination against bacterial infection using cell membrane-coated nanoparticles. *Bioconjugate Chemistry, 29*(3), 604–612. doi:10.1021/acs.bioconjchem.7b00692
37. Wang, F., Fang, R. H., Luk, B. T., Hu, C. J., Thamphiwatana, S., Dehaini, D., . . . Zhang, L. (2016). Nanoparticle-based antivirulence vaccine for the management of methicillin-resistant *Staphylococcus aureus* skin infection. *Advanced Functional Materials, 26*(10), 1628–1635. doi:10.1002/adfm.201505231
38. Wei, X., Beltran-Gastelum, M., Karshalev, E., Esteban-Fernandez de Avila, B., Zhou, J., Ran, D., . . . Zhang, L. (2019). Biomimetic micromotor enables active delivery of antigens for oral vaccination. *Nano Letters, 19*(3), 1914–1921. doi:10.1021/acs.nanolett.8b05051
39. Geisler, K. E., Elliott, T. C., & Saindon, J. M. (2019, January). Microbiological weapons: A potential consequence of CRISPR/Cas9 technologies. Homeland Defense & Security Information Analysis Center. Retrieved from https://www.hdiac.org/wp-content/uploads/HDIAC-Spotlight_Microbiological-Weapons-A-Potential-Consequence-of-CRISPR-Cas9-Technologies_January-2019-1.pdf
40. Dehaini, D., Wei, X., Fang, R. H., Masson, S., Angsantikul, P., Luk, B. T., . . . Zhang, L. (2017). Erythrocyte-platelet hybrid membrane coating for enhanced nanoparticle functionalization. *Advanced Materials, 29*(16), 1606209. doi:10.1002/adma.201606209

Ronnie H. Fang, Ph.D.
Assistant Project Scientist, University of California San Diego

Ronnie H. Fang is an assistant project scientist in the Department of NanoEngineering at the University of California San Diego. He received his Ph.D. in nanoengineering at the University of California San Diego. His research is focused on leveraging biomimetic nanoparticles for drug delivery and immunoengineering applications.

Liangfang Zhang, Ph.D.
Professor, University of California San Diego

Liangfang Zhang is a professor in the Department of NanoEngineering and Moores Cancer Center at the University of California San Diego. He received his Ph.D. in chemical and biomolecular engineering at the University of Illinois at Urbana-Champaign. His research aims to create cutting-edge biomimetic nanotechnologies and exploit them for various biomedical applications with a particular focus on drug delivery, biodetoxification, and vaccination.

CONFIDENCE AND TRUST IN HUMAN-MACHINE TEAMING

Aaron Celaya
& Nick Yeung

Following World War II and into the Cold War, the United States began utilizing offset strategies to gain strategic and operational advantage over adversaries. These strategies focused on technological advance as key to creating disparities in military and coercive power between opposing forces. Former U.S. Secretary of Defense Harold Brown explained this strategic approach to Congress in 1981 when he stated, “Technology can be a force multiplier, a resource that can be used to help offset numerical advantages of an adversary. Superior technology is one very effective way to balance military capabilities other than matching an adversary tank-for-tank or soldier-for-soldier [1].” Thus, the U.S.

Image Credit *Photo illustration created by HDIAC and adapted from Adobe Stock images and a DIA photo (available for viewing at: https://www.dia.mil/portals/27/Images/Phase%20I%20Redesign%20Files/Career%20Field%20Images/IT/170603-F-LW859-002.jpg?ver=2018-09-16-213142-767).*

developed tactical nuclear warheads and precision-guided weapons as key components of the first and second offset strategies. A similar emphasis on technological advance is evident in the current National Defense Strategy, released in January 2018, which anticipates further development of traditional warfighting technologies (nuclear forces, missile defense, forward forces), but crucially goes beyond these to also include Artificial Intelligence (AI) and other forms of algorithmic technology implementations [2]. Correspondingly, the Department of Defense (DoD) is calling for AI development, employment, and deployment to support decision making, intelligence operations, and additional capabilities via its new Artificial Intelligence Strategy, released in February 2019 [3].

There are many varieties of artificial systems used daily by companies, governments, universities, and individual citizens. Artificial systems can range from fixed, rules-based automation to evolutionary, judgement-based systems. Fixed automation consists of pre-programmed, strictly controlled and contained functions, such as scripts, macros, or robotic-process automation. By contrast, evolutionary, judgement-based algorithms are self-learning, autonomous, and potentially unbounded AI capabilities that typically rely on deep learning and statistical prediction, and can exhibit behaviors not programmed or even anticipated by their human designers [4, 5].

Much of the focus and interest in artificial systems is in increasing their capabilities: optimizing algorithms, increasing data quality, and expanding the domains to which they are applied. Recent years have seen striking progress in AI systems—from algorithms capable of teaching themselves superhuman chess-playing ability in a matter of hours [6]; to predictive maintenance via machine learning with military vehicles, allowing AI to flag failing vehicle parts before they break down in hostile territory [7]. However, our focus here is on an aspect of artificial system development that has gained less attention: the question of how people interact with these emerging technologies.

This issue is crucial for at least two reasons. First, in some areas it is agreed that there must be meaningful human control in place—effectively placing a limit on autonomy—and the respective roles of human and artificial systems must be considered. Second, and more broadly, artificial system applications (especially for AI) remain narrow and specialized, such that situations involving complex, ill-defined, strategic problems are likely to depend on human input for the foreseeable future. Thus, in the coming years, effective artificial systems deployment will not only depend on the quality of those systems, but also human operators' ability to work effectively with those artificial systems. The importance of human-machine teaming (HMT) is explicitly recognized in the emphasis within the DoD's AI strategy on "human-centered AI," and has been discussed in depth elsewhere in military strategy [8].

Some of the work to be done in the HMT domain will be in the vein of well-established human factors principles [9], namely, designing systems that are sensitive to human capabilities: people's unparalleled ability to communicate naturally with others, solve complex and ill-defined problems, and understand their world using a combination of domain expertise and common sense; especially versus striking limitations in human operators' cognitive capacity and ability to sustain attention, and in their susceptibility to cognitive bias. Understanding the relative strengths and limitations of human vs. artificial systems will remain critical to effective HMT in the coming years. Beyond this, however, we expect artificial systems to increasingly act as team members working alongside human operators, rather than as tools used by those operators—providing specialist skills and abilities that complement those provided by people.

As such, effective HMT will depend on factors that govern effective human teams—namely, trust and effective communication. Decades of research in psychology have documented the principles of trust, and there is a growing body of work applying these principles to understand what is shared and what is distinctive about human-machine trust. Here we briefly review this work and discuss future directions for this emerging field.

Human Trust

Current research regarding HMT is predicated on the assumption that human trust of an artificial system in any form is guided by factors that determine trust in normal human-human interactions [10]. Trust has been defined in a highly influential theory as "the willingness of a party to be vulnerable to the actions of another party based on the expectation that the other will perform a particular action important to the trustor, *irrespective of the ability to monitor or control that other party*" (emphasis added) [11]. As automation and AI increase in capability and complexity, human operators will find their abilities to monitor or possibly even control artificial systems lessened. Therefore, understanding the components of human-machine trust and factors that influence the user's desire to trust will be vital in designing and fielding future artificial systems.

To analyze the components of human-machine trust, we adapt the two-party trust model developed by Mayer et al. [11]. Their model consists of a trusting party (trustor) and a party to be trusted (trustee). In our adaptation, the trustor is our human operator and the trustee is the artificial system, whether it be automation or AI-based (see Figure 1). The three perceived trustworthiness factors of the trustee are their Ability, Benevolence, and Integrity (ABI). The trustee's Ability to perform its given task

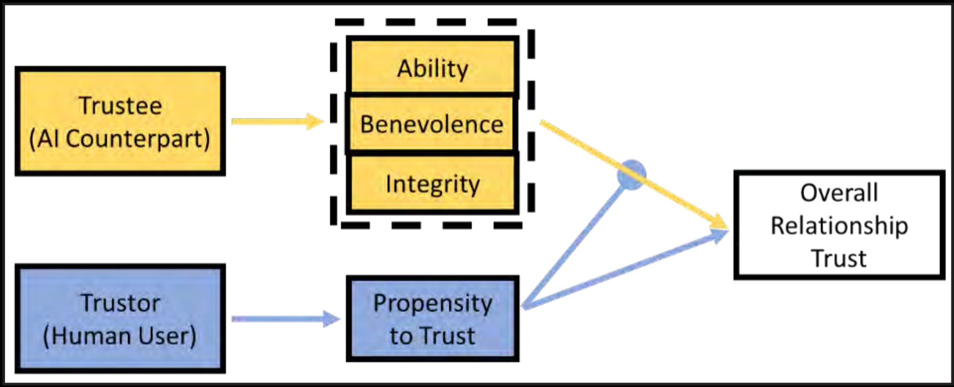


Figure 1. Trust factors, adapted from Mayer, Davis, & Schoorman [11]

well, as we will discuss later, will remain one of the most important factors in determining human trust. Benevolence and Integrity refer to the trustor's belief that the trustee wants to act in the trustor's interests and adheres to a set of principles that the trustor finds acceptable.

Meanwhile, the trustor is largely characterized by their general willingness to trust others and the relative weight they place on ability, benevolence, and integrity as key factors in winning and maintaining their trust. This *propensity to trust* is considered a stable characteristic of the trustor and, importantly, mediates the relationship between the trustee's ABI and overall relationship trust. Thus, the model emphasizes that individual characteristics and personality traits have a heavy influence in the HMT relationship.

Hoff and Bashir have developed a related model that applies specifically to the case of human trust in artificial systems [12]. Their model simplifies the critical determinants of trust to two factors—the ability of the system as experienced by the user ("learned trust," depicted in Figure 2) and the user's propensity to trust. They unpack the user's propensity into two components: dispositional trust and situational trust.

Dispositional trust is what the user brings to the interaction with automation, reflecting his or her culture, age, gender, personality, etc. As such, it varies diversely from person to person, rendering it a difficult concept to isolate, study, and effectively implement into future artificial systems. However, trends and correlations have begun to emerge in the past decade in ongoing research regarding characteristics affecting human-machine trust. This is especially true with the advent of specific assessment tools which measure human personality traits in relation to various types of artificial systems [13–16].

Developers and users of technologies should continue to support research that aims to uncover the trends that drive individual differences in the use of artificial systems. Further, in a military context, commanders should ensure that operators are trained to appropriately rely on their algorithmic counterparts, with trust that scales with the utility and flexibility of these systems—never blindly accepting and complying with their outputs. Future AI should have the ability to adapt to its user, much like computer systems today have user preferences. However, in the case of AI, user preferences will be determined from machine learning that works to maximize correct and confident output by its individual user. This will be driven by the aforementioned research efforts and could manifest in the form of individualized appearance, communication mode and style, etc., as well as advanced features such as information about its own reliability (confidence), as discussed below.

Inclusion of *situational trust* in Hoff and Bashir's model emphasises that people's reliance on automation will vary according to context. External factors, such as task difficulty, workload, perceived risks and benefits, organizational setting, and framing of the task can all affect trust in artificial systems. For example, Ross conducted an experiment that utilized decision aids in a video-based search-and-rescue task [17]. In this study, human users were able to adjust their reliance and trust on a single automated decision aid based on that aid's actual ability. However, appropriate reliability of a single decision aid was miscalculated when an additional decision aid was present and produced a different outcome (i.e., the decision aids had mixed reliability levels). In Ross' experiment, outcome variation due to external factors (i.e., having another decision aid present) resulted in a bias which significantly impacted human trust.

Trust in automation also depends on variable factors of the user, such as their subject matter expertise, mood, and attentional capability. For example, De Vries, Midden, and Bouwhuis experimentally observed that human participants who were more self-confident in a mapping task used automation less throughout the study even though their workload would have been lessened [18].

Hoff and Bashir's model provides a similarly detailed analysis of how trust reflects the perceived ability of an artificial system. Learned trust will depend partly on a user's initial pre-conceptions: their expectations, the reputation of the system and/or brand, prior experience with a similar technology, and their understanding of the system. For example, Bliss, Dunn, and Fuller assessed the impact of "hearsay" on the reliability of an automated agent prior to interaction [19]. This information inflated the perceived accuracy of the automated aid, which subsequently increased participants' initial reliance on the aid over that of a control group.

Once established, trust then changes dynamically over the course of system experience such that, through repeated interactions, a new reliance strategy is formed by the user. These interactions depend critically on the observed performance of the system [20], but human operators are far from perfect in evaluating system performance even when clear feedback is provided [21]. Moreover, in the

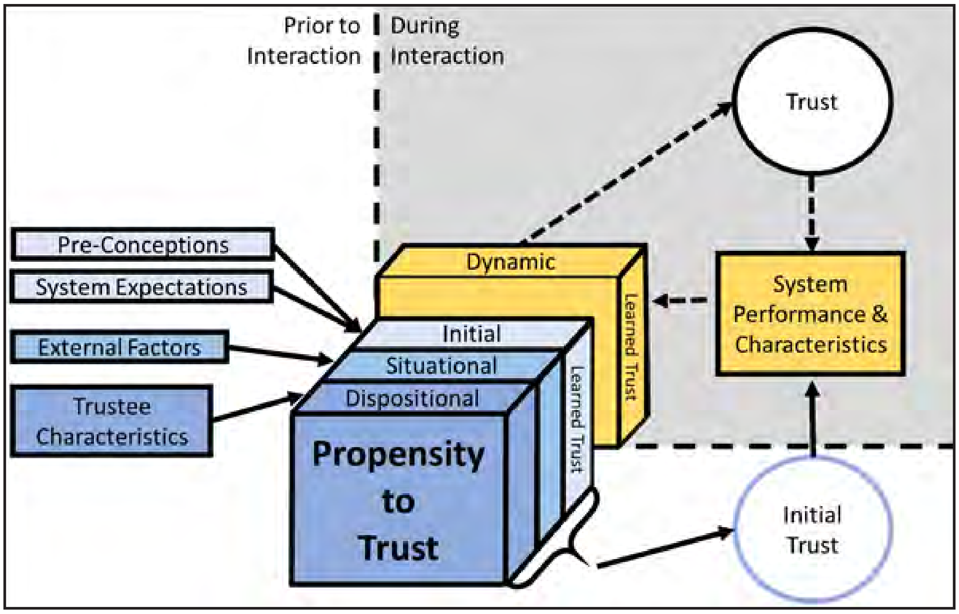


Figure 2. Integrated, three-layer trust model, adapted from Hoff & Bashir [12]

absence of feedback, human users have been shown to weigh advice differently based on if it agrees with them or not, regardless of accuracy [22].

In addition to system performance, trust also depends on usability features of the system, including its appearance, ease-of-use, communication style, transparency, and the operator's level of control. Dynamic learned trust has been demonstrated experimentally. For example, Metzger & Parasuraman conducted an experiment to see how experienced air traffic controllers responded to an automated aircraft collision warning system [23]. In their study, operators were first exposed to a warning system with perfect reliability, then the automated aid became less reliable. This change in reliability had a significantly negative effect on the user's trust of the automated aid.

Together, these models of trust highlight several conclusions relevant for human-machine teaming. First, users' reliance on artificial systems varies according to their trust in these systems: Technological developments will only provide military advantage to the extent that they are trusted by human operators. Second, trust depends not only on objective system performance, but also on the complex psychology of trust that crucially includes the users' propensity to trust according to their personality, prior experience, expectations, and their current situation. Third, users will form an initial reliance strategy that is created even before any experience with a particular system—biasing their trust even before any system interaction takes place. Thus, identifying specific aspects of dispositional, situational, and learned trust is of great importance to understanding how best to design and deploy new artificial systems in the service of effective HMT.

Trust in Human-Machine Teaming

Our recent research explores the themes identified above in the context of an AI decision aid for visual judgment tasks that are designed to parallel operationally-relevant tasks (e.g., is there a military installation in this satellite image?; is there suspicious activity in this radar map?). For example, while system performance remains the most important factor in determining human-machine trust, we observe large individual differences in actual use and reliance for artificial systems. We conducted an experiment where participants were required to complete a task with the assistance of either a human or a computer advisor. Even though human and computer advice was equally accurate and presented in the same format, we observed large individual differences in advisor influence and choice.

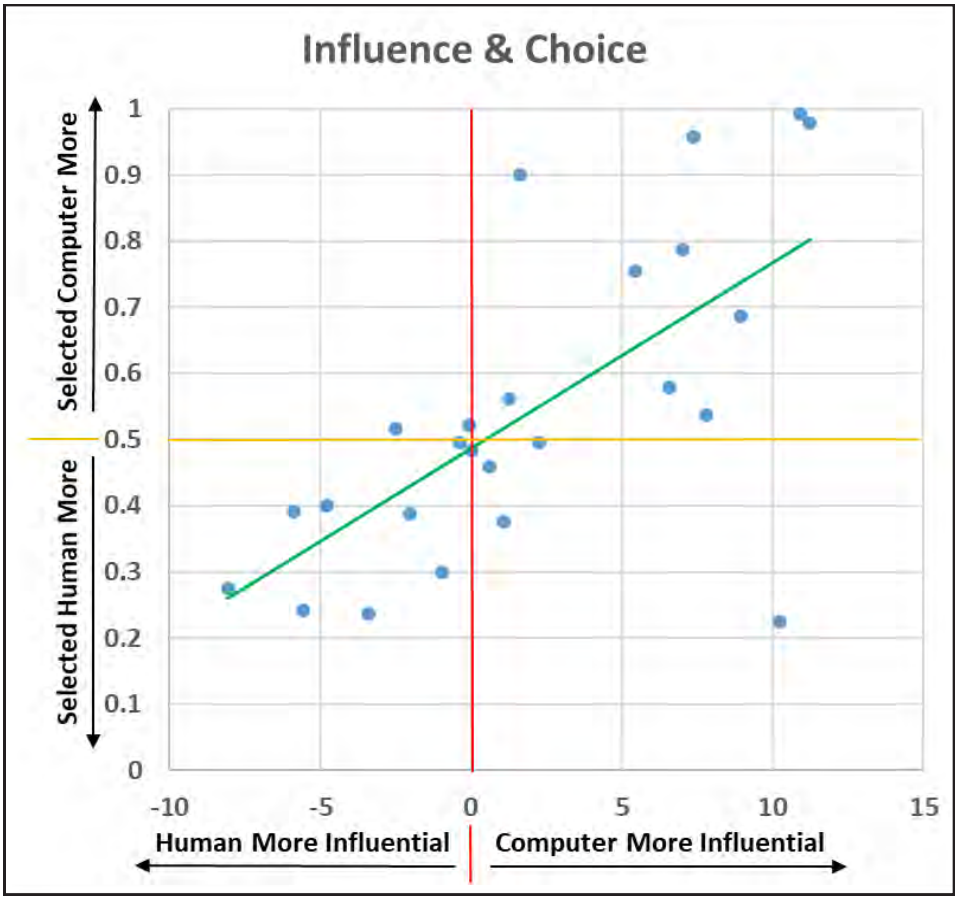


Figure 3. Plot of individual participants depicting advisor influence and choice. There is a strong correlation between an individual user's preference for one advisor over the other (y-axis values) and the relative influence of the two advisors' advice (x-axis). However, across individual users, there is substantial variability across users in whether this preference is for human or algorithmic advice (evident as wide scatter in datapoints across both axes).

Figure 3 depicts one of the significant findings from our study, showing variation across individuals (blue circles) according to their relative preference for choosing computer vs. human advice (y-axis) and being more influenced in their decision making by computer vs. human advice when provided (x-axis). People showed a strong tendency to choose the advisor they were more influenced by (datapoints generally fall on the green regression trend line), as one might expect, but there was a wide and varied difference regarding which advisor was more influential and preferred: Some participants overwhelmingly chose computer advice (e.g., with four people choosing computer advice more than 90% of the time), but others showed consistent aversion to the artificial system (with four others choosing it less than 30% of the time).

This finding brings to the forefront the importance of individual differences when it comes to propensity to trust. Singh, Molloy, & Parasuraman conducted ground-breaking research in identifying characteristics that could predict future automation-use traits [24]. They developed a psychometric measurement tool, the Complacency-Potential Rating Scale, which measures a person's individualized automation-induced complacency. Figure 3 shows that, all things being equal, people remain diverse in their preferences and usage strategies for artificial systems. A current focus of our research is on uncovering the traits that drive these individual differences when it comes to artificial systems employment—a potentially critical factor in the effectiveness of emerging technologies. As Paul Scharre aptly stated in 2014, "the winner…will not be who develops [the] technology first or even who has the best technology, but who figures out how to best use it [25]."

A second strand of our current research focuses on effective communication in decision making involving human-machine teams. This work focuses in particular on communication about uncertainty and confidence. Complex tasks are typically characterized by uncertainty—in the information provided, its provenance and reliability, the courses of action available, etc.

When addressing these problems, it is crucial to appropriately deal with uncertainty. And when it comes to group decision making, team members need to communicate and integrate their uncertainty—or its converse, their confidence. As we have known for over 100 years in psychology, but are just beginning to understand the significance of, people's confidence judgements about their own decisions convey very useful information about the reliability of those decisions [26]. People seem exquisitely sensitive to the information carried by statements of confidence.Consistent with this, evidence indicates that communication of confidence critically underpins human-human trust in decision making. Confidently expressed opinions have more influence [27] but trust is lost if this confidence proves to be unfounded [28]. More confident decision makers are less receptive to advice [29] and down-weight dissenting opinions [22].

Also, teams can make optimal decisions that outperform the best individual decision makers only if team members communicate confidence effectively [30]. Bahrami and colleagues used simple visual judgment tasks in their work to research which human-human communication model produced an optimal joint decision [30]. They showed that communication about confidence (which they defined as a person's subjective estimate of their probability of being correct) can improve team decision making, with the group significantly outperforming the best individual within the team.

In human-machine teams, communication is predominately provided only in terms of the answer or advice from the artificial system to the human user. HMT may remain critically limited if confidence cues are ignored or lacking. Specifically, the absence of artificial system confidence cues may give rise to overuse or misuse of the system through at least two distinct mechanisms: perceived overconfidence (of the artificial system) and attributed sources of uncertainty. For overconfidence, human advice is discounted when confidence is poorly calibrated (i.e., it correlates weakly with objective accuracy) [28, 27]. Therefore, if an artificial system teammate is perceived as being confident but incorrect, this could have a large negative impact on human user trust and reliance. Moreover, even though people are generally more influenced by advice that is given confidently (as opposed to unconfidently), they are not averse to uncertain advice. In particular, people show greater trust when advice acknowledges the uncertainty that is inherent in many complex decision making contexts [31, 32].

Both the communication of confidence itself and the manner in which it is expressed matter. In another of our experiments involving visual judgments with AI decision system support, we assessed preferences between two different computer algorithms. The advisors gave equally accurate advice, but one also communicated its confidence (e.g., a recommendation with 75% confidence) while the other did not (simply stating its answer). In Figure 4, the impact of communicated confidence is clear when it comes to advisor preference. Participants chose the advisor that gave them confidence information significantly more than the advisor that did not provide that information.

This simple study illustrates human user's sensitivity to subtle aspects of communication in HMT. Given that uncertainty is inherent in many complex task domains, implementing confidence communication capabilities into artificial systems may drastically impact the team building capability of human-machine teams. More broadly, these findings illustrate that it is crucial to understand how people interact with each other in effective teams (and indeed in ineffective teams), because their behavior in HMT contexts will reflect these established patterns of interaction.

Conclusion

Advances in AI promise to have a transformational impact on military strategy and capability in the coming years. However,

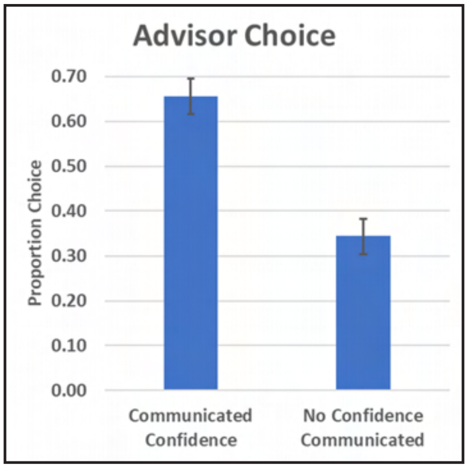


Figure 4. Study participants favored advisors that communicated confidence over advisors that did not.

fulfilling this promise depends on an understanding of human psychology to optimize human users' trust and reliance on AI systems. Thus, governments, militaries, or other organizations who employ these technologies would do well to ensure that artificial systems are developed in such a way that designing for *appropriate* user trust and reliance is considered in the original design and initial implementation of the new technology.

Effective HMT means going beyond simply focusing on ensuring that artificial systems have the best performance possible. System design should also consider that new technologies must incorporate principles that support effective interactions among human team members—particularly as the complexity of AI systems increases such that they come to take on more complex roles and tasks that historically would have been the preserve of human team members. In the context of strategic decision making, this means effective communication of confidence and uncertainty. Further, producers and requestors of these technologies should also lead the way in encouraging continuing research in the diverse domain of human trust.

Authors' note: *This work was partially funded by a grant from the U.S. Air Force Office of Scientific Research, European Office of Aerospace Research and Development. The views expressed are those of the authors and do not necessarily reflect the official policy or position of the Air Force, the DoD, or the U.S. Government.*

References

1. Brown, H. (1981, January). Department of Defense Annual Report *Fiscal Year 1982*. Washington, D.C.: Department of Defense. Retrieved from https://history.defense.gov/Portals/70/Documents/annual_reports/1982_DoD_AR.pdf?ver=2014-06-24-150904-113
2. Department of Defense. (2018). *Summary of the 2018 National Defense Strategy of the United States of America: Sharpening the American Military's Competitive Edge*. Retreived from https://dod.defense.gov/Portals/1/Documents/pubs/2018-National-Defense-Strategy-Summary.pdf
3. U.S. Department of Defense. (2018). *Summary of the 2018 Department of Defense Artificial Intelligence Strategy: Harnessing AI to Advance our Security and Prosperity*. Retreived from https://media.defense.gov/2019/Feb/12/2002088963/-1/-1/1/SUMMARY-OF-DOD-AI-STRATEGY.PDF
4. Krendl, P. (2017, July 25). Better with 'bots? Five questions to ask before automating oil and gas processes. Accenture Energy Blog. Retrieved from https://www.accenture.com/us-en/blogs/blogs-better-with-bots
5. Agrawal, A., Gans, J., & Goldfarb, A. (2018). *Prediction machines: The simple economics of artificial intelligence*. Boston: Harvard Business Review Press.
6. Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., ... & Chen, Y. (2017). Mastering the game of Go without human knowledge. *Nature, 550*(7676), 354. doi:10.1038/nature24270
7. Gregg, A. (2018, June 26). Army to use artificial intelligence to predict which vehicles will break down. *Washington Post*. Retrieved from https://www.washingtonpost.com/business/capitalbusiness/army-to-use-artificial-intelligence-to-predict-which-vehicles-will-break-down/2018/06/25/bfa1ef34-789f-11e8-93cc-6d3beccdd7a3_story.html?noredirect=on&utm_term=.901ce799e155
8. United Kingdom Ministry of Defence. (2018, May). Joint Concept Note (JCN) 1/18, Human-Machine Teaming. Retrieved from https://www.gov.uk/government/publications/human-machine-teaming-jcn-118
9. Wickens, C., Hollands, J., Banbury, S., & Parasuraman, R. (2016). *Engineering psychology and human performance (4th ed.)*. New York: Routledge.
10. de Visser, E. J., Pak, R., & Shaw, T. (2018). From 'automation' to 'autonomy': The importance of trust repair in human–machine interaction. *Ergonomics, 61*(10), 1409–1427. doi:10.1080/00140139.2018.1457725
11. Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995, July). An Integrative Model of Organizational Trust. *Academy of Management Review, 20*(3), 709–734. doi:10.2307/258792
12. Hoff, K., & Bashir, M. (2014, September). Trust in automation: Integrating empirical evidence on factors that influence trust. *Human Factors, 57*(3), 407–434. doi:10.1177/0018720814547570
13. Jian, J. -Y., Bisantz, A., & Drury, C. (2000). Foundations for an empirically determined scale of trust in automated systems. *International Journal of Cognitive Ergonomics, 4*(1), 53–71. doi:10.1207/S15327566IJCE0401_04
14. Davis, D. F. (1989, September). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319–340. doi:10.2307/249008
15. Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly, 27*(3), 425–478. doi:10.2307/30036540
16. Charalambous, G., Fletcher, S., & Webb, P. (2016). The development of a scale to evaluate trust in industrial human-robot collaboration. *International Journal of Social Robotics, 8*(2), 193–209. doi:10.1007/s12369-015-0333-8
17. Ross, J. (2008). Moderators of trust and reliance across multiple decision aids. *Electronic Theses and Dissertations, 3754*. Retrieved from https://stars.library.ucf.edu/etd/3754/
18. de Vries, P., Midden, C., & Bouwhuis, D. (2003). The effects of errors on system trust, self-confidence, and the allocation of control in route planning. *International Journal of Human Computer Studies, 58*(6), 719–735. doi:10.1016/S1071-5819(03)00039-9
19. Bliss, J., Dunn, M., & Fuller, B. (1995). Reversal of the cry-wolf effect: An investigation of two methods to increase alarm response rates. *Perceptual and Motor Skills, 80*(3_suppl), 1231–1242. doi:10.2466/pms.1995.80.3c.1231
20. Hancock, P., Billings, D., Schaefer, K., Chen, J., De Visser, E., & Parasuraman, R. (2011). A meta-analysis of factors affecting trust in human-robot interaction. *Human Factors, 53*(5), 517–527. doi:10.1177/0018720811417254
21. de Visser, E. J., Monfort, S. S., McKendrick, R., Smith, M. A., McKnight, P. E., Krueger, F., & Parasuraman, R. (2016, September). Almost human: Anthropomorphism increases trust resilience in cognitive agents. *Journal of Experimental Psychology: Applied, 22*(3), 331–349. doi:10.1037/xap0000092
22. Pescetelli, N., & Yeung, N. (2018, September). On the use of metacognitive evidence in feedback-free situations: Advice-taking and trust formation. *arXiv preprint*. Retrieved from https://arxiv.org/pdf/1809.10453
23. Metzger, U. & Parasuraman, R. (2005). Automation in future air traffic management: Effects of decision aid reliability on controller performance and mental workload. *Human Factors, 47*(1), 35–49. doi:10.1518/0018720053653802
24. Singh, I. L., Molloy, R., & Parasuraman, R. (1993). Automation-induced" complacency": Development of the complacency-potential rating scale. *The International Journal of Aviation Psychology, 3*(2), 111–122. doi:10.1207/s15327108ijap0302_2
25. Scharre, P. (2014). Robotics on the battlefield Part II: The coming swarm. Washington, DC: Center for a New American Security. Retrieved from https://www.cnas.org/publications/reports/robotics-on-the-battlefield-part-ii-the-coming-swarm
26. Henmon, V. (1911). The relation of the time of a judgment to its accuracy. *Psychological Review, 18*(3), 186–201. doi:10.1037/h0074579
27. Yaniv, I. (2004). Receiving other people's advice: Influence and benefit. *Organizational Behavior and Human Decision Processes, 93*(1), 1–13. doi:10.1016/j.obhdp.2003.08.002
28. Tenney, E., Maccoun, R., Spellman, B., & Hastie, R. (2007). Calibration Trumps Confidence as a Basis for Witness Credibility. *Psychological Science, 18*(1), 46–50. doi:10.1111/j.1467-9280.2007.01847.x
29. See, K., Morrison, E., Rothman, N., & Soll, J. (2011). The detrimental effects of power on confidence, advice taking, and accuracy. *Organizational Behavior and Human Decision Processes, 116*(2), 272–285. doi:10.1016/j.obhdp.2011.07.006
30. Bahrami, B., Olsen, K., Latham, P., Roepstorff, A., Rees, G., & Frith, C. (2010). Optimally Interacting Minds. *Science, 329*(5995), 1081–1085. doi:10.1126/science.1185718
31. Ülkümen, G., Fox, C., & Malle, B. (2016). Two Dimensions of Subjective Uncertainty: Clues From Natural Language. *Journal of Experimental Psychology: General, 145*(10), 1280–1297. doi:10.1037/xge0000202
32. Gaertig, C., & Simmons, J. P. (2018). Do people inherently dislike uncertain advice? *Psychological Science, 29*(4), 504–520. doi:10.1177/0956797617739369

Major Aaron W. Celaya

Air Force Institute of Technology Doctoral Student, New College, Department of Experimental Psychology, University of Oxford

Major Aaron W. Celaya is an Air Force Institute of Technology Doctoral Student, New College, Department of Experimental Psychology, University of Oxford, United Kingdom. He researches metacognition through decision confidence in human-machine teaming with an emphasis on algorithmic advisement sources. This research is conducted within the Attention and Cognitive Control Laboratory, part of the Department of Experimental Psychology at Oxford. Previously, he was the Research Center Director of Operations and Senior Instructor, Warfighter Effectiveness Research Center, Department of Behavioral Sciences and Leadership, Dean of Faculty, United States Air Force Academy.

Nick Yeung, Ph.D.

Professor of Cognitive Neuroscience and Head of the Attention and Cognitive Control Laboratory, University of Oxford

Nick Yeung is professor of cognitive neuroscience and head of the Attention and Cognitive Control Laboratory at the University of Oxford. He received his Ph.D. from the University of Cambridge in 2000, and has also held positions at Princeton University and Carnegie Mellon University. His research investigates human attention, memory, and decision making using a combination of behavioral, computational, and brain imaging techniques. His research funding has included projects with Defense Advanced Research Projects Agency, Intelligence Advanced Research Projects Activity, and the UK Defence Science and Technology Laboratory. He has published more than 70 articles.

(EYE) TRACKING THE INSIDER THREAT:

AN APPLICATION OF THE ACTIVE INDICATORS APPROACH

Gerald Matthews,
Lauren E. Reinerman-Jones,
& Ryan W. Wohleber

Security breaches perpetrated by insiders are a growing threat to a range of organizations, including the Department of Defense (DoD). Securing sensitive information remains of paramount concern. Through unlawful or unauthorized disclosure of sensitive information, insiders are a threat to national security and military operations [1]. The scale of the threat was recognized in DoD Directive 5205.16, *The DoD Insider Threat Program*. Issued in 2014 and revised in January 2017, the directive defines an insider to include any "person who has or had been granted eligibility for access to classified information or eligibility to hold a sensitive position." Activities perpetrated by insiders include unauthorized data exfiltration, destroying or tampering with critical data, eavesdropping on communications, and impersonating other users [2].

Defending against an insider threat (IT) is challenging, as the insider's familiarity with the organization's data structures and security procedures gives them a significant advantage. As defense operations become increasingly dependent on the use of information technology, access to sensitive information will increase for DoD personnel, thus increasing the potential for insider activity.

Strategies for IT defense take a variety of forms [2]. DoD's Defense Personnel and Security Research Center has recognized the need to leverage social and behavioral science to drive a comprehensive research plan and strategy for threat mitigation [3]. One approach is to monitor networks for anomalous events and unusual user behaviors, but such events do not necessarily imply malicious intent. In isolation, such passive indicators make reliable detection of ITs difficult due to a preponderance of false positives. Another approach is to identify environmental and personal factors that may predispose the person to engage in insider activity [4].

For example, an employee with anti-social personality traits, such as lack of conscience, who just incurred a large debt might be vulnerable to an adversary's offer to pay for sensitive information. An obstacle to this approach is the low base rate for malicious insider activities. According to a Defense Personnel and Security Research Center 2018 report, only a small fraction of people with personality flaws and/or difficult life circumstances have the potential to be induced to betray their country [3].

Image Credit | *Photo illustration created by HDIAC and adapted from Adobe Stock.*

Our team of researchers at the University of Central Florida (UCF) explored a detection strategy that may overcome limitations of existing methods for detecting problematic behaviors and people. We investigated an "Active Indicators" (AIs) strategy that aimed to provoke insiders into displaying psycho-physiological responses indicative of their malicious intent. Provoking the insider with a controlled stimulus overcomes limitations of passive monitoring for anomalous behavior. It provides a quick real-time test that utilizes a manageable data set instead of analyzing user access logs, which may be a lengthy process that slows response time to a threat. The AI approach also eliminates the need to make personal character judgments (which may not be valid).

The UCF research we will discuss in this article created immersive simulations of sensitive financial and intelligence work. Insiders were "provoked" by opportunities to gather information illicitly and security alerts. Their eye movement responses were assessed and analyzed to identify attentional and emotional metrics that might be diagnostic of the insider.

The IARPA SCITE Program

An alternative approach that complements network activity analysis and prediction of insider behavior has been developed by the Intelligence Advanced Research Projects Activity (IARPA). The Scientific Advances to Continuous Insider Threat Evaluation (SCITE) program [5] aimed to investigate countermeasures based on AIs—a class of diagnostic security tool. AIs are events or stimuli that elicit diagnostic behaviors that discriminate insiders from people performing legitimate actions. An example of an AI would be an email sent to employees of an organization warning of an imminent and thorough security check of all hard drives. Employees who suddenly delete numerous files from their hard drive would be showing a suspicious response—which could be used diagnostically, triggering further investigation of those involved.

The SCITE AI thrust supported research at GE Global Research, Raytheon BBN Technologies, and UCF. Researchers were tasked with identifying AIs that would produce high rates of insider detection relative to false positives. They were also required to test generalization of results across various environments. The UCF team focused on AIs that might elicit diagnostic psychophysiological responses, primarily from eye tracking, using immersive simulations of work environments. In this article, we examine this work and its implications.

Figure 1. Screenshots illustrating financial (top) and espionage (bottom) environments

From the Psychophysiology of Deception to Active Indicators

The psychology and neuroscience of deception have advanced considerably from the traditional polygraph or "lie detector test." Theories of deception [6, 7] recognize multiple indicators of deceptive behavior and intent that include both emotional and cognitive responses. On the emotional side, deception is associated with general stress and specific emotions, such as fear and guilt. A polygraph detects the increased sympathetic arousal associated with these responses. On the cognitive side, deception changes the ways in which information is processed. People attempt to regulate their responses to avoid detection; for example, they may be attentive to cues that others are suspicious of them. Deception also elevates cognitive workload because it takes mental effort to maintain a false narrative.

Most scientific research into deception is based on highly-structured interrogation procedures where targets are aware that questions put to them are designed to identify suspicious responses. However, the focus of IT detection efforts is different. The need is to evaluate large numbers of individuals unobtrusively as they perform their normal work duties, without formal questioning. Ex-

Scenario	Concept	Key AIs (Financial Environment)
1. Opportunity	Insider has time-limited opportunity to access prohibited information source	• Instruction from handler • Confederate available for contact • Password supplied by confederate • Illicit file access • Deliver info to handler
2. Person-of-Interest	Insider covertly monitors files for information on named person of interest	• Instruction from handler • File linked to person-of-interest appears • Deliver folder name to handler
3. Hacking Tool	Insider uses software tool to illicitly access computer systems under the direction of an external handler	• Instruction from handler • First use of hacking tool • Spike in network activity resulting from hacking • Failure in system access • Termination of episode
4. Accomplice	Insider works with an accomplice to illicitly access computer systems under the direction of an external handler	• Instruction from handler • First interaction with accomplice • Spike in network activity resulting from accomplice actions • Failure in system access • Termination of episode

Table 1. Examples of AIs in four simulated scenarios

isting research provides some examples of how research can be transitioned from formal lie detection to more indirect evaluations. For example, a study published in 2014 [8] designed an automated screening kiosk that tested for eye movements orienting toward words such as “bomb” and “gun.” Our research builds on such findings by integrating AIs and detection of diagnostic responses into an ongoing workflow.

Figure 2. Target image for U.S. insider (espionage Person-of-Interest scenario)

Simulation Methodology

One of the challenges for conducting controlled, experimental studies of IT is the artificial nature of the laboratory environment, in which little is at stake for the participant. We met this challenge by designing realistic, immersive simulations that produce psychological fidelity to real-world settings. That is, the simulation elicits responses analogous to those occurring in corresponding real environments [9].

Simulations were built for two environments vulnerable to ITs: financial work in a bank, and analysis work for military intelligence. In each case, participants’ main task was to perform normal work under the direction of an off-site supervisor, reconstructing corrupted accounts and analyzing possible terrorist plots, respectively. Detailed scripts for interacting with the supervisor were developed for each simulation to support immersion. Figure 1 illustrates the two environments. Each task performed by participants was assigned to a separate window in the display. For example, in the financial scenario, the participant uses the dialogue box to interact with the supervisor, searches for account information in file cabinets and computer files, and interacts with other employees via a chat window.

To motivate insider participants, we had them play the role of an insider working for the U.S. government embedded in a foreign organization. In financial scenarios, participants were U.S. tax agents seeking information on undisclosed foreign accounts held in a Swiss bank. In espionage scenarios, participants were U.S. agents infiltrated into a Russian counter-terrorism intelligence unit that was obtaining information about plots against the United States. Depending on the scenario, insider participants were also required to surreptitiously perform a range of other activities, such as opening prohibited files and exfiltrating data to an accomplice. Control participants performed matched activities, such as legitimately opening protected files and communicating with personnel outside the organization. Careful matching of activities allowed us to attribute any psychophysiological response differences between insiders and controls to insider intent rather than to differences in tasking.

We ran eight studies, representing four different scenarios replicated in financial and military intelligence environments. Table 1 provides examples of the AIs built into each scenario. In scenarios 3 and 4, we scripted a narrative in which the insider's activities became progressively more risky, eliciting warnings from the organization's cybersecurity team. Typically, multiple AIs were presented as a sequence of events supporting an illicit action. For example, in the "Person-of-Interest" scenario (espionage version), the insider is briefed on the identity of a suspected terrorist and how to recognize his presence in a building. In one instance, the terrorist is known to travel in a distinctive vehicle. The image shown in Figure 2 identifies the location of the terrorist which the participant must relay to an outside U.S. handler—thus, the appearance of the image serves as an AI.

The primary response measures used for detection of insider intent were derived from eye tracking, based on the deception literature. Figure 3 illustrates the experimental set up. Our eye tracker utilized an infrared light source that illuminates the eye. A camera records the reflection of the infrared from the whole eye, including the pupil. The sequence of camera images is then analyzed to provide multiple eye tracking metrics.

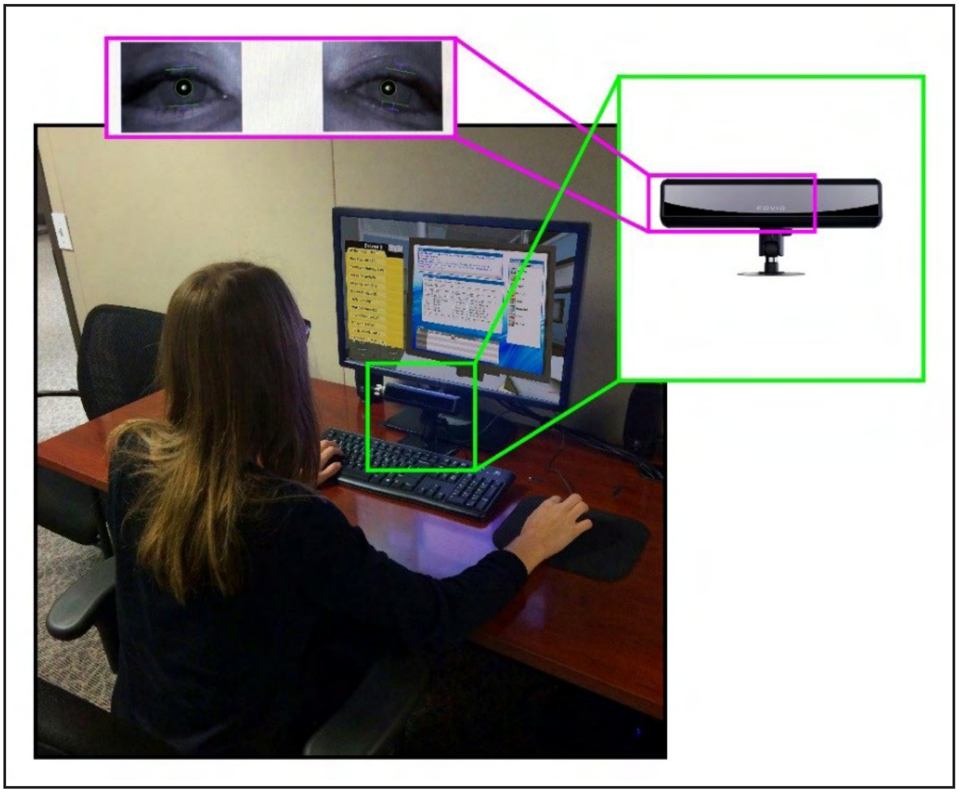


Figure 3. Lab workstation showing placement of eye tracker

Movements of the eye are defined by fixations (where the eye dwells on a location) and saccades (where the eye moves rapidly between different locations). The frequencies and durations of fixations and saccades are measured. The frequency of fixations on particular task windows can also be assessed, providing an index of attention to each task component. Eye tracking software also analyzes the diameter of the pupil, which is sensitive to both cognitive workload and stress. In order to motivate attempts to avoid detection, participants were told explicitly that their eye movements were being monitored and suspected insiders who were detected would face severe penalties.

For Scenario 4, we added a thermal camera to measure temperature declines in the perinasal area of the face, which is associated with perspiration. This metric provides an off-the-body assessment of sympathetic activation, which is sensitive to a variety of stress factors, including deception [10].

Key Results

Findings from the research are too numerous to list in full; we have published reports on the first two scenarios [11, 12]. Across the different studies, we found that multiple response measures were diagnostic of insider intent, including frequency, duration and direction of eye fixations, pupil diameter, and thermographic response. We identified three main types of indicator:

- *Strategic concealment of vigilance*. The first two scenarios required the insider to monitor one of the task windows without attracting attention from supervisors. The Opportunity task required watching a chat program contact list of online associates for the appearance of an accomplice who would supply a password. In the Person-of-Interest task, participants monitored building images supplied from a drone feed. Insiders showed a pattern of shorter duration and/or less frequent fixations on the task window, as though they were trying to peek at the window surreptitiously. Such covert vigilance could be controlled consciously or unconsciously. Consciously controlled eye movements tend to have longer durations than unconscious ones. The insider effect tended to be stronger for longer-duration fixations (150–900 ms)—suggesting insiders utilized a deliberate strategy to disguise their interest in the task window, given that they were aware of being monitored.

- *Monitoring for cues to being detected*. Instructions highlighted the presence of cybersecurity and insiders were aware that their activities might lead to their detection. In all four scenarios, we found evidence that insiders showed increased fixation on task windows that could provide cues to detection. In Scenarios 1 and 2, following critical AIs, insiders tended to show increased fixation on the dialogue box that supported their conversations with their supervisor, who periodically mentioned security concerns. In Scenarios 3 and 4, we included additional AIs suggesting risk of detection, including a monitor showing spikes in network activity and system access errors—both of which were elicited by the insider's hacking attempts. Toward the end of the task, as security appeared to be closing in on the insider, we found shifts in attention to these sources of cues following AIs.

Figure 4 illustrates the insider effect. The upper panel shows fixation durations following three AIs, as the insider begins hacking and causes a spike on the network activity monitor. The first and third AIs in the sequence elicited significantly longer fixation durations in insiders. The lower panel shows the distribution of fixation durations in each group. Using values of >230ms as an index to identify possible insiders, we accomplished a

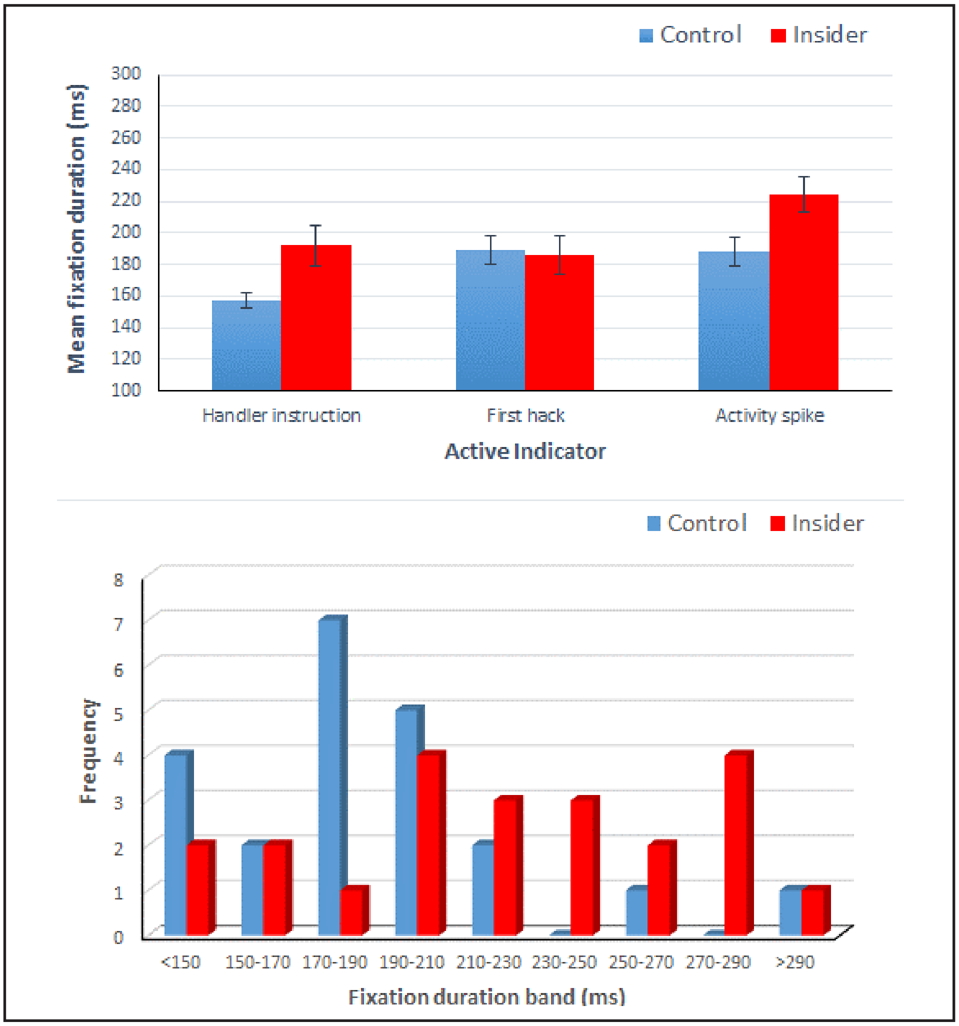


Figure 4. Mean fixation duration for eye movements following three AIs (top) and frequency distribution of responses to third AI (bottom), in control and insider participants

true positive rate of detection of 46% of the insiders. Few of the controls exceeded the cutoff; the false positive rate was 9%. Discrimination was imperfect, but the difference in the two rates supports the utility of the index.

- *Stress response*. On the basis of previous studies, we anticipated elevated stress response in insiders. In fact, changes in attention evidenced in fixation data appeared to be more effective for insider detection. Given the limitations of laboratory studies, it may be difficult to reproduce the levels of stress which a real insider might experience in a real-world scenario.

Nevertheless, we found a few instances of elevated stress, especially in scenarios 3 and 4, in which the narrative emphasized the possibility of being detected. For example, in scenario 4, thermographic facial monitoring indicated greater sympathetic arousal in insiders.

Toward Applications

The AI methodology outlined here does not provide a magic bullet for identifying ITs in the DoD or its contractors. Even successful indicators do not accurately detect all ITs. As illustrated in Figure 4, while many insiders may respond diagnostically to a given AI, others do not. Also, given the low base rate for IT activity, false positives will inevitably occur.

However, the diagnosticity of eye tracking responses is strong enough to suggest a place for them in countermeasures to insiders. As all DoD personnel with at least a Secret clearance have access to Distribution B-E material (over four million individuals held security clearances as of October 2016 [13]), an immediate screening process is needed in order to identify a smaller subset who should be monitored or investigated for possible IT. Use of AIs, in conjunction with other insider-detection methods [2], may contribute to an efficient, automated screening process flagging a relatively high proportion of insiders in the screened sample. Such efforts should be conducted within a systematic framework for investigating, detecting, and reacting to IT events [14].

Designing an AI-based system requires ingenuity and domain knowledge. For example, if the target population is military analysts, AI stimuli might be incorporated into the documents and images they work with. “Honeytokens” [15]—digital entities of special interest to the insider—can be utilized. Our work suggests that attempts to disguise interest in the honeytoken are more diagnostically actionable than eye tracking indicators of interest itself. The increasing availability of intelligent decision aids also provides an avenue for application. A decision aid might incorporate monitoring for cyberattack and report such efforts to users and their superiors. AIs suggesting detection of illicit activity may provoke diagnostic eye movement responses. Use of a human-like avatar for the aid might heighten the insider’s concerns about being detected. There is considerable scope for future research to validate such methods.

Such methods are more workable in contexts in which user activities can be sufficiently structured to interpret eye tracking data. In some cases, interfaces may be designed to facilitate interpretation (e.g., by separating different data sources and task windows). Users’ awareness of being monitored is also critical for success; they must be motivated to disguise their attention to illicit information sources and to monitor for security countermeasures. Looking forward, this system can play a key role in identifying and tracking potential threats to national security from malicious insiders.

Acknowledgements

This research featured in this article is based on work supported by the Office of the Director of National Intelligence (ODNI), IARPA, via IARPA R&D Contracts, contract number 2016-16031500006.

The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the ODNI, IARPA, or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright annotation thereon.

References

1. What is "Insider Threat?". (2016). CHIPS – *The Department of the Navy's Information Technology Magazine*. Retrieved from https://www.doncio.navy.mil/CHIPS/ArticleDetails.aspx?ID=8350
2. Sanzgiri, A., & Dasgupta, D. (2016). Classification of insider threat detection techniques. *Proceedings of the 11th Annual Cyber and Information Security Research Conference, 16*. doi:10.1145/2897795.2897799
3. Jaros, S. L. (2018, October). *A Strategic Plan to Leverage the Social & Behavioral Sciences to Counter the Insider Threat* (Technical Report No. PERSEREC-TR-18-16, OPA-2018-082). Seaside, CA: Defense Personnel and Security Research Center. Retrieved from https://www.dhra.mil/Portals/52/Documents/perserec/reports/TR-18-16-Strategic-Plan.pdf
4. Zimmerman, R. A., Friedman, G. M., Munshi, D. C., Richmond, D. A., & Jaros, S. L. (2018). *Modeling Insider Threat from the Inside and Outside: Individual and Environmental Factors Examined Using Event History Analysis* (Technical Report No. PERSEREC-TR-18-14, OPA-2018-065). Seaside, CA: Defense Personnel and Security Research Center. Retrieved from https://www.dhra.mil/Portals/52/Documents/perserec/reports/TR-18-14_Modeling_Insider_Threat_From_the_Inside_and_Outside.pdf
5. Intelligence Advanced Research Projects Activity. (n.d.). SCITE program. Retrieved from www.iarpa.gov/index.php/research-programs/scite
6. Burgoon, J. K., & Buller, D. B. (2015). Interpersonal deception theory: Purposive and interdependent behavior during deceptive interpersonal interactions. In L. A. Baxter & D. O. Braithwaite (Eds.), *Engaging Theories in Interpersonal Communication: Multiple Perspectives* (pp. 349–362). Los Angeles, CA: Sage Publications.
7. Zuckerman, M., DePaulo, B. M., & Rosenthal, R. (1981). Verbal and nonverbal communication of deception. *Advances in Experimental Social Psychology, 14*, 1–59. doi:10.1016/s0065-2601(08)60369-x
8. Twyman, N. W., Lowry, P. B., Burgoon, J. K., & Nunamaker Jr, J. F. (2014). Autonomous scientifically controlled screening systems for detecting information purposely concealed by individuals. *Journal of Management Information Systems, 31*(3), 106–137. doi:10.1080/07421222.2014.995535
9. Ortiz, E., Reinerman-Jones, L., & Matthews, G. (2016). Developing an Insider Threat Training Environment. *Advances In Intelligent Systems And Computing*, 267-277. doi: 10.1007/978-3-319-41932-9_22
10. Pavlidis, I., Eberhardt, N. L., & Levine, J. A. (2002). Human behaviour: Seeing through the face of deception. *Nature, 415*(6867), 35. doi:10.1038/415035a
11. Matthews, G., Reinerman-Jones, L.E., Wohleber, R., & Ortiz, E. (2017). Eye tracking metrics for insider threat detection in a simulated work environment. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 61*(1), 202–206. doi:10.1177/1541931213601535
12. Matthews, G., Wohleber, R., Lin, J., Reinerman-Jones, L., Yerdon, V., & Pope, N. (2018). Cognitive and affective eye tracking metrics for detecting insider threat: A study of simulated espionage. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 62*(1), 242–246. doi:10.1177/1541931218621056
13. National Counterintelligence and Security Center. (2016). *Fiscal Year 2016 Annual Report on Security Clearance Determinations*. Retrieved from https://fas.org/sgp/othergov/intel/clear-2016.pdf
14. Raytheon (2009). *Best Practices for Mitigating and Investigating Insider Threats*. (White Paper #2009-144). Salt Lake City, UT: Raytheon Company. Retrieved from https://www.raytheon.com/sites/default/files/capabilities/rtnwcm/groups/iis/documents/content/rtn_iis_whitepaper-investigati.pdf
15. Spitzner, L. (2003, July 17). Honeytokens: The other honeypot. SecurityFocus. Retrieved from https://www.symantec.com/connect/articles/honeytokens-other-honeypot

Gerald Matthews, Ph.D.
Research Professor, Institute for Simulation and Training, University of Central Florida (UCF)

Gerald Matthews is a research professor at the UCF Institute for Simulation and Training. He received his Ph.D. in experimental psychology from the University of Cambridge in 1984. He is an associate editor for Human Factors and Personality and Individual Differences. He is also an International Association for Applied Psychology (IAAP) Fellow and a past president of the International Society for the Study of Individual Differences (ISSID). His research interests include human factors engineering for cybersecurity, the neuroergonomics of stress, workload and fatigue, and individual differences in cognition and emotion.

Lauren E. Reinerman-Jones, Ph.D.
Research Professor, Institute for Simulation and Training, University of Central Florida (UCF)

Lauren Reinerman-Jones is the Prodigy Lab Director at the UCF Institute for Simulation and Training. She received her Ph.D. in human factors psychology from the University of Cincinnati in 2008. She is the chair of the DoD Human Factors and Ergonomics Technical Advisory Group Trust in Autonomy Sub-TAG, and a recipient of the Admiral Leland S. Kollmorgen Spirit of Innovation Award for Augmented Cognition. Her research centers on interdisciplinary collaborations using physiological measures for resilience, human-robot interaction, decision making, socio-cultural team dynamics, automation, virtual to live environment comparisons, and the assessment of emotional and cognitive states.

Ryan Wohleber, Ph.D.
Associate in Human Factors, Institute for Simulation and Training, University of Central Florida (UCF)

Ryan Wohleber is a research associate in the Institute for Simulation and Training at UCF. He holds degrees in industrial design (B.S.) and psychology (M.A.) from University of Cincinnati, and degrees in modeling & simulation (M.S., Ph.D.) from UCF. He was awarded a Combined Arms Center – Training organizational coin by Deputy Commanding General and STE CFT Director, Major General Gervais, in recognition of excellence in support of RVCT user assessment and technology readiness assessment efforts. His research interests include attitudes, implicit measures, stress, fatigue, emotional intelligence, deception, and decision making for various human factors applications.

RS-OCT: Bimodal Diagnostic System for Classification of Burn Injuries

Priya Ganapathy,
Lakshmi Priya Rangaraju,
Gautam Kunapuli,
Dayna Every,
Oscar D. Ayala,
Anita Mahadevan-Jansen,
& Jacob Yadegar

According to the Centers for Disease Control and Prevention, burns are one of the leading causes of nonfatal injury in the United States, across all age ranges [1]. Thermal injuries account for approximately 10% of combat casualties in the modern era [2], and these injuries are increasingly incurred in tandem with blast or explosion wounds [3]. For instance, burns comprised 5% of all casualties evacuated from Operation Iraqi Freedom and Operation Enduring Freedom between April 2003 and April 2005 [3]. Due to improved combat burn care and shorter response times, burn injury survival rates have risen since the Vietnam era, but thermal injuries from explosions are likely to pose a continued threat to U.S. forces.

In a battlefield setting, wounded warfighters must be triaged in order to identify the priority of medical treatment and evacuation. Early and close assessment of burn degree (i.e., severity level) informs sound decisions on the type of treatment to be given. However, such burn analysis can be highly subjective, and the unstable nature of burn wounds can further complicate the decision-making process. Research has long shown that early treatment decisions regarding burn injuries result in shorter hospital stays and lower rates of infection [4, 5]—underlining the importance of rapid assessment.

To date, Department of Defense (DoD) entities, such as U.S. Army Medical Research and Materiel Command and the Military Medical Photonics Program, have sponsored research and development in non-invasive techniques like terahertz imaging, infrared spectroscopy, laser Doppler imaging, and reflectance mode confocal microscopy [6]. However, current methods for burn triage are mostly visual and checklist-based [7].

To counter these challenges, we present a novel and portable battlefield point-of-care diagnostic system that combines two complementary optical technologies, Optical Coherence Tomography (OCT) and Raman Spectroscopy (RS), with a clinical decision support system. This system can (a) promptly and accurately classify burn severity, and (b) recommend a course of action for urgent treatment. Applying point-of-care diagnostics to burn injuries will provide medics or the first line of field caregivers an essential tool for providing optimal treatment to injured warfighters [8].

Limitations of Existing Diagnostic Techniques

An accurate assessment of burn severity (i.e., superficial, superficial-partial, deep-partial, or full thickness) is crucial in determining the type of intervention required to treat a burn victim. A conventional method of identifying burn severity is through visual perception, which is mainly based on the total body surface area (TBSA). Although superficial and full-thickness types of burns are visually discernable, it can be difficult even for experienced surgeons to accurately distinguish superficial-partial and deep-partial thickness burns [9, 10]. Currently, laser Doppler imaging (LDI) is often used as a non-invasive diagnostic tool for estimating burn degree [11]. LDI assesses burn depth by measuring the perfusion index (PI), defined as the measure of microcirculatory changes in skin obtained through Doppler shift of reflected laser light.

Observational studies have reported a good correlation between LDI's PI and burn depth.

Image Credit | *Photo illustration created by HDIAC and adapted from a U.S. Army photo by Joseph Wilbanks (available for viewing at: https://www.army.mil/article/152515/innovative_treatments_offer_hope_for_burn_victims) and a UtopiaCompression Corp. photo.*

However, due to the influence of reactive vasoconstriction, LDI's diagnostic accuracy is highest when used more than 24 hours after the onset of injury [12]. Furthermore, changes in PI are susceptible to conditions such as anemia or infection, leading to obscure perfusion maps. Physical barriers, such as the curvature of tissue surface, scanning distance, and thickness of topical dressings, also influence LDI's outcomes [13, 14]. A survey study of burn centers in the United States reported in 2014 that surgeons depended mostly on clinical examination (60%), LDI (6%), and biopsy (4%) to estimate burn depth [12]. The study also highlighted cost (72%), restricted availability of access (63%), and the absence of up-to-date evidence (35%) as the three main barriers to incorporating new technologies [15].

Assessing Burn Severity with a Complementary "Bimodal" System

Definitive treatment and monitoring of a typical burn wound requires assessment of changes in skin thickness (i.e., the amount of skin tissue damaged and extent of layers involved) along with functional aspects of the underlying skin components. Capturing structural morphology and the chemical nature of skin tissue is of great significance in order to effectively analyze burn wounds. We propose a diagnostic system that can precisely identify burn severity based on features combined from two complementary optical techniques: OCT and RS.

OCT is an imaging technique which uses the principle of low-coherence interferometry to capture cross-sectional images of biological tissues with micro resolution. This allows for the visualization of structures up to 1–2 mm under the surface of highly irregular, light-scattering tissues such as skin. OCT produces a 2D image known as a B-scan, which is composed of a series of lateral A-scans. Several B-scans acquired over an area enable 3D reconstruction of the target region.

Because intensity-varied pixels in an OCT image represent the amount of light scattered from target tissue, changes in the optical properties of a target area of tissue are reflected in the image. *Birefringence* is known to degrade due to the loss of collagen caused by a burn injury, thereby restricting the penetration of light. Consequently, a converse relation between the cross-sectional depth estimate (derived from the OCT image) and the severity of the burn can be evaluated [16]. Furthermore, analyzing intensity-based textural measures may help differentiate burn characteristics.

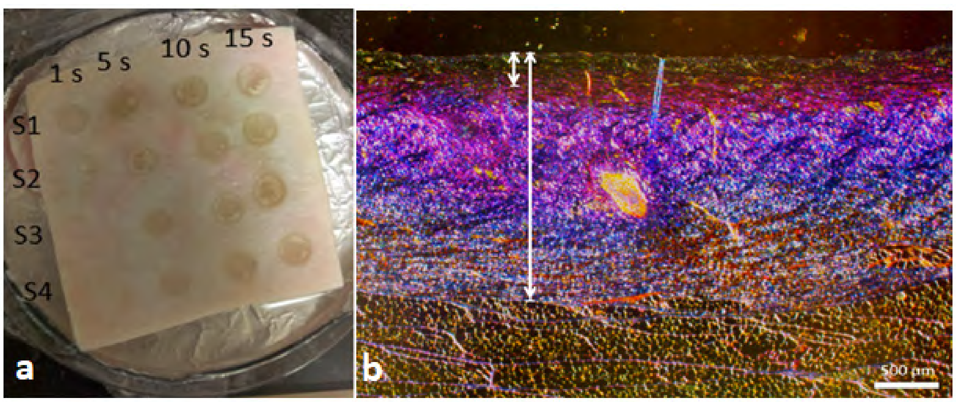


Figure 1. (a) Sample of cadaver porcine skin showing burn wounds created at 1s, 5s, 10s, and 15s durations of time (burns created at 1s were not used in the study). (b) Histological image of 10s burn with coagulated (short line) to total (long line) collagen ratio of less than 0.35—labeled as SPT wound.

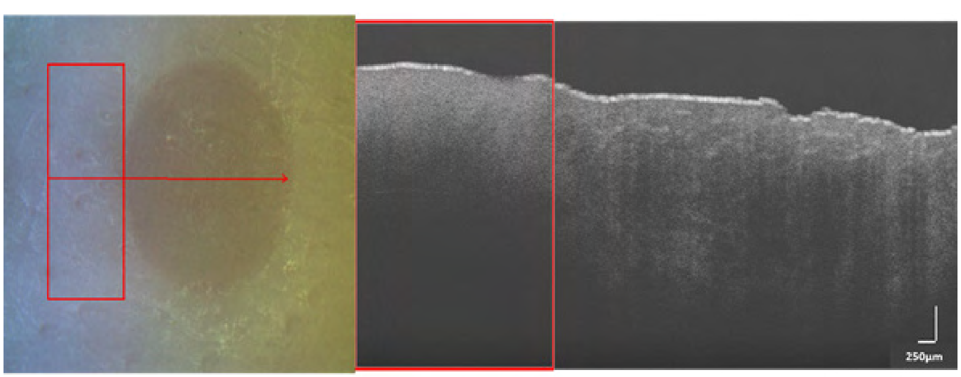


Figure 2. (left) White light image of a SPT burn with red arrow indicating linear scan region from left to right. (right) B-scan acquired from OCT (250 µm) over the region shown in left image. Rectangular red box in both images represents the control (normal) skin region and the region outside it covers burn portion.

OCT has been effectively employed to study different skin conditions in terms of anatomy and physiology [17], where specific characteristic features, such as the thickening of the epidermis, signal attenuation in the dermis, and homogenous signal distribution from OCT scans, helped recognize skin tumors and inflammation [18, 19]. Other variants of OCT, namely polarization-sensitive OCT and OCT angiography, have also proved useful in studies for evaluating different skin abnormalities and burns [20–22].

RS relies on the shift in wavelength of the inelastically scattered light (Raman shift). This takes place when electromagnetic waves interact with molecules of a given target, revealing their chemical structures. The chemical sensitivity of biomolecules (e.g., proteins, lipids, nucleic acids) can be used to analyze their corresponding spectral changes. Since it is known that constituent proteins in skin are denatured as a result of a burn, the respective spectra may be correlated with burn severity [23].

The utility of RS in diagnosing skin conditions has been demonstrated in multiple studies, where researchers have shown associated changes in lipid molecules of the *corneum stratum* layer with Raman spectra [24], distinguished malignant from benign skin lesions with high sensitivity [25], and used confocal RS to provide differential diagnoses of melanoma, basal cell carcinoma, and other non-melanoma skin cancers [26, 27]. The ability of RS to extract hydration-relevant details from skin and heterotopic ossification due to burns highlights the significance of employing RS to assess burn wounds [27, 28].

An amalgamation of these two techniques can produce a system that leverages the complementary capabilities of each. As such, microstructural details from OCT wide-field images, along with spectral features corresponding to biochemical specificity of underlying skin molecules, can together accurately and reliably diagnose the severity of burns. Experiments utilizing this RS-OCT system have already successfully recognized cancer skin lesions and tissues of liver, kidney, and small intestine [29, 30].

Burn type	Mean Intensity (Mean ± SD)		Variance (Mean ± SD)		Entropy (Mean ± SD)		Smoothness (Mean ± SD)		Depth [mm] (Mean ± SD)	
	Control	Burn	Control	Burn	Control	Burn	Control	Burn	Control	Burn
Full-thickness (FT) (N=26)	126.90 ± 7.29	110.81 ± 6.34	0.48 ± 0.04	0.37 ± 0.06	2.48 ± 0.14	1.86 ± 0.40	0.19 ± 0.03	0.12 ± 0.03	0.60 ± 0.05	0.44 ± 0.12
Deep partial-thickness (DPT) (N=22)	124.30 ± 5.78	109.18 ± 6.39	0.48 ± 0.03	0.31 ± 0.09	2.33 ± 0.21	2.09 ± 0.51	0.19 ± 0.02	0.09 ± 0.05	0.56 ± 0.06	0.52 ± 0.15
Superficial partial-thickness (SPT) (N=28)	126.41 ± 7.42	112.55 ± 3.98	0.49 ± 0.05	0.35 ± 0.05	2.28 ± 0.14	2.68 ± 0.22	0.19 ± 0.03	0.11 ± 0.03	0.55 ± 0.04	0.69 ± 0.07

Table 1. Average values of features computed from OCT scans over control and burn regions for each burn type

RS-OCT-based Identification of Burn Severity

Experimental Setup

UtopiaCompression Corporation (UC) collaborated with Vanderbilt University (VU) in Nashville, Tennessee, to develop a reliable burn classification system by combining RS and OCT optical techniques. Ex vivo experiments on porcine skin to collect data on burn wounds were conducted at VU, whereas UC carried out the data processing, analysis, and development of classification models.

As part of the study, burn wounds were inflicted on cadaver porcine skin samples by following a protocol set by Ponticorvo et al. [31], wherein a soldering rod heated to 140 degrees Celsius was used to create burn wounds for contact durations of 5, 10, and 15 seconds. Considering the sample size criterion required for valid statistical analysis and inference, 26 wounds for each contact time were generated (5s: N=26, 10s: N=26, 15s: N=26) (see Figure 1a), and a histological test using trichrome staining was performed to confirm burn severity.

Skin's exposure to heat denatures its collagen structure, resulting in coagulation that corresponds with burn degree [32, 33] and contact time. Consequently, the staining procedure, enhancing collagen in a unique color, helps reveal the amount of coagulation. A ratio of coagulated collagen to total collagen thickness was measured manually from a histological slide using calipers, and a severity level was assigned to each wound according to the ratio value [31]. Burns with values (a) less than 0.35 were labeled as superficial-partial thickness (SPT); (b) between 0.35 and 0.65 as deep-partial thickness (DPT); and (c) above 0.65 as full thickness (FT) wounds (see Figure 1b). Thus, based on the histology results, the final data sample consisted of 28-SPT, 22-DPT, and 26-FT burns.

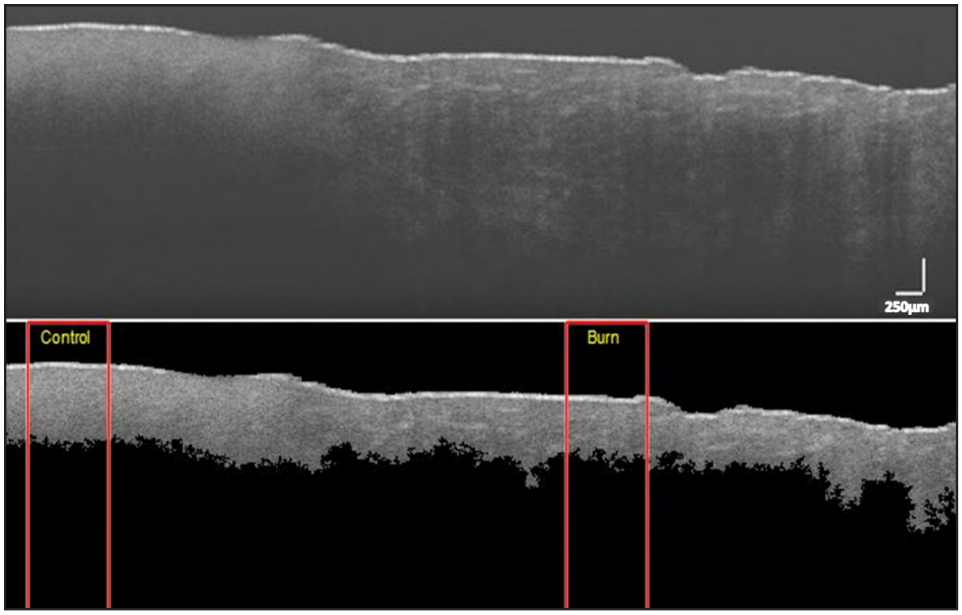


Figure 3. OCT image of superficial-partial thickness burn. (top) Unprocessed image. (bottom) Segmented image. Rectangular red boxes indicate 300-column region used for extracting features (Depth: Control = 0.66mm, Burn = 0.58mm; Entropy: Control = 2.56, Burn = 2.27) (Scale: 250µm).

After the burn wounds were created, each spot was scanned under OCT and RS. Commercial OCT equipment (*Thorlabs*, Telesto TEL1300V2-BU) with a depth resolution of 5.5 µm (in air) was used to acquire images at 0.201 seconds acquisition time and 76 kHz imaging speed. A single image was formed by averaging 20 horizontal B-scans of 9mm length. The scan region included a small portion of normal skin adjacent to burnt skin, so that the features extracted from the burnt portion could be compared with that of normal skin (see Figure 2). The RS device used in the study was a prototype developed by the VU team [34] that was built using a 785nm laser source at 80mW, a fiber optics probe (EmVision LLC, USA), 10 collection fibers (300µm), and an excitation fiber (200 µm). The device was calibrated for wavelength and intensity before scanning the sample. The effective scanning area was 19mm^2.

Normal spectra (i.e., control, before burn creation) and burn spectra (after burn creation) were obtained four times at each spot with an integration time of three seconds and 20 accumulations. Post-processing steps such as noise reduction techniques and fluores-

Performance Measures	Logistic Regression			Linear SVM			Random Forest		
	OCT	RS	RS-OCT	OCT	RS	RS-OCT	OCT	RS	RS-OCT
AUC-ROC: Class-1 (FT burns)	0.90	0.95	0.96	0.84	0.88	0.91	0.92	0.96	0.98
AUC-ROC: Class-2 (DPT burns)	0.80	0.87	0.89	0.64	0.82	0.85	0.76	0.90	0.89
AUC-ROC: Class-3 (SPT burns)	0.84	0.98	0.99	0.76	0.93	0.94	0.85	0.97	0.97
Average AUC-ROC (FT, DPT and SPT burns)	**0.87**	**0.94**	**0.95**	**0.76**	**0.88**	**0.90**	**0.85**	**0.95**	**0.95**
Overall Accuracy (%)	**67.11**	**80.26**	**86.84**	**68.42**	**84.21**	**86.88**	**69.74**	**81.58**	**78.95**

Table 2. Performance measures of three selected classifiers each trained with three data sets. RS features used as input are computed relative to control spectra (Method-I). Accuracy and average AUC-ROC values under each case are bolded.

Performance Measures	Logistic Regression			Linear SVM			Random Forest		
	OCT	RS	RS-OCT	OCT	RS	RS-OCT	OCT	RS	RS-OCT
AUC-ROC: Class-1 (FT burns)	0.90	0.92	0.93	0.84	0.89	0.91	0.92	0.95	0.97
AUC-ROC: Class-2 (DPT burns)	0.80	0.89	0.86	0.64	0.81	0.87	0.76	0.92	0.93
AUC-ROC: Class-3 (SPT burns)	0.84	0.99	0.99	0.76	0.98	0.98	0.85	0.99	0.99
Average AUC-ROC (FT, DPT and SPT burns)	**0.87**	**0.94**	**0.93**	**0.76**	**0.90**	**0.92**	**0.85**	**0.95**	**0.96**
Overall Accuracy (%)	**67.11**	**82.89**	**84.21**	**68.42**	**86.84**	**89.47**	**69.74**	**85.53**	**85.53**

Table 3. Performance measures of three selected classifiers each trained with three data sets. Absolute RS features are used as inputs to the classifiers (Method-II). Accuracy and average AUC-ROC values under each case are bolded.

cence background subtraction were applied, which primarily involved binning data to half the spectral resolution and noise filtering with second order Savitzky-Golay filter (window size=2*spectral resolution), and modified polynomial-fitting algorithm for auto-fluorescence subtraction [35, 36]. The final spectra were mean-normalized to account for any variation in intensity.

Data Analysis and Classification

Data analysis included evaluating grayscale images acquired from OCT and spectra from RS. First, OCT images were processed to segment skin region from background, with image processing steps including contrast enhancement, thresholding by Otsu's method, maximum component determination using flood-fill algorithm, and morphological refinement [37]. Second, depth [16] and statistical textural features (mean, variance, entropy, and smoothness) [38] were computed for non-zero pixels of the segmented images over two 300-column regions corresponding to control and burn regions (see Figure 3). Averages of all computed values for different burn types are listed in Table 1. The differences of mean depth between control and burn regions were statistically significant for SPT and FT burns ($p<0.01$), but insignificant for DPT burns ($p=0.23$). Considering burn regions, only depth and entropy values significantly decreased from SPT to FT ($p<0.01$).

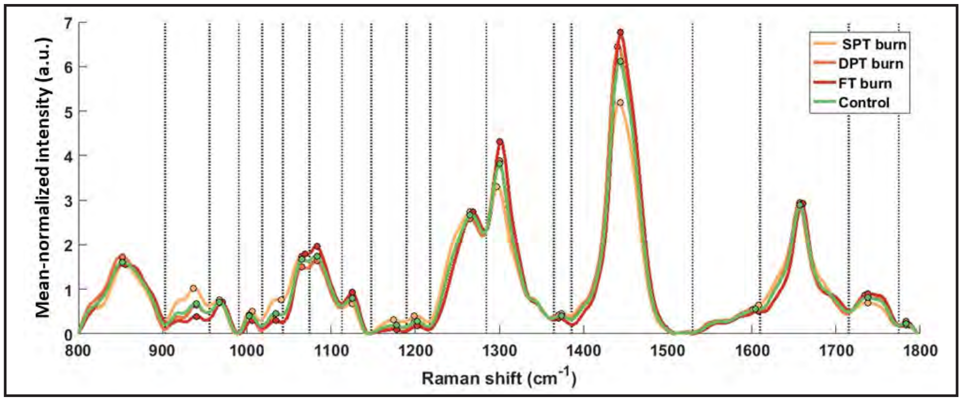


Figure 4. Mean normalized Raman spectra averaged across all spots for each burn type. 18 segments (bands) separated by dotted delimiters, drawn at local minima, were obtained using a peak detection algorithm. All relative features (Method I) (N=54) were computed across each of the 18 segments, and peak intensities in each segment are marked with colored circles.

The normalized Raman spectra, acquired four times at each burn spot, were reduced to one spectrum per spot via averaging. The spectra, 1D array, consisted of varying intensity values at different wavenumbers; however, we used a range between 800–1,800cm^{-1} for computing features relevant to skin components. For simplicity in computational flow, the selected range of wavenumbers was divided into bands using a peak-detection algorithm by binning the wavenumbers between two local minima into one band, thus resulting in 18 bands/segments for one spot (see Figure 4).

Accounting for possible spectral artifacts, RS features were computed relative to control spectra (Method I – Relative RS features, N=54/spot). Nevertheless, in real-world situations, acquiring control spectra may not be possible; therefore, another set of features from only burn spectra (Method II – Absolute RS features, N=21/spot) were computed (see Figure 5).

In order to classify burn severity with high reliability, the number of features and training data size were the major factors used to short-list three candidate machine learning (ML) classifiers: logistic regression (LR), linear support vector machines (linear SVM), and random forests (RF). Each classification model was implemented on Waikato Environment for Knowledge Analysis (WEKA) software (University of Waikato, Hamilton, New Zealand), using tenfold cross validation. Accuracy and area under the receiver operator characteristic (AUC-ROC) curve were used to gauge classifiers' performance [39]. Figure 6 illustrates workflow of the evaluation method that we adapted to identify the robustness of OCT and RS features in distinguishing burns. Each of the three classifiers was tested under two feature sets (Method I and Method II), and the model evaluation metrics were recorded (see Table 2 and Table 3, respectively).

When the three classifiers were trained using only OCT features, accuracy and AUC-ROC ranging from 67–69% and 0.76–0.85 were measured, respectively. The classifier models performed consistently with both, relative (accuracy: 80–84%; AUC-ROC: 0.88–0.95) and absolute (accuracy: 83–87%; AUC-ROC: 0.90–0.95) RS features ($p>0.05$). Similarly, when combined OCT and RS features were used to train classifiers with the two feature sets (RS-OCT(I): relative; RS-OCT(II): absolute), the classifiers' performance measures were not statistically different at $p=0.05$.

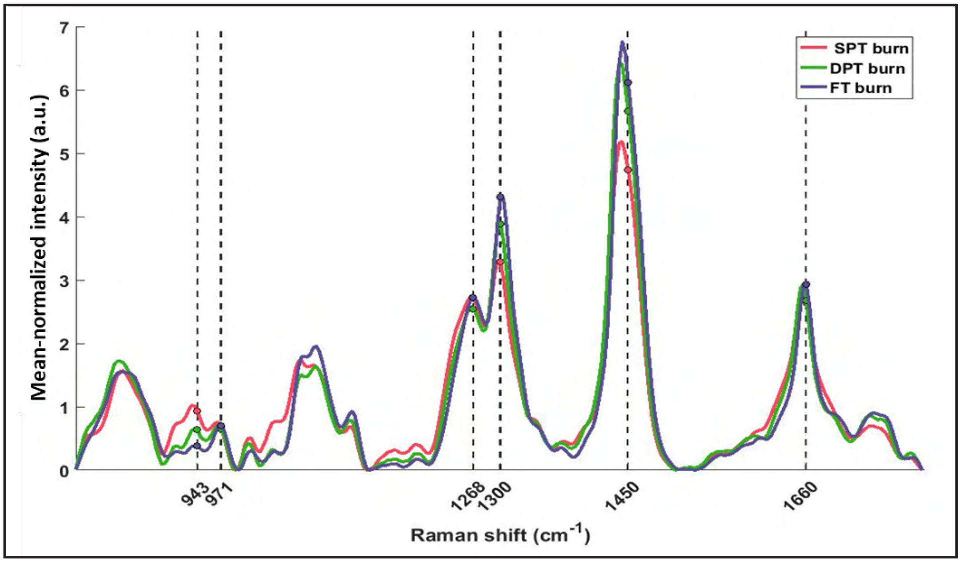


Figure 5. Mean normalized spectra averaged across all spots of different burn severity. Three of the absolute features (Method II) were computed at wavenumbers clinically relevant to skin tissue (marked with dashed lines). Peak intensity ratios were calculated as: 943/971cm^{-1} (N-Cα-C/C-C proline ring), 1,300/1,268cm^{-1} (C-H bending/Amide III), and 1,450/1,660cm^{-1} (CH_2 bending/ Amide I – C=O stretch).

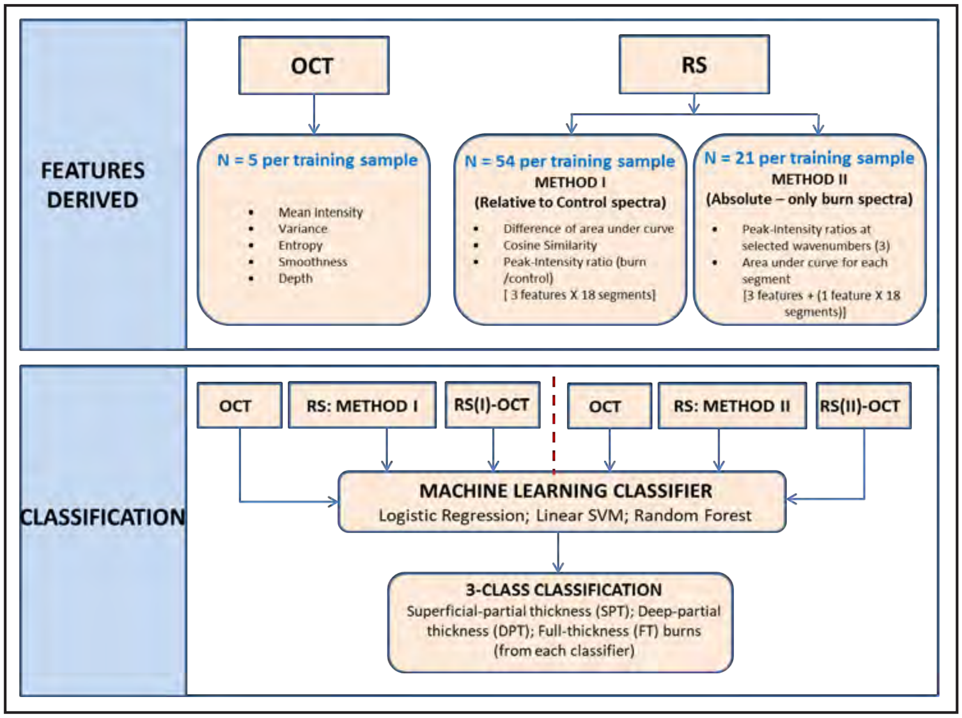


Figure 6. The Features Derived segment denotes all the features derived from OCT images and Raman spectra are listed in the box. The Classification segment denotes the identification of burn degree using three selected classifiers. Each model was trained with five sets of features: OCT, RS (I), RS (II), RS (I)-OCT, and RS (II)-OCT.

Discussion

Our primary aim through the ex vivo study was to demonstrate the ability of an RS-OCT bimodal system to reliably identify the severity of burn wounds. Toward this effort, we achieved a streamlined analytical procedure to process data acquired from the two modalities. We also developed a classification model using ML algorithms to identify burn severity.

The average achieved accuracy of 85% and AUC-ROC=0.94, recorded using a combined pool of RS and OCT features across all classifiers, succinctly indicates the reliability of the bimodal system. Among the three classifiers, LR and RF (both, AUC-ROC=0.92) performed moderately higher than linear SVM (AUC-ROC=0.86) with all types of input features (RS, OCT, RS-OCT). The separability measure using OCT image features was 0.83, whereas RS features alone provided a higher measure of 0.93—thus demonstrating the significance of the biochemical aspects of skin tissue.

Further, equivalent performance measures observed between relative and absolute RS features (average AUC-ROC across all classifiers: RS (I) = 0.92; RS (II) = 0.93) indicate that the features computed only from burn regions are capable of effectively revealing details pertaining to burn severity. As a result, measures from control regions can be disregarded, thereby reducing computational cost. Although RS alone classifies burn wounds with good accuracy, OCT features cannot be excluded, as it is evident from the repetitive appearance of OCT features (namely, depth and entropy) among the top 10 features picked by various feature selection techniques on WEKA. In addition, the inelastic scattering of RS has weak characteristic rendering it as a point-and-shoot method, which is compensated by the ability of OCT scans to cover a wider region. This underscores the value of combining the two optical techniques.

Outlook

In future work, we aim to perform in vivo studies to account for potential influencing factors, such as presence of blood flow, hair, occurrence of edema, and respiration. Additionally, to enhance the bimodal system's capability, we plan to develop a unified RS-OCT probe, where the Raman spectra will be acquired based on the guidance provided by OCT. A compact system will also be easy to use without needing specific training or skill.

Toward this effort, effective registration and data fusion algorithms will be developed and current processing methods will be modified accordingly. Furthermore, decision-level ML algorithms, where results from individual modalities are fused to classify burns, will be explored and optimized to improve the accuracy of the system [40]. As part of an advanced diagnostic system, UC also plans to include decision-based intervention support sub-systems that can provide information on the type of treatment to be given based on burn severity.

In conclusion, satisfactory discernibility established by the ex vivo study signifies the effectiveness of the bimodal diagnostic system in objectively identifying severity of burns and fosters future in vivo studies. Such a system which can efficiently discern burns, especially superficial-partial and deep-partial burns, will be a promising diagnostic tool for medics on the battlefield that will aid in prompt decision-making on treatment steps, thereby reducing mortality rates and improving warfighter burn injury treatment.

References

1. Centers for Disease Control and Prevention. (2019, January 18). Leading causes of nonfatal injury records, 2000–2017. U.S. Department of Health and Human Services. Retrieved from https://webappa.cdc.gov/sasweb/ncipc/nfilead.html
2. Chung, K. K., Blackbourne, L. H., Wolf, S. E., White, C. E., Renz, E. M., Cancio, L. C., ... & Barillo, D. J. (2006). Evolution of burn resuscitation in operation Iraqi freedom. *Journal of Burn Care & Research, 27*(5), 606–611.doi:10.1097/01.BCR.0000235466.57137.f2
3. Kauvar, D. S., Wolf, S. E., Wade, C. E., Cancio, L. C., Renz, E. M., & Holcomb, J. B. (2006). Burns sustained in combat explosions in Operations Iraqi and Enduring Freedom (OIF/OEF explosion burns). *Burns, 32*(7), 853–857. doi:10.1016/j.burns.2006.03.008
4. Gray, D. T., Pine, R. W., Harnar, T. J., Marvin, J. A., Engrav, L. H., & Heimbach, D. M. (1982). Early surgical excision versus conventional therapy in patients with 20 to 40 percent burns: a comparative study. *The American Journal of Surgery, 144*(1), 76–80. doi:10.1016/0002-9610(82)90605-5
5. Sjöberg, F., Danielsson, P., Andersson, L., Steinwall, I., Zdolsek, J., Östrup, L., & Monafo, W. (2000). Utility of an intervention scoring system in documenting effects of changes in burn treatment. *Burns, 26*(6), 553-559. doi:10.1016/S0305-4179(00)00004-8
6. Burmeister, D. M., Ponticorvo, A., Yang, B., Becerra, S. C., Choi, B., Durkin, A. J., & Christy, R. J. (2015, September). Utility of spatial frequency domain imaging (SFDI) and laser speckle imaging (LSI) to non-invasively diagnose burn depth in a porcine model. *Burns, 41*(6), 1242–1252. doi:10.1016/j.burns.2015.03.001
7. Rylander, M. N. (personal communication with HDIAC, May 7, 2019). Used with permission.
8. Atiyeh, B. S., Gunn, S. W. A., & Hayek, S. N. (2007). Military and civilian burn injuries during armed conflicts. *Annals of Burns and Fire Disasters, 20*(4), 203. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3188083/
9. Heimbach, D. M., Afromowitz, M. A., Engrav, L. H., Marvin, J. A., & Perry, B. (1984). Burn depth estimation—man or machine. *The Journal of Trauma, 24*(5), 373–378.
10. Riordan, C. L., McDonough, M., Davidson, J. M., Corley, R., Perlov, C., Barton, R., ... & Nanney, L. B. (2003). Noncontact laser Doppler imaging in burn depth analysis of the extremities. *The Journal of Burn Care & Rehabilitation, 24*(4), 177–186. doi:10.1097/01.BCR.0000075966.50533.B0
11. Monstrey, S., Hoeksema, H., Verbelen, J., Pirayesh, A., & Blondeel, P. (2008). Assessment of burn depth and burn wound healing potential. *Burns, 34*(6), 761–769. doi:10.1016/j.burns.2008.01.009
12. Hoeksema, H., Van de Sijpe, K., Tondu, T., Hamdi, M., Van Landuyt, K., Blondeel, P., & Monstrey, S. (2009). Accuracy of early burn depth assessment by laser Doppler imaging on different days post burn. *Burns, 35*(1), 36–45. doi:10.1016/j.burns.2008.08.011
13. Kaiser, M., Yafi, A., Cinat, M., Choi, B., & Durkin, A. J. (2011). Noninvasive assessment of burn wound severity using optical technology: A review of current and future modalities. *Burns, 37*(3), 377-386. doi:10.1016/j.burns.2010.11.012
14. Droog, E. J., Steenbergen, W., & Sjöberg, F. (2001). Measurement of depth of burns by laser Doppler perfusion imaging. *Burns, 27*(6), 561–568. doi:10.1016/S0305-4179(01)00021-3
15. Resch, T. R., Drake, R. M., Helmer, S. D., Jost, G. D., & Osland, J. S. (2014). Estimation of burn depth at burn centers in the United States: a survey. *Journal of Burn Care & Research, 35*(6), 491–497. doi:10.1097/BCR.0000000000000031
16. Ganapathy, P., Tamminedi, T., Qin, Y., Nanney, L., Cardwell, N., Pollins, A., ... & Yadegar, J. (2014). Dual-imaging system for burn depth diagnosis. *Burns, 40*(1), 67–81. doi:10.1016/j.burns.2013.05.004
17. Olsen, J., Themstrup, L., & Jemec, G. B. (2015). Optical coherence tomography in dermatology. *Giornale Italiano di Dermatologia e Venereologia, 150*(5), 603–615.
18. Silver, R., Helms, A., Fu, W., Wang, H., Diaconu, D., Loyd, C. M., ... & Ward, N. L. (2012). Using optical coherence tomography for the longitudinal non-invasive evaluation of epidermal thickness in a murine model of chronic skin inflammation. *Skin Research and Technology, 18*(2), 225–231. doi:10.1111/j.1600-0846.2011.00558.x
19. Welzel, J. (2001). Optical coherence tomography in dermatology: A review. *Skin Research & Technology, 7*(1), 1–9. doi:10.1034/j.1600-0846.2001.007001001.x
20. Pierce, M. C., Sheridan, R. L., Park, B. H., Cense, B., & de Boer, J. F. (2004). Collagen denaturation can be quantified in burned human skin using polarization-sensitive optical coherence tomography. *Burns, 30*(6), 511–517. doi:10.1016/j.burns.2004.02.004
21. Themstrup, L., De Carvalho, N., Nielsen, S. M., Olsen, J., Ciardo, S., Schuh, S., ... & Jemec, G. B. (2018). In vivo differentiation of common basal cell carcinoma subtypes by microvascular and structural imaging using dynamic optical coherence tomography. *Experimental Dermatology, 27*(2), 156–165. doi:10.1111/exd.13479
22. Gong, P., Es'haghian, S., Wood, F. M., Sampson, D. D., & McLaughlin, R. A. (2016). Optical coherence tomography angiography for longitudinal monitoring of vascular changes in human cutaneous burns. *Experimental Dermatology, 25*(9), 722–724. doi:10.1111/exd.13053
23. Hettiaratchy, S., & Dziewulski, P. (2004). Pathophysiology and types of burns. *The BMJ, 328*(7453), 1427–1429. doi:10.1136/bmj.328.7453.1427
24. Caspers, P. J., Lucassen, G. W., Wolthuis, R., Bruining, H. A., & Puppels, G. J. (1999). In vitro and in vivo Raman spectroscopy of human skin. *Biospectroscopy, 4*(S5), S31–S39. doi:10.1002/(SICI)1520-6343(1998)4:5+<S31::AID-BSPY4>3.0.CO;2-M
25. Lui, H., Zhao, J., McLean, D., & Zeng, H. (2012). Real-time Raman spectroscopy for in vivo skin cancer diagnosis. *Cancer Research, 72*(10), 2491–2500. doi:10.1158/0008-5472.CAN-11-4061
26. Caspers, P. J., Bruining, H. A., Puppels, G. J., Lucassen, G. W., & Carter, E. A. (2001). In vivo confocal Raman microspectroscopy of the skin: noninvasive determination of molecular concentration profiles. *Journal of Investigative Dermatology, 116*(3), 434–442. doi:10.1046/j.1523-1747.2001.01258.x
27. Lieber, C. A., Majumder, S. K., Ellis, D. L., Billheimer, D. D., & Mahadevan-Jansen, A. (2008, September). In vivo nonmelanoma skin cancer diagnosis using Raman microspectroscopy. *Lasers in Surgery and Medicine, 40*(7): 461–467. doi:10.1002/lsm.20653
28. Peterson, J. R., Okagbare, P. I., De La Rosa, S., Cilwa, K. E., Perosky, J. E., Eboda, O. N., ... & Wang, S. C. (2013). Early detection of burn induced heterotopic

ossification using transcutaneous Raman spectroscopy. *Bone, 54*(1), 28–34. doi:10.1016/j.bone.2013.01.00
29. Patil, C. A., Kirshnamoorthi, H., Ellis, D. L., van Leeuwen, T. G., & Mahadevan-Jansen, A. (2011). A clinical instrument for combined raman spectroscopy-optical coherence tomography of skin cancers. *Lasers in Surgery and Medicine, 43*(2), 143–151. doi:10.1002/lsm.21041
30. Liu, C. H., Qi, J., Lu, J., Wang, S., Wu, C., Shih, W. C., & Larin, K. V. (2015). Improvement of tissue analysis and classification using optical coherence tomography combined with Raman spectroscopy. *Journal of Innovative Optical Health Sciences, 8*(4), 1550006. doi:10.1142/S1793545815500066
31. Ponticorvo, A., Burmeister, D. M., Yang, B., Choi, B., Christy, R. J., & Durkin, A. J. (2014). Quantitative assessment of graded burn wounds in a porcine model using spatial frequency domain imaging (SFDI) and laser speckle imaging (LSI). *Biomedical Optics Express, 5*(10), 3467–3481. doi:10.1364/BOE.5.003467
32. Thomsen, S. (1991). Pathologic analysis of photothermal and photomechanical effects of laser–tissue interactions. *Photochemistry and Photobiology, 53*(6), 825–835. doi:10.1111/j.1751-1097.1991.tb09897.x
33. Chvapil, M., Speer, D. P., Owen, J. A., & Chvapil, T. A. (1984). Identification of the depth of burn injury by collagen stainability.

Priya Ganapathy, Ph.D.
Senior Research and Development Scientist, UtopiaCompression Corporation

Dr. Ganapathy is a senior research and development scientist at UtopiaCompression Corporation (Ph.D., Wright State University, Dayton, OH). She has worked extensively as PI on multiple biomedical projects (NIH/DoD grants) related to development of image processing and machine learning algorithms for traumatic brain injury detection, burn injury diagnosis, blood loss, and physiological monitoring. She has collaborated extensively with biomedical device developers and clinicians from multiple academic university-affiliated hospitals in the past six years for complete hardware and software integration. She is also instrumental in transitioning in-house training assessment technologies to military training platforms.

Lakshmi Priya Rangaraju
Research and Development Engineer, UtopiaCompression Corporation

Ms. Rangaraju earned her M.S. in bioengineering from University of Illinois at Chicago (UIC). Her thesis work concentrated on the application of optical coherence tomography to diagnose diabetic retinopathy. She has also worked as Graduate Research Assistant at UIC, performing retinal image processing and analysis. At UtopiaCompression, she has worked on developing image processing algorithms, 3D modelling, and testing machine learning models for biomedical projects on diagnosis of burn wound and traumatic brain injury. She is also engaged in graphical user interface development.

Gautam Kunapuli, Ph.D.
Research Associate Professor, Department of Computer Science, University of Texas at Dallas (UTD)

Gautam Kunapuli is a research associate professor in the Department of Computer Science at UTD and is part of the STAtistical Relational LearnING (Starling) Lab. He has over 10 years of experience in machine learning and optimization, specifically human-in-the-loop learning, knowledge-based and advice-taking learning algorithms, scalable learning algorithms for hard machine-learning problems, and applications in medical informatics. Before joining UTD, he was Senior R&D Scientist at UtopiaCompression Corporation, applying state-of-the-art machine learning in diverse domains including medical informatics, social networking, text analytics, behavior mining, and educational data mining. He received his Ph.D. in mathematics from Rensselaer Polytechnic Institute in 2008 and was a postdoctoral research associate at the University of Wisconsin-Madison from 2008-2013.

Oscar D. Ayala, Ph.D.
Scientist, Cynosure

Dr. Oscar D. Ayala earned his B.S. in biomedical engineering from the University of Texas at Austin in 2012 and his Ph.D. in biomedical engineering from Vanderbilt University in 2019. Dr. Ayala was a National Defense Science and Engineering Graduate (NDSEG) Fellow while at Vanderbilt where he utilized Raman spectroscopy to characterize bacteria to improve clinical diagnostics of infectious diseases. Dr. Ayala is currently a scientist at Cynosure, a Hologic company, where he studies the effects of energy-based technologies for medical aesthetic applications.

Jacob Yadegar, Ph.D.
CEO, UtopiaCompression Corporation

Dr. Jacob Yadegar, CEO, UtopiaCompression Corporation (UC) has over 30 years of R&D experience in academia and industry where he led projects on computational intelligence. Before founding UC, he served as Chief Scientist and VP of R&D for Woosh, Mohomine, and Orincon in the US. At UC, he has generated computational intelligence technical proposals and acted as PI for numerous federal projects worth over $30M. Dr. Yadegar served as a leading scientist at British Telecom Research Lab where he led $2M projects in network management using distributed AI. He served on the faculty at Imperial College of Science, Technology and Medicine – University of London, and the Open University, and post-doctoral at University of Cambridge. He has supervised M.S. and Ph.D. students and published numerous papers at conferences, universities, and research centers. Dr. Yadegar has contributed to and is one of the holders of several patents at UC and one with British Telecom.

Not pictured: Dayna Every and Anita Mahadevan-Jansen

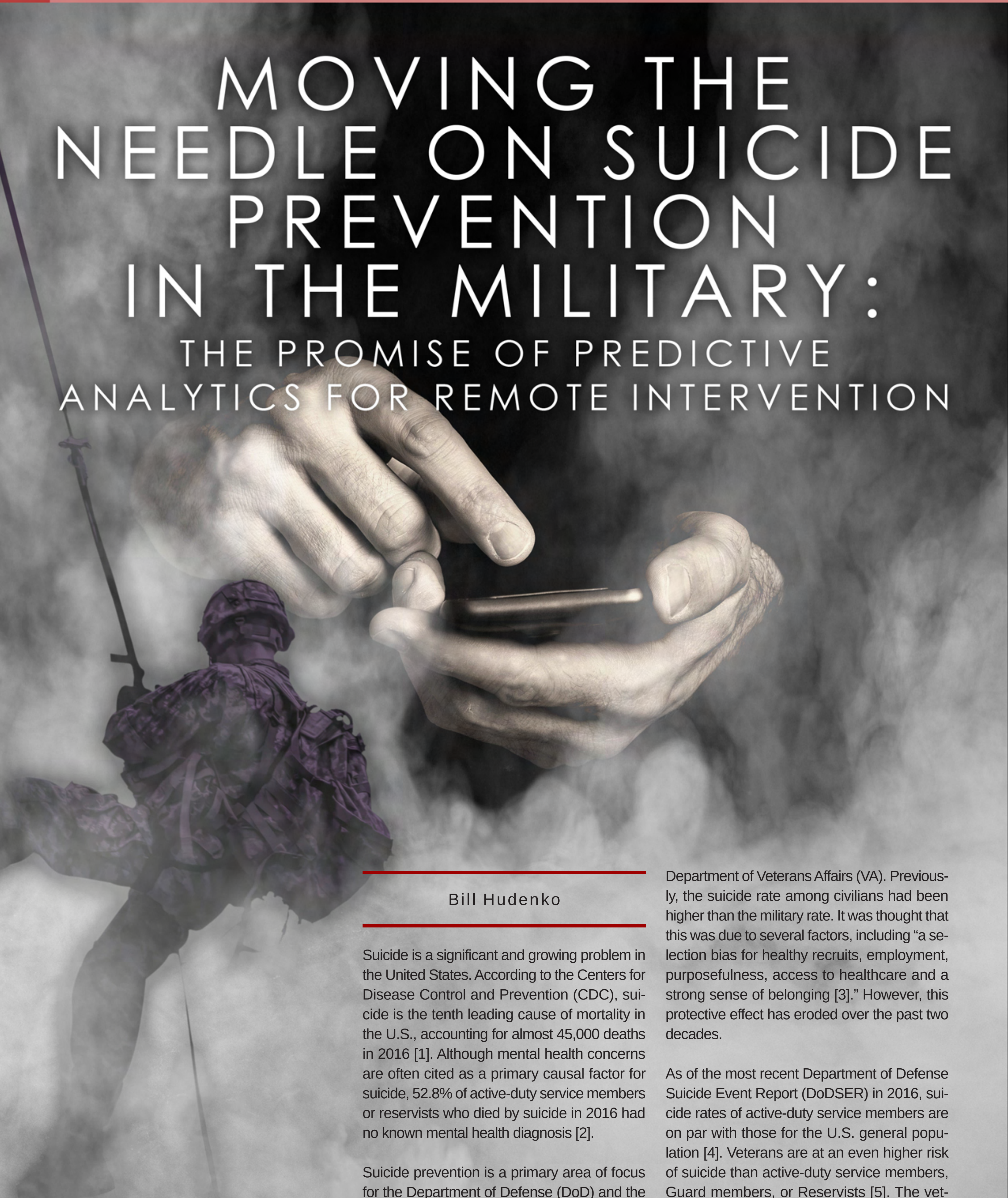
Discover the value of sharing your DoD-funded research...

R&E Gateway Powered by DTIC

https://go.usa.gov/xEe7P

MOVING THE NEEDLE ON SUICIDE PREVENTION IN THE MILITARY:

THE PROMISE OF PREDICTIVE ANALYTICS FOR REMOTE INTERVENTION



Bill Hudenko

Suicide is a significant and growing problem in the United States. According to the Centers for Disease Control and Prevention (CDC), suicide is the tenth leading cause of mortality in the U.S., accounting for almost 45,000 deaths in 2016 [1]. Although mental health concerns are often cited as a primary causal factor for suicide, 52.8% of active-duty service members or reservists who died by suicide in 2016 had no known mental health diagnosis [2].

Suicide prevention is a primary area of focus for the Department of Defense (DoD) and the Department of Veterans Affairs (VA). Previously, the suicide rate among civilians had been higher than the military rate. It was thought that this was due to several factors, including "a selection bias for healthy recruits, employment, purposefulness, access to healthcare and a strong sense of belonging [3]." However, this protective effect has eroded over the past two decades.

As of the most recent Department of Defense Suicide Event Report (DoDSER) in 2016, suicide rates of active-duty service members are on par with those for the U.S. general population [4]. Veterans are at an even higher risk of suicide than active-duty service members, Guard members, or Reservists [5]. The vet-

Image Credit | *Photo illustration created by HDIAC and adapted from a U.S. National Guard photo by Jarod Dye (available for viewing at: https://www.nationalguard.mil/Resources/Image-Gallery/News-Images/igphoto/2001643610/) and Adobe Stock.*

eran suicide rate in 2016 was approximately 1.5 times greater than that of the non-veteran general population [5].

Current Solutions

The VA and DoD have allocated significant resources to reduce suicide rates in active-duty and veteran populations. Multidimensional and multidisciplinary approaches to suicide prevention employed by the DoD include mental health awareness training for all military members; mental health treatment referral guidelines; post-suicide intervention programs; life skills training; and increased staffing of mental health providers [6–11].

In addition to bolstering traditional 24/7 call lines such as the VA's Veterans Crisis Line and increasing access to "Telemental" health services, researchers across the country have turned to emerging technologies to facilitate the prevention of suicide. This research has primarily focused on creating machine learning algorithms to identify predictive risk factors in the electronic health records of patients [12, 13]. Other approaches to detect suicide risk include using natural language processing (NLP) of text, social media, or auditory conversations [14]; passive sensing of mobile phone data [15]; and developing machine learning algorithms to evaluate risk from self-reported data [16, 17].

Intervention strategies that employ technology include the use of mobile health applications to reduce suicide rates [18] and the development of algorithms to identify optimal suicide prevention strategies [19]. The military has adopted a number of these strategies to improve suicide risk detection and prevention. For example, initiatives such as the VA's REACH VET, education-based mobile apps, and the VA's "Make the Connection" outreach campaign have all improved access to resources for at-risk individuals. Despite substantial resources devoted to reducing suicide rates in active-duty and veteran populations, suicide rates remain unacceptably high. Though some initiatives have demonstrated efficacy in reducing rates, the critical question remains: why are suicide rates still increasing?

Barriers to Suicide Rate Reduction

Unfortunately, there are substantial barriers to reducing suicide rates in both active-duty and veteran populations. Although a multitude of factors contribute, there are at least four primary challenges that have significantly impeded progress: access to care, stigma, aversion to help-seeking, and timeliness of interventions.

Access to Care

Data from the VA Office of Mental Health and Suicide Prevention indicate that approximately 70% of at-risk veterans who attempt suicide have not visited VA facilities within the year prior to their attempt [20]. Similarly, nearly one quarter of veterans live in remote, rural areas without easy access to support [21]. Practical barriers include geographical location, availability of resources, access to transportation, time constraints, financial limitations, and the general inconvenience of seeking out professional help [22, 23]. These factors may present as non-trivial obstacles that decrease the likelihood of mental health service utilization.

Stigma

One of the most cited barriers to service utilization in the general population is the stigma associated with mental health treatment [22]. People who experience mental and behavioral health symptoms may experience both public stigma and self-stigma [24]. Public stigma consists of negative perceptions held by others about mental health service utilization and about those who experience mental health challenges. This stigma is particularly acute in military populations, due to concerns about leadership, chain of command, and concerns that service utilization will have a negative impact on career advancement or lead to the loss of a security clearance [25, 26].

Self-stigma consists of the individual's personal beliefs about mental health issues and service utilization. If an individual has internalized negative perceptions of mental health distress and treatment, it may lead to diminished self-esteem and create a barrier to seeking needed services [24, 27–29].

Aversion to Help-Seeking

Data suggest that both active-duty and veteran populations are hesitant to seek help. In a recent DoD Office of People Analytics study of 14,088 active-duty service members, of those who had experienced suicidal ideation or a suicide attempt since joining the military, 43.3% reported that they had not sought help for their concerns [30]. Among non-help-seekers, most (70.2%) never considered talking to someone about their suicidal ideation. Factors associated with non-help-seeking included being an officer; being male; not being able to identify suicide risk factors nor knowing how to take appropriate action to help; and being concerned about the career impact of seeking mental health treatment [30].

This aversion to help-seeking is likely due to a military culture that values self-sacrifice [31]. Individuals are encouraged to deny their natural instinct for self-preservation and instead to place their own well-being second to their mission. Therefore, mental health concerns may be viewed as a sign of weakness and a liability to others.

Timeliness of Intervention

Many technology-based solutions developed to reduce suicide rates in the military implement previously discussed methods to detect suicide risk. These include scanning the electronic health record for factors that predict increased risk [12, 13], using NLP of social media to identify risk [14], or identifying biomarkers of suicidal ideation [32].

Unfortunately, very few of these efforts have been linked with intervention methods that might avert imminent suicidal gestures. Though services such as the Veterans Crisis Line now offer texting options in addition to 24/7 telephone availability, all of these services require a common element: the service member must be willing to ask for help. Although many service members and veterans do use help lines, the data suggest that an overwhelming majority of individuals at greatest risk of harm do not seek aid at the time of their attempt [12].

Suicide Reduction: A New Model

We have developed a novel solution that we call "Voi Reach." This approach is designed to address all four of the primary barriers to suicide prevention. Our technology provides the user with an easy-to-use mobile app that connects them with a behavioral health coach, and his or her natural support network (e.g., friends, family) to increase safety. We augment this support network with artificial intelligence (AI) and continuous risk monitoring via NLP.

This application provides the first solution capable of predictively identifying an imminent risk of suicide before the user asks for help. This gives DoD medical staff and other first responders the capacity to send an active rescue within moments of an individual needing help. Perhaps most importantly, our technology circumvents major barriers that have previously stymied progress toward the reduction of suicide rates.

Access to care: The app assists veterans and active-duty service members who are not connected with healthcare services. This is particularly helpful for veterans who live in remote areas, and cannot easily access healthcare services because of geographic or physical constraints. Although the app does require the use of an iOS- or Android-capable smartphone, figures from 2014 indicate that up to 79% of all surveyed active-duty Army personnel [33] and up to 90% of veterans [34] own a smartphone. For service members who are capable of using a smartphone and who accept the approach, they are matched with a trained behavioral health coach who provides 24/7 in-app messaging.

The app is built on a patient engagement messaging platform that has serviced over 200,000 patient-provider exchanges to date. Behavioral health coaches are supported by in-app AI algorithms to increase engagement. Our system utilizes proprietary algorithms to prompt the coach on when and how to reach out to the user to maximize engagement. This elevates the quality of care beyond what is currently available through crisis lines, because the coach is more knowledgeable about how to develop and sustain a meaningful relationship with the user.

Stigma/aversion to help-seeking: Though both public and self-stigma will continue to be cultural barriers to accessing care in the military, we believe that our approach has the potential to substantially reduce such stigma: access to care through our app is more private than what is typically afforded through in-person medical or psychotherapy services. Furthermore, with trained veterans serving as behavioral healthcare coaches, military personnel will have more confidence in shared understanding with care providers.

Perhaps most importantly, however, the app provides veterans or active-duty service members the option to invite natural support network members into their care circle. Through a combination of psychoeducation, AI-directed messaging, and direct intervention by behavioral health coaches, this system seamlessly integrates the care of supporting members into the life of an at-risk individual. Our app is consequently a very scalable healthcare solution because a single behavioral health coach can cover hundreds of military personnel.

Timeliness of intervention: Unlike other systems, Voi Reach does not require a veteran to overtly ask for help, as it predicts imminent risk using AI. This solves one of the biggest problems associated with suicide prevention. The app uses advanced NLP to identify at-risk communications within the platform, or through social media (if the user grants permission). This NLP system is built on a specific type of supervised machine-learning system called genetic programming (i.e., a computerized system that can learn to recognize patterns associated with a known outcome) [35].

The system was constructed by converting the free-text records of veterans with known suicide attempts into word or word-phrase datasets, or numerical counts of how often a given word or phrase appeared in a patient record. The derived models then identified the combination of words that were associated with suicide. The data were then analyzed using a machine-learning algorithm to generate predictive models. This approach allows a remote behavioral health coach to be aware instantly when a service member transitions to high risk. Furthermore, a remote coach can then send out emergency services for an active rescue within five minutes of the status change.

Conclusion

While many technological solutions have been developed to address the problem, suicide rates among active-duty service members and veterans continue to rise. We believe this is due primarily to an inability to stay effectively connected to military personnel and appropri-

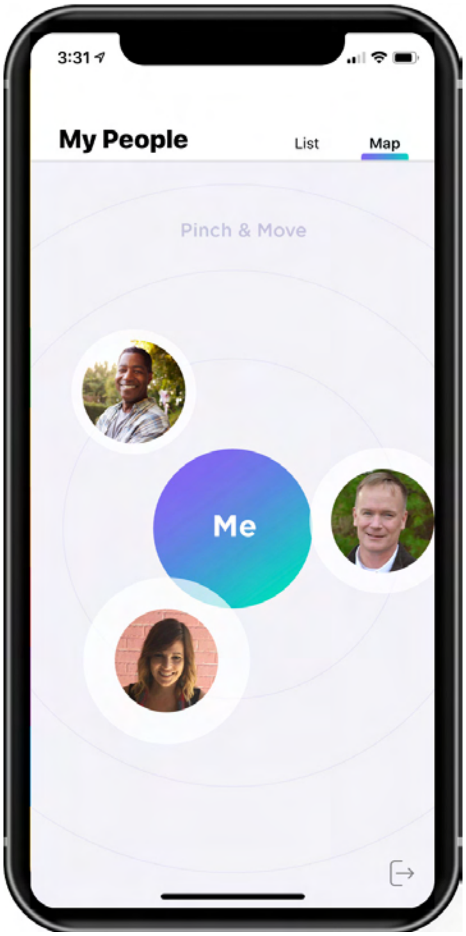


Figure 1. Screenshot from the Voi Reach app.

ately monitor them for risk long-term. Unlike other solutions, our AI-enabled app-based approach is the first to predict imminent risk, which then allows a behavioral health coach to rapidly reach out to a service member after a risk status change. This is in contrast to other approaches that require those in need of help to ask for it—a problem that, unless we overcome it, will inhibit any significant decline in suicide rates. It is our hope that new, novel applications of technology gain acceptance so that we can move the needle on the tragic and preventable loss of life in the military.

References

1. Centers for Disease Control and Prevention. (2017, March 17). Leading causes of death. National Center for Health Statistics. Retrieved from https://www.cdc.gov/nchs/fastats/leading-causes-of-death.htm
2. Department of Defense. (2018, June 20). *Department of Defense Suicide Event Report (DoDSER), Calendar Year 2016 Annual Report*. Office of the Under Secretary of Defense for Personnel and Readiness. Retrieved from https://www.dspo.mil/Portals/113/Documents/DoDSER%20CY%202016%20Annual%20Report_For%20Public%20Release.pdf?ver=2018-07-02-104254-717
3. Department of Veterans Affairs and Department of Defense. (2013, June). *VA/DoD Clinical Practice Guideline for Assessment and Management of Patients at Risk for Suicide (Version 1.0)*. Assessment and Management of Risk for Suicide Working Group. Retrieved from https://www.healthquality.va.gov/guidelines/MH/srb/VADODCP_SuicideRisk_Full.pdf
4. Centers for Disease Control and Prevention. (2018, September 29). Preventing suicide. National Center for Injury Prevention and Control, Division of Violence Prevention. Retrieved from https://www.cdc.gov/violenceprevention/suicide/fastfact.html?CDC_AA_refVal=https%3A%2F%2Fwww.cdc.gov%2Fviolenceprevention%2Fsui-

cide%2Fconsequences.html

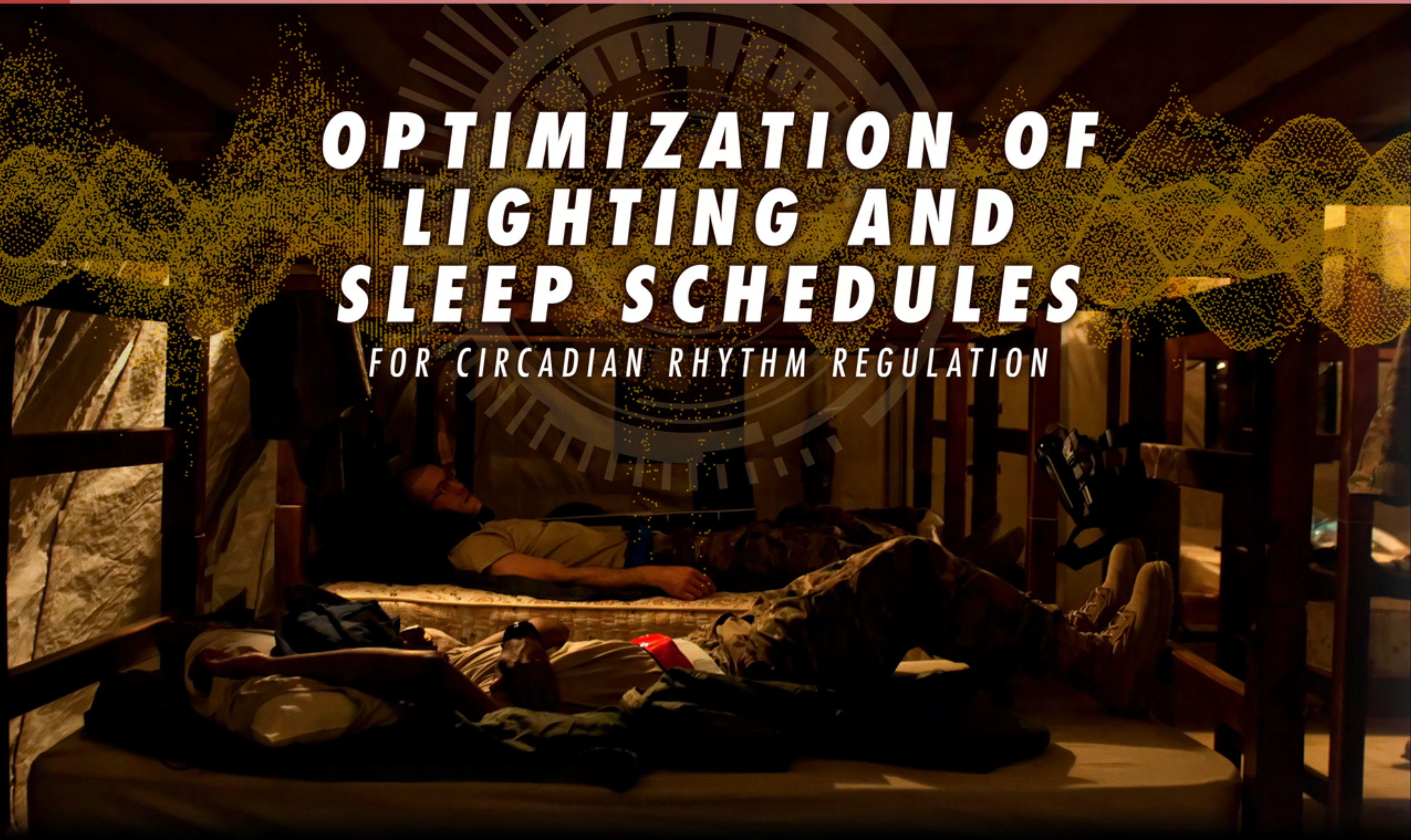
5. Department of Veterans Affairs. (2018, September). *VA National Suicide Data Report 2005–2016*. Office of Mental Health and Suicide Prevention. Retrieved from https://www.mentalhealth.va.gov/docs/data-sheets/OMHSP_National_Suicide_Data_Report_2005-2016_508.pdf
6. Bagley, S. C., Munjas, B., & Shekelle, P. (2010). A systematic review of suicide prevention programs for military or veterans. *Suicide and Life-Threatening Behavior, 40*(3), 257–265. doi:10.1521/suli.2010.40.3.257
7. James, L. C., & Kowalski, T. J. (1996). Suicide prevention in an Army infantry division: A multi-disciplinary program. *Military Medicine, 161*(2), 97–101. doi:10.1093/milmed/161.2.97
8. Jones, D. E., Kennedy, K. R., Hawkes, C., Hourani, L. A., Long, M. A., & Robbins, N. L. (2001). Suicide prevention in the Navy and Marine Corps: Applying the public health model. *Navy Medicine, 92*(6), 31–36.
9. Knox, K. L., Litts, D. A., Talcott, G. W., Feig, J.C., & Caine, E. D. (2003). Risk of suicide and related adverse outcomes after exposure to a suicide prevention programme in the US Air Force: Cohort study. *Bmj, 327*(7428), 1376-0. doi:10.1136/bmj.327.7428.1376
10. McDaniel, W. W., Rock, M., & Grigg, J. R. (1990, April). Suicide prevention at a United States Navy training command. *Military Medicine, 155*(4), 173–175.
11. Medical and Dental Care: Members on Duty Other than Active Duty for a Period of More than 30 Days. 10 U.S.C. § 1074a et seq.
12. Department of Veterans Affairs. (2017, April 3). VA REACH VET initiative helps save veterans lives: Program signals when more help is needed for at-risk veterans. News release. Retrieved from https://www.va.gov/opa/pressrel/includes/viewPDF.cfm?id=2878
13. Kessler, R. C., Hwang, I., Hoffmire, C. A., Mccarthy, J. F., Petukhova, M. V., Rosellini, A. J., . . . Bossarte, R. M. (2017). Developing a practical suicide risk prediction model for targeting high-risk patients in the Veterans Health Administration. *International Journal of Methods in Psychiatric Research, 26*(3). doi:10.1002/mpr.1575
14. Poulin, C., Shiner, B., Thompson, P., Vepstas, L., Young-Xu, Y., Goertzel, B., . . . Mcallister, T. (2014). Predicting the risk of suicide by analyzing the text of clinical notes. *PLoS ONE, 9*(1). doi:10.1371/journal.pone.0085733
15. Allen, N. B., Nelson, B. W., Brent, D., & Auerbach, R. P. (2019). Short-term prediction of suicidal thoughts and behaviors in adolescents: Can recent developments in technology and computational science provide a breakthrough? *Journal of Affective Disorders, 250*, 163–169. doi:10.1016/j.jad.2019.03.044
16. Bernecker, S. L., Zuromski, K. L., Gutierrez, P. M., Joiner, T. E., King, A. J., Liu, H., . . . Kessler, R. C. (2018, December). Predicting suicide attempts among soldiers who deny suicidal ideation in the Army Study to Assess Risk and Resilience in Servicemembers (Army STARRS). *Behaviour Research and Therapy*. doi:10.1016/j.brat.2018.11.018
17. Gradus, J. L., King, M. W., Galatzer-Levy, I., & Street, A. E. (2017). Gender Differences in Machine Learning Models of Trauma and Suicidal Ideation in Veterans of the Iraq and Afghanistan Wars. *Journal of Traumatic Stress, 30*(4), 362–371. doi:10.1002/jts.22210
18. Jaroszewski, A. C., Morris, R. R., & Nock, M. K. (2019). Randomized controlled trial of an online machine learning-driven risk assessment and intervention platform for increasing the use of crisis services. *Journal of Consulting and Clinical Psychology, 87*(4), 370–379. doi:10.1037/ccp0000389
19. Vaiva, G., Walter, M., Arab, A. S., Courtet, P., Bellivier, F., Demarty, A. L., . . . Libersa, C. (2011). ALGOS: The development of a randomized controlled trial testing a case management algorithm designed to reduce suicide risk among suicide attempters. *BMC Psychiatry, 11*(1). doi:10.1186/1471-244x-11-1
20. Department of Veterans Affairs. (2018, June). Facts about veteran suicide: June 2018. Office of Mental Health and Suicide Prevention. Retrieved from https://www.mentalhealth.va.gov/docs/FINAL_VA_OMHSP_Suicide_Prevention_Fact_Sheet_508.pdf
21. Holder, K. A. (2017, January). Veterans in rural America: 2011–2015. U.S. Department of Commerce, Economics and Statistics Administration. Retrieved from https://www.census.gov/content/dam/Census/library/publications/2017/acs/acs-36.pdf
22. Hom, M. A., Stanley, I. H., & Joiner, T. E. (2015). Evaluating factors and interventions that influence help-seeking and mental health service utilization among suicidal individuals: A review of the literature. *Clinical Psychology Review, 40*, 28–39. doi:10.1016/j.cpr.2015.05.006
23. Mojtabai, R., Olfson, M., Sampson, N. A., Jin, R., Druss, B., Wang, P. S., ... & Kessler, R. C. (2010). Barriers to mental health treatment: Results from the National Comorbidity Survey Replication. *Psychological Medicine, 41*(8), 1751–1761. doi:10.1017/s0033291710002291
24. Corrigan, P. W., & Watson, A. C. (2002, February). Understanding the impact of stigma on people with mental illness. *World Psychiatry, 1*, 16–20. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1489832/
25. Vogt, D. (2011). Mental health-related beliefs as a barrier to service use for military personnel and veterans: A review. *Psychiatric Services, 62*(2), 135–142. doi:10.1176/ps.62.2.pss6202_0135
26. Zinzow, H. M., Britt, T. W., Pury, C. L., Raymond, M. A., Mcfadden, A. C., & Burnette, C. M. (2013). Barriers and facilitators of mental health treatment seeking among active-duty Army personnel. *Military Psychology, 25*(5), 514–535. doi:10.1037/mil0000015
27. Corrigan, P. (2004). How stigma interferes with mental health care. *American Psychologist, 59*(7), 614–625. doi:10.1037/0003-066x.59.7.614
28. Pattyn, E., Verhaeghe, M., Sercu, C., & Bracke, P. (2014). Public stigma and self-stigma: Differential association with attitudes toward formal and informal help seeking. *Psychiatric Services, 65*(2), 232–238. doi:10.1176/appi.ps.201200561
29. Pietrzak, R. H., Johnson, D. C., Goldstein, M. B., Malley, J. C., & Southwick, S. M. (2009). Perceived stigma and barriers to mental health care utilization among OEF-OIF veterans. *Psychiatric Services, 60*(8), 1118–1122. doi:10.1176/ps.2009.60.8.1118
30. Ho, T., Hesse, C.M., Osborn, M.M., Schneider, K.G., Smischney, T.M., Carlisle, B.L., Beneda, J.G., & Schwerin, M.J. (2018, July). *Mental Health and Help-Seeking in the U.S. Military: Survey and Focus Group Findings* (PERSEREC-TR_18-10, OPA-2018-048). Department of Defense, Defense Personnel and Security Research Center, Office of People Analytics. Retrieved from https://www.dhra.mil/Portals/52/Documents/perserec/reports/TR-18-10%20Estimating%20and%20Understanding%20the%20Non-Help%20Seeking%20Population%20in.pdf
31. Bryan, C. J., Jennings, K. W., Jobes, D. A., & Bradley, J. C. (2012). Understanding and Preventing Military Suicide. *Archives of Suicide Research, 16*(2), 95–110. doi:10.1080/13811118.2012.667321
32. Bossarte, R., Claassen, C. A., & Knox, K. (2010). Veteran Suicide Prevention: Emerging Priorities and Opportunities for Intervention. *Military Medicine, 175*(7), 461–462. doi:10.7205/milmed-d-10-00050
33. Mercado, J.E., & Spain, R.D. (2014, July). *Evaluating Mobile Device Ownership and Usage in the U.S. Army: Implications for Army Training* (Research Report #1974). United States Army Research Institute for the Behavioral and Social Sciences. Retrieved from https://apps.dtic.mil/dtic/tr/fulltext/u2/a603886.pdf
34. Mcinnes, D. K., Fix, G. M., Solomon, J. L., Petrakis, B. A., Sawh, L., & Smelson, D. A. (2015). Preliminary needs assessment of mobile technology use for healthcare among homeless veterans. *PeerJ*, 3. doi:10.7717/peerj.1096
35. Looks, M. (2007, July). Meta-optimizing semantic evolutionary search. *Proceedings of the 9th Annual Conference on Genetic and Evolutionary Computation*. Retrieved from https://dl.acm.org/citation.cfm?doid=1276958.1277086

Bill Hudenko, Ph.D.
Chief Executive Officer, Voi
Assistant Professor of Psychiatry and Adjunct Assistant Professor of Psychological & Brain Sciences, Dartmouth College

Dr. Bill Hudenko has significant experience in the fields of both mental health and technology. Dr. Hudenko is a licensed psychologist, a researcher, and a professor who holds a joint appointment as a faculty member at Dartmouth's Geisel School of Medicine and Dartmouth's Department of Psychological and Brain Sciences. Dr. Hudenko received his B.A. from the University of Michigan and his Ph.D. in Clinical Psychology from Vanderbilt. His research focuses on the use of technology to improve mental health delivery and patient outcomes. He is also an experienced web designer who has served as a software engineer and database administrator for the National Center for Post-Traumatic Stress Disorder, and has designed and implemented websites for numerous other organizations. Dr. Hudenko is the former CEO of Incente, LLC. Prior to his work at Dartmouth, Dr. Hudenko was a professor at both Ithaca College and Cornell University. Dr. Hudenko has broad expertise in clinical psychology with an emphasis in child psychopathology and family systems.

OPTIMIZATION OF LIGHTING AND SLEEP SCHEDULES

FOR CIRCADIAN RHYTHM REGULATION



A. Agung Julius,
Jiawei Yin,
& John T. Wen

The human circadian rhythm evolved as a mechanism by which living beings can synchronize their biological processes with the light and dark pattern of the terrestrial day. The circadian rhythm effectively functions as a master clock that regulates these processes. In humans, the circadian rhythm is tightly linked to various physiological processes, such as sleep, metabolism, and neuro-behavioral processes [1, 2].

Disruption of the circadian rhythm is known to have negative impacts on health, ranging from fatigue in travelers with jetlag to an increased risk of cardiac disease and cancer in rotating shift workers [3, 4]. Furthermore, malalignment of the circadian phase and related neurobehavioral states (such as alertness) with the timing of critical tasks may lead to lower performance and higher risk of failure.

The sleep process in humans is very tightly connected to the circadian rhythm. The sleep drive, for example, is known to be modulated by the circadian rhythm [5–7]. Sleep is critical to health, as it allows the body to recuperate and regenerate cells. It is also tied to neurocognitive performance; the lack of or mistiming of sleep have been empirically linked to degeneration of neurocognitive performance [8, 9] and disruption of circadian rhythm regulation [10, 11].

Many Department of Defense (DoD) missions necessarily disrupt natural circadian cycles and sleep. For example, airlift and tanker missions can impair reaction times, lower attention spans, diminish memory recall, and increase human errors and elevate risk [12]. Sleep duration is also found to affect cognitive performance [13]. Other research has shown that sleep cycle disruption may be more of a contributing cause than a symptom of PTSD experience in veterans [14]. Additionally, research conducted by the U.S. Naval Submarine Medical Research Laboratory, concludes that special operations forces may be able to phase-lock their circadian rhythms when traveling across time zones, via judicious control of light exposure [15].

In the circadian rhythm research literature, several mathematical models have been developed to capture the dynamics of circadian rhythm and sleep. These models are typically formulated to match experimental results from human subject studies. These models allow us to make quantitative predictions about, for example, the impact of lighting to a subject's circadian rhythm, or the effects of circadian rhythm and sleep scheduling on a subject's neurobehavioral states, such as alertness.

The models typically incorporate some parameters, such as the free-running period of the subject's circadian rhythm, which may be adapted to individual subjects, leading to personalized circadian and sleep health solutions. In this article, we present a novel approach to mathematically modeling circadian rhythm and sleep activity and regulation.

Mathematical Modeling

Light is a strong synchronizer of the circadian rhythm. Generically, the circadian rhythm models are ordinary differential equations of the form

$$dx/dt = F(x,u), \qquad (1)$$

where x is the state of the circadian oscillator, and u represents the impact of light on the circadian oscillator. Because light can only attain non-negative intensity and its impact to the cir-

Image Credit | *Photo illustration created by HDIAC and adapted from a U.S. Air Force photo by David Salanitri (available for viewing at: https://www.afcent.af.mil/News/Photos/igphoto/2000153843/) and Adobe Stock.*

cadian oscillator has a saturation regime, u is typically constrained as:

$$0 \leq u(t) \leq u_{max}, \quad (2)$$

for all $t \geq 0$.

These models come in various orders (i.e., levels of complexity). The highest order models could contain up to several dozens of state variables. These are typically based on the biochemical processes in the circadian gene regulation [16, 17]. Lower order models are typically empirical, such as the third order model by Kronauer et al. [18, 19], which uses a van der Pol oscillator to represent the circadian rhythm of core body temperature. A first order model is used to represent the dynamics of the phase of the circadian rhythm, which is a good approximation of the higher order dynamics under lighting conditions with low variation (e.g., a dark or dimly lit environment) [20].

The dynamics of the sleep and neurobehavioral states are coupled to that of the circadian rhythm. Their relationship is described in a family of models called the *two-process models*. These models link the dynamics of sleep drive and alertness to the circadian phase in a hybrid or multi-modal dynamical system:

$$\frac{ds}{dt} = \begin{cases} G_s(s), \text{if the subject is asleep,} \\ G_a(s), \text{if the subject is awake.} \end{cases} \quad (3)$$

Here, s denotes the sleep and neurobehavioral states. Note that the dynamics of $s(t)$ depends on whether the subject is asleep or awake. The transitions between the sleep and awake modes are determined by sleepiness. Sleepiness is quantified as variable $B(t)$ and modeled as a function of the circadian and sleep states:

$$B(t) = H(x(t),s(t)). \quad (4)$$

The switching between sleep and wakefulness can be modeled as spontaneous, with the subject going to sleep when sleepiness ($B(t)$) reaches a certain high level (B_{max}) and waking-up when it reaches a certain low level (B_{min}). An alternative approach is to assume that these switches can be scheduled in a constrained manner. For example, the subject can be scheduled to go to sleep when $B(t)$ is between B_{max} and B^H_{max} and has to go to sleep when $B(t)$ reaches B^H_{max} for the schedule to avoid excessive sleep deprivation. Conversely, the subject can be scheduled to wake up when $B(t)$ is between B_{min} and B^H_{min} and will wake up spontaneously when $B(t)$ reaches B^H_{min}. This is illustrated in Figure 1. The sleep state impacts the circadian state dynamics because the light stimuli cannot act on the subject while asleep.

Alertness, quantified as $A(t)$, is modeled as a function of the sleep and neurobehavioral states [21–24]:

$$A(t) = K(x(t),s(t)). \quad (5)$$

The factors that impact alertness in this model are sleep debt (whether the subject is lacking sleep), sleep inertia (whether the subject just recently gained wakefulness), and circadian phase (the subject's biological clock).

Optimal Regulation of the Circadian and Sleep Process

Regulation of the circadian rhythm has been the subject of a great deal of research in recent years. From the perspective of biomathematics, this problem is typically expressed as an optimal control problem of a system with nonlinear dynamics. One of the main problems in optimal circadian rhythm regulation is the *quickest entrainment problem*, discussed below.

Quickest Entrainment Problem (QEP): Given the dynamics of the circadian state in Equations (1–2), an initial condition $x(0) = x_0$, and a reference circadian state trajectory $x_{ref}(t)$, determine the input $u(t)$ that minimizes the entrainment time T. Here, the entrainment time T is defined as the earliest time such that $x(T) = x_{ref}(T)$.

The quickest entrainment problem can capture, for example, the problem of eliminating jetlag as quickly as possible. In this case, the reference $x_{ref}(t)$ is the desired circadian state trajectory (without the jetlag).

QEP can be solved using optimal control techniques, resulting in both *open-loop* optimal lighting strategies [25, 26], where the optimal

Papers	Problem Addressed				Model Used†
	QEP	QEP-A	QEP-S	AOP	
[20]	√				1st order model of circadian phase dynamics
[25] [26] [28]	√				2nd or 3rd order Kronauer circadian model
[29]	√				high order models of circadian dynamics
[30]		√			1st order model of circadian phase dynamics with 2nd order model of sleep dynamics
[31]		√	√		3rd order Kronauer circadian model with 2nd order model of sleep dynamics
[27]				√*	2nd order Kronauer circadian model with 2nd order model of sleep dynamics

Table 1. Summary of papers that address the problems described in this article

† circadian dynamics refers to Eq. (1), sleep dynamics refers to Eq. (3)
** without the sleep scheduling constraints*

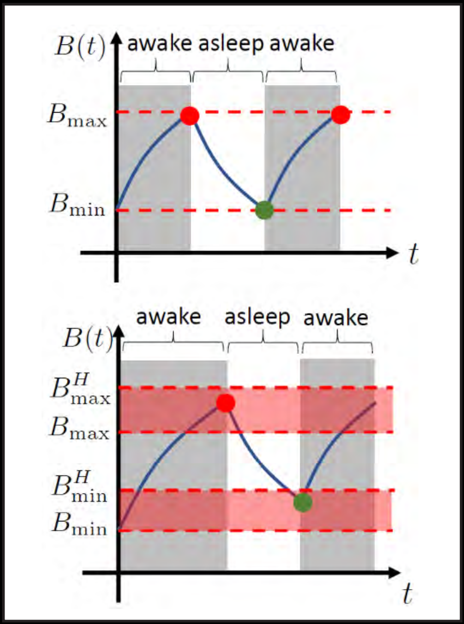


Figure 1. Illustrations of the spontaneous (top) and scheduled (bottom) sleep-wake switching. Red and green circles represent the times when the subject goes to sleep and wakes up, respectively.

light signal is computed offline, and *feedback* optimal lighting strategies, where the optimal light signal is computed as a function of the current circadian state [20]. Note that the formulation of QEP ignores the dynamics of the sleep and neurobehavioral states ($s(t)$). Therefore, the resulting optimal solutions may be impractical from the sleep perspective, as they may require the subject to stay awake for an extended period of time regardless of their sleepiness. To address this limitation, QEP can be modified to include the sleep dynamics. There are two variations of the new formulation, one where sleep is considered autonomous, and one where sleep can be scheduled, as discussed in the previous section.

Quickest Entrainment Problem with Autonomous Sleep (QEP-A): Given the dynamics of the circadian, sleep, and neurobehavioral states in Equations (1–4), where the switching of the sleep dynamics is autonomous, an initial condition $x(0) = x_0$, $s(0) = s_0$, and a reference circadian state trajectory $x_{ref}(t)$, determine the input $u(t)$ that minimizes the entrainment time T. Here, the entrainment time T is defined as the earliest time such that $x(T) = x_{ref}(T)$.

Quickest Entrainment Problem with Schedulable Sleep (QEP-S): Given the dynamics of the circadian, sleep, and neurobehavioral states in Equations (1–4), an initial condition $x(0) = x_0$, $s(0) = s_0$, and a reference circadian state trajectory $x_{ref}(t)$, determine the input $u(t)$ and the sleep schedule that meets the scheduling constraint, such that the entrainment time T is minimized. Here, the entrainment time T is defined as the earliest time such that $x(T) = x_{ref}(T)$.

QEP-A and QEP-S are more challenging to solve compared to QEP because the associated systems dynamics (a) have higher orders, and (b) is non-smooth because of the switching. Further, QEP-S is more challenging to solve compared to QEP-A due to the fact that QEP-S has more optimization variables and constraints (i.e., the sleep schedule).

Nevertheless, we recently developed a functional gradient descent method to solve these problems (samples of the solutions can be seen in Figure 2). The circadian rhythm model used to obtain these results are the third order Kronauer model [19]. The trajectories of the two circadian oscillator states and their respective references (with the subscript r) are shown in each graph. We can observe that, as expected, the solutions of QEP take the least amount of entrainment time compared to those of QEP-A and QEP-S. On the other hand, the solutions of QEP-A take the longest entrainment time because of the imposed constraint due to the autonomous sleep assumption.

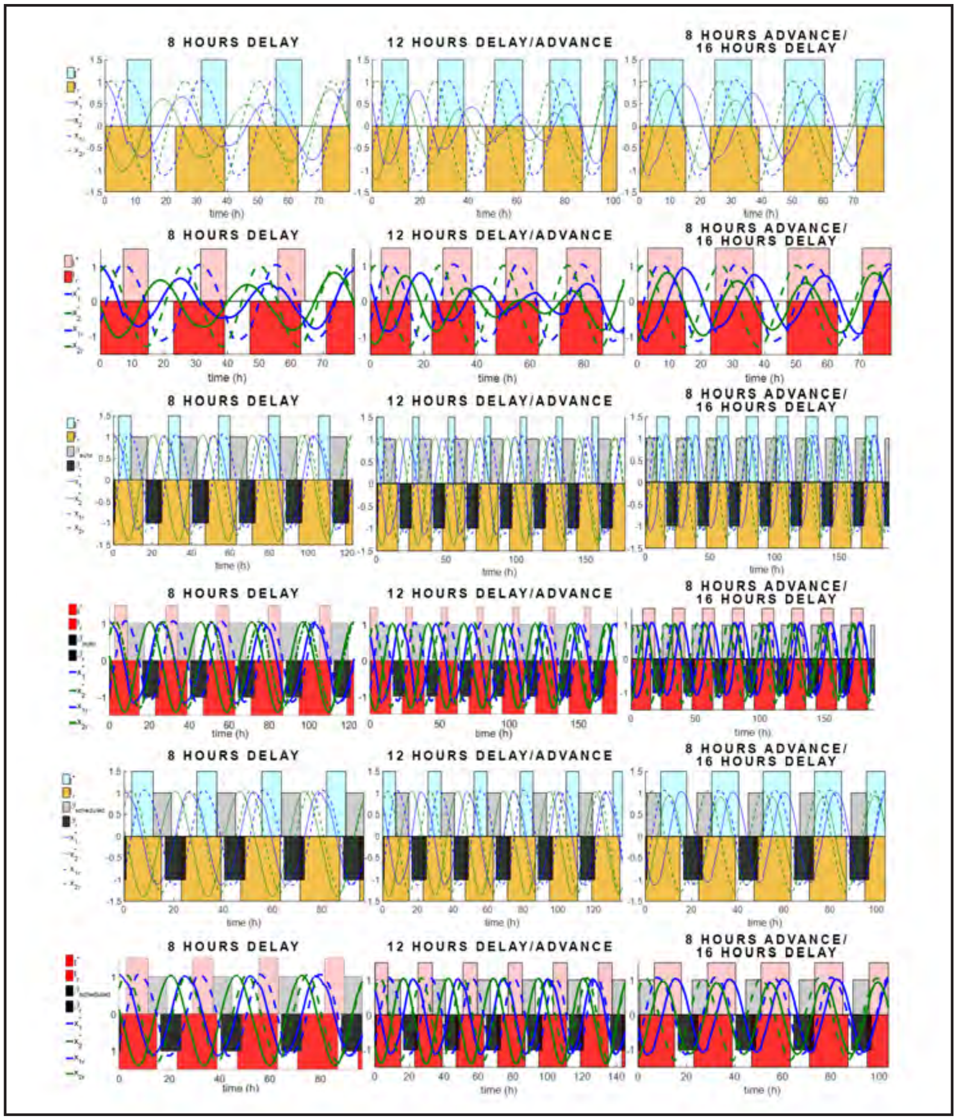


Figure 2. Solutions of QEP (top row), QEP-A (middle row), and QEP-S (bottom row), for various jetlag cases. All simulations are run until entrainment. Pink bands represent when the optimal control input (lighting) is on. Red bands represent when the natural (reference) daylight is on (assuming 16h/8h light/dark pattern). Light gray bands represent the sleep intervals of the subject. Dark gray bands represent the sleep intervals of the reference.

Another type of optimal control problems that can be formulated and solved for this system is the alertness optimization problem. The idea is to have the subject be maximally alert during a prescribed time interval, e.g., between given T_1 and T_2. Such a problem can be formulated as follows.

Alertness Optimization Problem (AOP): Given the dynamics of the circadian, sleep, and neurobehavioral states in Equations (1–5), an initial condition $x(0) = x_0$, $s(0) = s_0$, determine the input $u(t)$ and the sleep schedule that meets the scheduling constraint, such the minimum alertness A_{min} is maximized. Here, A_{min} is defined as

$$A_{\min} \triangleq \min_{T_1 \le t \le T_2} A(t).$$

AOP can also be solved using a functional gradient descent method [27].

Conclusion

Joint mathematical models for the circadian, sleep, and neurobehavioral processes are important tools for addressing problems related to the regulation of circadian rhythm, sleep, and optimization of alertness. These problems are related to practical DoD interests, such as the quick(est) elimination of jetlag and optimization of alertness in warfighters. Mathematically, they can be formulated as optimal control problems, for which there are solution techniques based on the calculus of variations.

Further research into circadian rhythm activity and sleep cycle regulation may help DoD maximize warfighter readiness and improve the force's overall health and resilience to the rigors of military training and deployment.

Acknowledgements

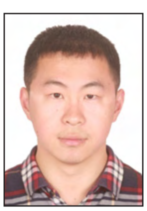
The research in this article was supported by the National Science Foundation through the Smart Lighting Engineering Research Center (EEC-0812056) and in part by the U.S. Army Research Office through grant number W911NF-17-1-0562 and by New York State under NYSTAR contract C130145. The authors would like to thank Dr. Jiaxiang Zhang and Dr. Wei Qiao for their contributions to some of the results reported in this article, and Dr. Daniel Forger for valuable discussions related to optimal entrainment of circadian rhythms.

References

1. Smolensky, M. H., & Lamberg, L. (2001). *The body clock guide to better health: How to use your body's natural clock to fight illness and achieve maximum health*. New York: H. Holt.
2. Foster, R. G., & Kreitzman, L. (2017). *Circadian rhythms: A very short introduction*. Oxford, United Kingdom: Oxford University Press.
3. Knutsson, A. (2003). Health disorders of shift workers. *Occupational Medicine, 53*(2), 103–108. doi:10.1093/occmed/kqg048
4. Figueiro, M. G., & White, R. D. (2013). Health consequences of shift work and implications for structural design. *Journal of Perinatology, 33*(S1). doi:10.1038/jp.2013.7
5. Borbély, A. A. (1982). A two process model of sleep regulation. *Human Neurobiology, 1*(3), 195–204.
6. Daan, S., Beersma, D. G., & Borbely, A. A. (1984). Timing of human sleep: Recovery process gated by a circadian pacemaker. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology, 246*(2). doi:10.1152/ajpregu.1984.246.2.r161
7. Achermann, P. (2004, March). The two-process model of sleep regulation revisited. *Aviation, Space, and Environmental Medicine, 75*(3), A37–A43.
8. Durmer, J. S., & Dinges, D. F. (2005). Neurocognitive consequences of sleep deprivation. *Seminars in Neurology, 25*(01), 117–129. doi:10.1055/s-2005-867080
9. Alhola, P., & Polo-Kantola, P. (2007). Sleep deprivation: Impact on cognitive performance. *Neuropsychiatric Disease and Treatment, 3*(5), 553–567.
10. Archer, S. N., Laing, E. E., Möller-Levet, C. S., Veen, D. R., Bucca, G., Lazar, A. S., . . . Dijk, D. (2014). Mistimed sleep disrupts circadian regulation of the human transcriptome. *Proceedings of the National Academy of Sciences, 111*(6). doi:10.1073/pnas.1316335111
11. Archer, S. N., & Oster, H. (2015). How sleep and wakefulness influence circadian rhythmicity: Effects of insufficient and mistimed sleep on the animal and human transcriptome. *Journal of Sleep Research, 24*(5), 476–493. doi:10.1111/jsr.12307
12. Criss, A G. (2018). *Circadian rhythm disruption: A comparative analysis of enumeration for the mobility Air Force*. Graduate Research Paper, Air Force Institute of Technology, Wright-Patterson AFB, Ohio. Retrieved from https://apps.dtic.mil/dtic/tr/fulltext/u2/1064133.pdf
13. Richards, A. (2017). *When sleep isn't perfect: Risk and resilience for cognitive consequences of imperfect sleep duration and suboptimal timing of tasks in circadian rhythm*. U.S. Army Medical Research and Materiel Command (Report W81XWH-16-1-0259). Retrieved from https://apps.dtic.mil/dtic/tr/fulltext/u2/1057801.pdf
14. Kilgore, W. D. (2015). A *nonpharmacologic method for enhancing sleep in PTSD*. U.S. Army Medical Research and Materiel Command (Report W81XWH-14-1-0570). Retrieved from https://apps.dtic.mil/dtic/tr/fulltext/u2/1007313.pdf
15. Chabal, S., Couturier, K., Dyche, J., Soutiere, S., Figuiero, M., & Plitnick, B. (2018 January). *Circadian rhythm phase locking for traveling special forces operators: Using light exposure to maintain time zone entrainment*. Naval Submarine Medical Research Laboratory (Technical Report NSMRL/F1209/TR-2018-1318). Retrieved from https://apps.dtic.mil/dtic/tr/fulltext/u2/1054520.pdf.
16. Leloup, J., & Goldbeter, A. (2003). Toward a detailed computational model for the mammalian circadian clock. *Proceedings of the National Academy of Sciences, 100*(12), 7051–7056. doi:10.1073/pnas.1132112100
17. Becker-Weimann, S., Wolf, J., Kramer, A., & Herzel, H (2004). A model of the mammalian circadian oscillator including the EV-ERBα module. *Genome Informatics*, 15(4), 3–12. Retrieved from https://www.jstage.jst.go.jp/article/gi1990/15/1/15_1_3/_pdf
18. Kronauer, R. E., Forger, D. B., & Jewett, M. E. (2000). Quantifying human circadian pacemaker response to brief, extended, and repeated light stimuli over the phototopic range. *Journal of Biological Rhythms, 14*(6), 500–515.
19. Jewett, M. E., Forger, D. B., & Kronauer, R. E. (1999). Revised limit cycle oscillator model of human circadian pacemaker. *Journal of Biological Rhythms, 14*(6) 493–499.
20. Qiao, W., Wen, J. T., & Julius, A. A. (2017). Entrainment control of phase dynamics. *IEEE Transactions on Automatic Control, 62*(1), 445–450.
21. Achermann, P., & Borbely, A. A. (2003). Mathematical models of sleep regulation. *Frontiers in Bioscience, 8*, 683–693.
22. Jewett, M. E., & Kronauer, R. E. (1999). Interactive mathematical models of subjective alertness and cognitive throughput in humans. *Journal of Biological Rhythms, 14*(6), 588–597.
23. Rajdev, P., Thorsley, D., Rajaraman, S., Rupp, T. L., Wesensten, N. J., Balkin, T. J., & Reifman, J. (2013). A unified mathematical model to quantify performance impairment for both chronic sleep restriction and total sleep deprivation. *Journal of Theoretical Biology, 331*, 66–77. doi:10.1016/j.jtbi.2013.04.013
24. Ramakrishnan, S., Lu, W., Laxminarayan, S., Wesensten, N. J., Rupp, T. L., Balkin, T. J., & Reifman, J. (2015). Can a mathematical model predict an individual's trait-like response to both total and partial sleep loss? *Journal of Sleep Research, 24*(3), 262–269. doi:10.1111/jsr.12272
25. Zhang, J., Qiao, W., Wen, J. T., & Julius, A. A. (2016). Light-based circadian rhythm control: Entrainment and optimization. *Automatica, 68*, 44–55. doi:10.1016/j.automatica.2016.01.052
26. K. Serkh and D. B. Forger, "Optimal schedules of light exposure for rapidly correcting circadian misalignment," PLOS Computational Biology, vol. 10, no. 4, p. e1003523, 2014.
27. Serkh, K., & Forger, D. B. (2014). Optimal schedules of light exposure for rapidly correcting circadian misalignment. *PLoS Computational Biology, 10*(4). doi:10.1371/journal.pcbi.1003523
28. Zhang, J., Wen, J. T., & Julius, A. A. (2013). Optimal and feedback control for light-based circadian entrainment. *52nd IEEE Conference on Decision and Control*. doi:10.1109/cdc.2013.6760287
29. Yin, J., Julius, A. A., & Wen, J. T. (2019, March). Rapid circadian entrainment in models of circadian genes regulation. arXiv:1902.09059. Retrieved from https://arxiv.org/pdf/1902.09059.pdf
30. Julius, A. A., Yin, J., & Wen, J. T. (2017). Time-optimal control for circadian entrainment for a model with circadian and sleep dynamics. *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*. doi:10.1109/cdc.2017.8264355
31. Julius, A. A., Yin, J., & Wen, J. T. (n.d.). *Optimization of light exposure and sleep schedule for circadian entrainment*. Unpublished manuscript.

A. Agung Julius, Ph.D.
Associate Professor of Electrical, Computer, and Systems Engineering, Rensselaer Polytechnic Institute (RPI)

A. Agung Julius is an associate professor in the Department of Electrical, Computer, and Systems Engineering at RPI. He earned his Ph.D. in applied mathematics from the University of Twente (Netherlands) in 2005. Dr. Julius' research interests include systems and control, systems biology, stochastic models in systems biology, control of biological systems, hybrid systems, and mathematical systems theory. At Rensselaer, he received the School of Engineering Award for Research Excellence (2016) and the James M. Tien'66 Early Career Award for Faculty (2017). He also received an NSF CAREER award in 2010.



Jiawei Yin
Ph.D. Candidate and Research Assistant, Rensselaer Polytechnic Institute (RPI)

Jiawei Yin received his B.S. and M.S. in marine engineering from Harbin Engineering University (China) in 2012 and 2015, respectively. He is currently working toward his Ph.D. in electrical engineering, and working as a research assistant in the Light Enabled Systems and Application (LESA) Engineering Research Center at RPI. His research interests include circadian modeling and entrainment, nonlinear control and optimization, and noise and vibration control.

John Wen, Ph.D.
Head of Electrical, Computer, and Systems Engineering, Rensselaer Polytechnic Institute (RPI)

Dr. John Wen was with Fisher Controls from 1981-1983, Jet Propulsion Laboratory from 1985-1988, and has since been the Head of Electrical, Computer, and Systems Engineering with RPI. He was the Director for the Center for Automation Technologies and Systems from 2005-2013, Head of Industrial & Systems Engineering from 2013-2018, ASEE/NASA Summer Faculty Fellow in 1993, Japan Society for the Promotion of Science Senior Visiting Scientist in 1997, and Oversea Assessor of the Chinese Academy of Sciences from 2004-2009. He was awarded the IEEE Control Systems Society Transition to Practice Award in 2013, and has been an IEEE Fellow since 2001.

Calendar of Events

October 2019

October 8 - 10 • **National Harbor, MD**
HDS Defense TechConnect Fall Summit and Expo

October 14 - 16 • **Washington, DC**
HDS AUSA Annual Meeting and Exposition

October 22 - 24 • **Annapolis, MD**
HDS 24th Annual Expeditionary Warfare Conference

October 23 - 24 • **Alexandria, VA**
HDS Insider Threat Symposium

November 2019

November 5 - 7 • **Nashville, TN**
CBRN CBRNe Convergence 2019

November 6 - 7 • **Alexandria, VA**
HDS 7th Annual EOD/IED & Countermine Symposium

November 18 - 21 • **Cincinnati, OH**
HDS 2019 CBD S&T Conference

November 19 - 21 • **Norfolk, VA**
CBRN 11th US EPA International Decontamination R&D Conference

December 2019

December 10 - 11 • **Alexandria, VA**
HDS Digital Forensics for National Security

December 10 - 12 • **San Diego, CA**
M, CBRN 5th Annual Sensors Summit

December 11 - 12 • **Alexandria, VA**
HDS Special Operations and Irregular Warfare Symposium

Call for Papers

HDIAC **is now accepting abstracts and articles for consideration for future publications. For more information, contact the Publications Team at** publications@hdiac.org.

The HDIAC Journal is a quarterly publication, focusing on novel developments and technology in the Alternative Energy, Biometrics, CBRN Defense, Critical Infrastructure Protection, Cultural Studies, Homeland Defense and Security, Medical, and Weapons of Mass Destruction focus areas.

Articles must be relevant to one of the eight focus areas and relate to Department of Defense applications.

Articles should be submitted electronically as a Microsoft Word document.

We require a maximum of 2,500 words.

All submissions must include graphics or images (300 DPI or higher in JPG or PNG format) to accompany the article. Photo or image credit should be included in the caption.