Objective Pain Identification and Monitoring for Fighter Pilots

Rebecca Kwasinski, Jeffrey Hullfish, Adam Lynch, Brent Winslow
Design Interactive, Inc.
Orlando, FL
Rebecca.Kwasinski@designinteractive.net, Jeff.Hullfish@designinteractive.net, Adam.Lynch@designinteractive.net, Brent.Winslow@designinteractive.net

Adam Faurot
SPEAR Human
Tallahassee, FL
adam@spearhhp.com

Mitchell Ruble
Hylaine
Charlotte, NC
mruble@hylaine.com

ABSTRACT

Fighter pilots face unique occupational stressors, including extreme gravitational forces, long and stressful missions, and rigorous physical and mental training. The physical demands can take their toll on the pilot's body, as evidenced by the numerous health issues such as neck and back problems that have plagued the fighter pilot population for years. To help address this problem, several groups within the United States Air Force (USAF) are implementing wellness and human performance optimization programs. Most approaches are centered around increasing physical therapy and wellness efforts. While implementing these programs will undoubtedly make a positive impact towards the goal of rapid recovery and full operating capacity, there is an opportunity to create and implement artificial intelligence (AI) algorithms that can collect, analyze, store and present objective pain data. This would deliver critical information to provide better insights and enhance fighter pilots' rehabilitation decisions while also providing a unique and personalized approach for each fighter pilot.

An unobtrusive physiological pain classifier was created for use by the USAF fighter pilot community by first collecting physiological measures (electrocardiogram [ECG]) from healthy, adult (N=41) participants during baseline and pain-induction tasks. These raw ECG signals were used to derive a series of cardiovascular features including time domain, frequency domain, and non-linear features. Using logistic regression, these features classified pain at an accuracy level of 79.6%. Field data collection is currently underway with the 56th Training Squadron at Luke Air Force Base (AFB) to determine classification accuracy and ruggedness in operational environments. For this effort, the classifier was integrated with a smartwatch and mobile application for classification in on- and off-duty environments. This data will ultimately assist the pilots and their medical staff in building a more robust, individualized physical therapy program.

ABOUT THE AUTHORS

Ms. Rebecca Kwasinski is a Data Scientist at Design Interactive, Inc. (DI) and has over five years of expertise in non-invasive physiological research, data collection, data processing, and data analysis. Rebecca earned her bachelor's degree in biomedical engineering with a minor in chemistry from Florida International University where she worked as a research fellow in the Optical Imaging Laboratory focusing on near-infrared imaging of lower extremity ulcers to provide a quantitative assessment of the wound healing process. Upon gradution, Rebecca worked as a Project Engineer for ECRI where she worked to develop and execute test protocols and publish key findings to evaluate a wide variety of medical technologies based on the device performance, safety, workflow, interoperability, user/patient experience, cybersecurity, and maintenance. At DI, Rebecca focuses on utilizing non-invasive physiological sensor data to develop classifiers of the human state.

Dr. Jeffrey Hullfish is a Senior Data Scientist and the Portfolio Manager for Digital Twins R&D at Design Interactive, Inc. (DI). He has over seven years of experience studying signals from the peripheral and central nervous systems. His primary research focus is reverse engineering the structure and function of the human brain via the study of neurological and psychiatric disorders. He led a clinical pilot study in collaboration with Interacoustics A/S, a Denmark-based medical device company, to improve the early detection of peripheral nerve tumors using auditory brainstem responses; this resulted in his successful master's thesis. He went on to earn his Ph.D. from the University of Texas at Dallas, where he studied the neuropathology of tinnitus using functional magnetic resonance imaging (fMRI). At DI, he works primarily with physiological data collected non-invasively from human subjects. Using these data, he develops and implements algorithms for the real-time classification of cognitive and physiological status; recent examples include both stress and pain.

Mr. Adam Lynch is a Senior Project Engineer at DI and has over 9 years of project management experience in both hardware manufacturing and software development. Adam has over 17 years of military experience as an officer in the Marine Corps Reserve and is currently serving as the Battalion Commander of 4th Air Naval Gunfire Liaison Company.

Dr. Brent Winslow is Chief Scientist at DI and has over 15 years of experience in rehabilitation, training, biomedical engineering, and applied neuroscience. Dr. Winslow developed patented approaches to human stress quantification in natural environments, and has tested such approaches in law enforcement, first response, special operations, and with active duty service members. Dr. Winslow is active in advisory councils and advocacy groups from brain injury to bleeding disorders, has published articles from in vivo electrophysiology to military training, and has presented work in various international bioengineering, neuroscience, and imaging conferences.

Mr. Adam Faurot is the Chief Commercial Officer (CCO) at SPEAR Human Performance, and has been a key influencer in the Human and Economic Performance Optimization industry. After finishing a career in professional baseball with the Milwaukee Brewers and Boston Red Sox organizations, Adam co-founded TITUS and has been instrumental in both services and solutions including the largest childhood obesity program in the U.S., a pioneering human factors data platform-SPEAR, impacting thousands of athletes, public safety and military personnel.

Mr. Mitchell Ruble is a consultant in delivery services where he works with a team to provide application development and business intelligence recommendations. Mitchell has a strong background in rehabilitation with over 6 years of experience where he has led software development efforts that focused on gamification of rehabilitation including a myoelectric pre-prosthetic training system for upper limb amputees and a rehabilitation-based game for children with cerebral palsy. He has also led several efforts focused on mobile application development including a mobile tool for detecting stress and providing stress mitigation strategies in cognitive behavioral therapy and developing a classifier for detecting learner engagement in classroom settings.

Objective Pain Identification and Monitoring for Fighter Pilots

Rebecca Kwasinski, Jeffrey Hullfish, Adam Lynch, Brent Winslow
Design Interactive, Inc.
Orlando, FL
Rebecca.Kwasinski@designinteractive.net, Jeff.Hullfish@designinteractive.net, Adam.Lynch@designinteractive.net, Brent.Winslow@designinteractive.net

Adam Faurot
SPEAR Human
Tallahassee, FL
adam@spearhhp.com

Mitchell Ruble
Hylaine
Charlotte, NC
mruble@hylaine.com

BACKGROUND

Fighter pilots face unique occupational stressors, including extreme gravitational forces, long and stressful missions, and rigorous physical and mental training. The physical demands can take their toll on the pilot's body, as evidenced by the numerous health issues such as neck and back problems that have plagued the fighter pilot population for years (Kikukawa, Tachibana, and Yagura 1995). The lifetime prevalence rate of reported overall spinal (including cervical, thoracic, and lumbar) disorders among fighter pilots has been found to range from 89% (Kikukawa, Tachibana, and Yagura 1995) to 93% (Rintala et al. 2015). When the quantity of acute musculoskeletal pain episodes was investigated (Kikukawa, Tachibana, and Yagura 1995), it was found that nearly one third of fighter pilots (average age= 33 years) reported more than 10 episodes of musculoskeletal pain during their career, and the mean recovery time for each episode was eight days.

Musculoskeletal disorders may lead to temporary (Knudson et al. 1988) or permanent (Mccrary, Bf, and Van Svoc 1995) flight disqualification and thereby affect a pilot's career and result in the loss of predicted working years. Musculoskeletal disorders represent one of the most common reasons for permanent medical flight disqualification among United States Air Force (USAF) pilots and navigators after cardiovascular and neurological disorders (Mccrary, Bf, and Van Svoc 1995).The most common diagnostic categories in the group of musculoskeletal disorders are chronic (neck or back) pain or discus related (herniated nucleus pulposus) problems. Early career limitations and, in the worst-case scenario, permanent flight disqualification affect squadrons' human resources and operational capability. As such, anecdotal evidence suggests that many pilots conceal or otherwise do not report pain, leading to exacerbation of the injury. Fully trained fighter pilots transferred to desk jobs or to flying non-high-performance aircraft (NHPA) are a loss for the USAF , both economically and operationally. It is therefore important to detect the pain a fighter pilot experiences in an effort to identify injuries early and rehabilitate them (Honkanen et al. 2019).

Along these lines, in June 2019, Gen. Mike Holmes of Air Combat Command (ACC) disclosed that USAF is looking for ways to incorporate more preventative medical care for pilots, including athletic trainers, massage therapists, strength and conditioning coaches, and physical therapists in fighter units and squadrons. ACC is investigating personalized methods to prevent and rehabilitate Service members through overall wellness improvements and early management of neck and back pain. To help address this problem, several groups within USAF are implementing wellness and human performance optimization programs. Most approaches are centered around increasing physical therapy efforts and wellness personnel. The 56th Operational Medical Readiness Squadron at Luke Air Force Base (AFB) has developed the Tactical Integrated Training and Nutrition (TITAN) arena for fighter pilots, a one-of-a-kind training facility with a dedicated human performance team focusing on enhancing mission capabilities and providing quality of life measures for pilots in and out of the cockpit. While implementing these programs will undoubtedly make a positive impact towards the goal of full operating capacity, there is a unique opportunity to create and implement AI algorithms that can collect, analyze, store and present objective pain data. This would deliver critical information to provide better insights and enhance fighter pilots' treatment decisions while also providing a unique and personalized approach for each fighter pilot.

Pain is an unpleasant sensory and emotional experience, which is induced by noxious stimuli that are detected by nociceptive neurons (Raja et al. 2020). Both forms of pain are associated with stereotypical physiological alterations (Cowen et al. 2015) including increasing cardiovascular activity such as heart rate, blood pressure and heart rate variability (Cowen et al. 2015; Koenig et al. 2014; Lowery, Fillingim, and Wright 2003; Streff et al. 2010), respiration rate and depth (Jafari et al. 2017), and electrodermal activity (Ledowski et al. 2007; Loggia, Juneau, and Bushnell 2011). Advances in non-invasive, persistent, physiological monitoring devices may allow for real-time, objective pain sensing and mitigation. In the current study, a physiological dataset associated with pain induction was leveraged to develop an algorithm capable of detecting pain.

METHODS

Experimental Procedure

In order to develop an objective pain algorithm leveraging physiological data, healthy, adult, (N=41) participants were fitted with a 3-lead electrocardiogram (ECG), sampled at 500 Hz, and wirelessly sent to an MP-150 system running AcqKnowledge software (Biopac Systems, Goleta CA). All methods involving participants were approved by an independent Institutional Review Board (Copernicus Group, Durham, NC). The sociodemographic factors of the participants in the initial classifier study are listed in Table 1. Forty-one novice participants [thirty-eight male; average age 21.8 ± 2.4 (standard deviation; SD) years] completed and received payment of $100 USD for participation in the study. All participants were recruited from the community and met minimum requirements including age (18-30), normal visual acuity, and no medical conditions such as endocrine disorders. Among the 11 individuals who did not complete the CPT, the average duration was 2.5 ± 0.80 (SD) min.

Table 1. Sociodemographic factors of the study sample.

	%	n
Gender		
Male	92.7	38
Female	7.3	3
Age Group		
18-21	51.2	21
22-25	39	16
26-30	9.8	4
Education		
High School Diploma	26.8	11
Some College/University	51.2	21
University Degree	21.4	9

Following placement of the ECG, a 5-minute recording of baseline (BL) physiological activity was taken while participants remained seated. Participants then underwent the cold pressor test (CPT), consisting of up to 3 minutes of non-dominant hand immersion in ice cold water under experimenter observation. While the CPT was initially developed as a clinical cardiovascular challenge to monitor changes to heart rate and blood pressure (Streff et al. 2010), it is also considered a reliable experimental method for controlled pain induction (Mitchell, MacDonald, and Brodie 2004). Physiological measures were captured throughout the baseline and pain-induction phases.

Physiological Measurements

The data used for this modeling effort was based on interbeat (RR) intervals calculated from the raw ECG data using AcqKnowledge software. The QRS complex represents the electrical impulse as it spreads through the ventricles in advance of ventricular depolarization. The RR interval is the time interval between consecutive R waves of the QRS signal on the ECG (Figure 1). The R-peak is usually used as the fiducial point due to its readily distinguishable amplitude (Peltola 2012). This point is typically not influenced by noise or low amplitudes as other points on the ECG.

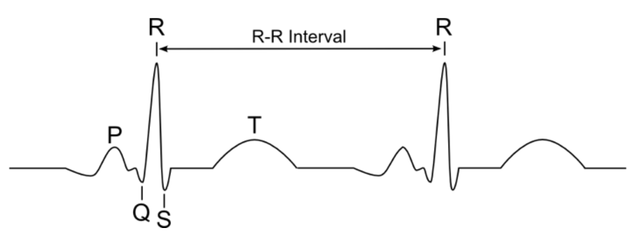

Figure 1. RR interval representation on an ECG

In addition to ECG, photoplethysmography (PPG) can be utilized to derive heart rate measurements by measuring changes in blood volume under the skin (Bent et al. 2020). The integration of PPG in wearable devices such as smartwatches allows for better ease of use in comparison to ECG since it does not require electrodes and wire lead placement on the body. Both ECG and PPG signals contain waves associated with heart cycle

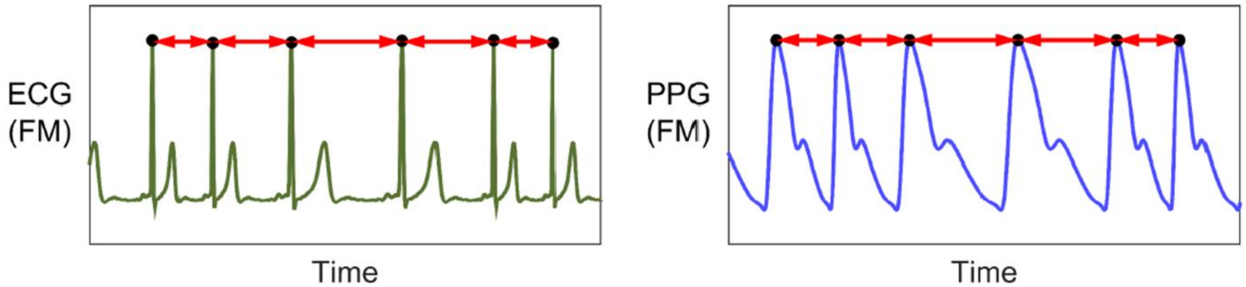

Figure 2. Feature-based extraction techniques for deriving respiration from ECG and PPG including: baseline wander (top); amplitude modulation (middle); or frequency modulation (RSA, bottom). Derived from (Charlton et al. 2018).

information, allowing for derivation of heart rate (bpm), interbeat intervals (s), and time and frequency domain metrics (Figure 2). The ECG RR interval data is comprised of continuous, time-series waveforms (f_s = 500 Hz) where the value changed only when a new ECG R peak is detected.

The RR intervals were separated into 60-second, non-overlapping epochs for reduction of noise on the classification accuracy and to coordinate with the temporal dynamics of pain (Schneider et al. 2018). This RR interval data was analyzed to convert the time-series data into featurized observations for use as model inputs. These features (N= 46) fell into one of four categories: (1) respiration (n=4); (2) time-domain heart rate variability (n=10); (3) frequency-domain heart rate variability (n=7); and (4) nonlinear heart rate variability (n=4). After calculating the features for each epoch, the features were baseline normalized. This was done to reduce the confounding influence of interindividual variability on the classifier. This normalization process began by pairing each participant's BL epochs with all of their other epochs. The features were then subtracted between each of these BL/BL and BL/CPT epoch pairs and the absolute values of the differences were taken. The resulting data that were used as classifier inputs therefore describe the magnitude of the difference between an epoch of data and a physiological baseline for a given participant.

RR interval data was used to derive respiratory information by leveraging the respiratory sinus arrhythmia (RSA). RSA is heart rate variability (HRV) in synchrony with respiration, by which the RR interval on an ECG is shortened during inspiration and prolonged during expiration (Yasuma and Hayano 2004). First, the RR intervals were linearly interpolated to create a uniformly sampled time-series (fs = 4 Hz). This interpolated time-series data was then filtered using a Butterworth bandpass filter in the range of .2–.8 Hz to isolate the frequency components relevant for respiration. Then, a peak detection algorithm was run on the bandpass-filtered signal to detect breaths. The inter-breath intervals—that is, the respiratory equivalent of interbeat intervals (IBI)—were then calculated (in ms) for each of the detected peak/breath. Finally, the IBIs were converted to respiration rate (in breaths per minute). The respiratory features used to create the pain classifier were calculated based on the instantaneous respiration rate data. These included a total of four features, all in the time domain: (1) minimum; (2) maximum; (3) mean and (4) standard deviation of respiration rate, in breaths/min.

An open source toolbox, pyHRV, was utilized to calculate HRV features (Gomes 2021). The toolbox bundles a selection of functions to compute time domain, frequency domain, and nonlinear HRV features. The time domain features contain statistical information from two different classes: (a) features derived directly from the RR intervals or instantaneous heartrate, (b) those derived from the differences between successive RR intervals. An overview of the time domain features that were calculated is presented in Table 2.

Frequency domain analysis started with the computation of a power spectral density (PSD), followed by splitting the PSD into frequency bands and deriving a series of overall features related to specific frequency bands including very low frequency (VLF; 0–.04 Hz), low frequency (LF; .04–.15 Hz), high frequency (HF; .15–.40 Hz), and very high frequency (VHF; .40–3.0 Hz). Segmenting the PSD into specific frequency bands provides a method to determine dominant frequencies in an RR interval series, which are ultimately linked to autonomic responses of a subject's body. An overview of the frequency domain features that were calculated are presented in Table 3.

Feature Engineering

Table 2. HRV time-domain features

Feature	Unit	Description
RRI	*ms*	RR Interval features (min, max, mean, max, count)
Δ RRI	*ms*	Successive RR interval differences features (min, max, mean)
HR	*bpm*	Heart rate features (min, max, min)
SDRR	*ms*	Standard deviation of the RR intervals
RMSSD	*ms*	Root mean square of successive RR interval differences
SDSD	*ms*	Standard deviation of successive RR interval differences
RR50	-	Number of RR interval differences greater than 50 ms
pRR50	-	Ratio between RR50 and total number of RR intervals
RR20	-	Number of RR interval differences greater than 20 ms
pRR20	-	Ratio between RR20 and total number of RR intervals

Table 3. HRV frequency-domain features

Feature	Unit	Description
Total Power	ms^2	Power over all FC
Absolute Power	ms^2	Power of each FC
Relative Power	*%*	Relative power of each FC
Log (Power)	-	Natural logarithm of FC's absolute power
Normalized Power	-	Normalized powers of each LF and HF component
Peak Frequency	*Hz*	Frequency where maximum power of the FC occurs
LF/HF Ratio	*%*	Ratio between the LF and HF components

The nonlinear features are intended to enhance nonlinear characteristics and unpredictability of RR interval series, which are caused by the different complex physiological dynamics of the human body that lead to HRV (e.g. sympathetic vs. parasympathetic (Billman 2011; Shaffer and Ginsberg 2017). The Poincaré plot is a scatter plot where a given RR interval is plotted against its successor RR interval. It is a graphical tool for HRV analysis of a RR interval

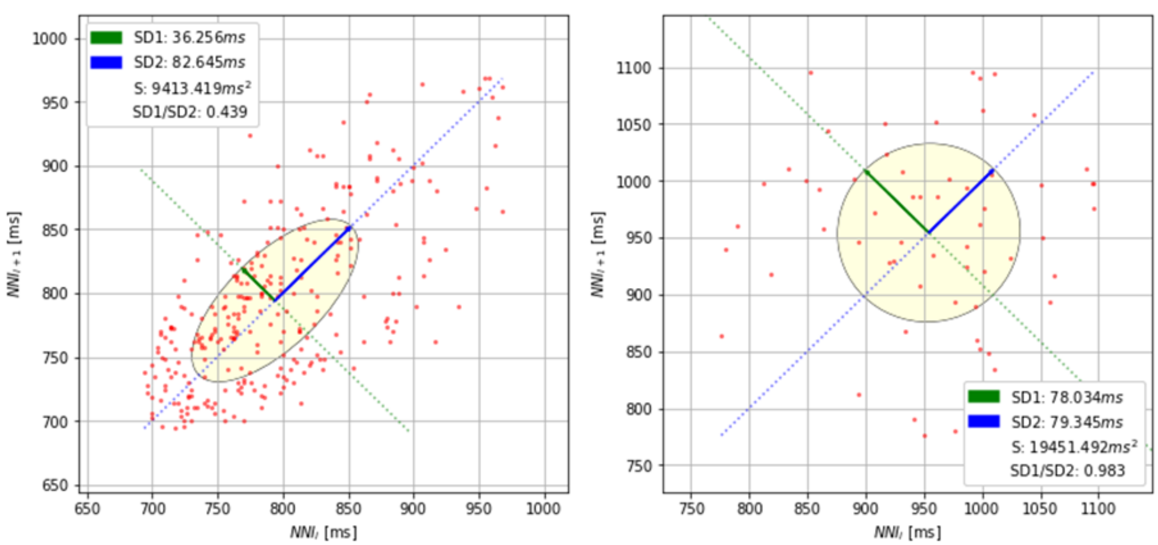

Figure 3. Poincaré plots for a subject at BL (left) and a subject undergoing CPT/ experiencing pain (right). There is a clear linear relationship between the given RR interval and its successor RR interval during the BL whereas there is no clear relationship while the subject is undergoing CPT. This subject halted the CPT early, indicating a high level of pain experienced.

dataset, which allows for a rapid first judgment of a subject's health, as the shape of the scatter plot provides a visual representation of the overall HRV (Tayel and AlSaba 2015). A Poincaré plot depicting a sample participant's BL and CPT is shown in Figure 3 and an overview of the nonlinear Poincaré plot features that were calculated are presented in Table 4. The cluster in Figure 3 when the subject is undergoing CPT appears in the top right corner. Such appearance is an evidence of the parasympathetic nervous system having a high impact on subject's heart rate. Often, the main cluster in the upper right position tends to be rather wide reflecting the fact that HRV becomes higher with lower heart rates/higher RR (CardioPoint 2014). Additionally, the width of this cluster shows how respiratory sinus arrhythmia and parasympathetic nervous system contributes to the total HRV. The wider the cluster, the higher the effect of respiratory arrhythmia to the total HRV (CardioPoint 2014).

Model Selection

Modeling efforts began by generating training, test, and holdout subsets to be used as inputs for the model with a 60%, 20%, and 20% split, respectively. Efforts were made to ensure that each of the training, test, and holdout subsets was representative of the entire dataset. This meant that each of the subsets included either 60% (training set) or 20% (test and holdout sets) of the total participants as well as the total BL and CPT epochs. For example, the representative percentages of subjects ended the CPT early and consequently have shorter pain recordings were considered across each of the subsets. A holdout subset was created to prevent overfitting (Brownlee 2016). The holdout set was not used for model fitting or model selection, otherwise, this would result in an unrealistically optimistic estimate of the performance of the method (Murphy 2012).

To account for the class imbalance between the BL and CPT epochs, the Synthetic Minority Oversampling Technique (SMOTE) was utilized within scikit-learn on the training set. SMOTE is a statistical technique for increasing the number of cases in a dataset in a balanced way (Chawla et al. 2002). The module works by generating new instances from existing minority cases that are supplied as inputs. The new instances are not just copies of existing minority cases; instead, the algorithm takes samples of the feature space for each target class and its nearest neighbors, generating new examples that combine features of the target case with features of its neighbors. This approach increases the features available to each class and makes the samples more general (Chawla et al. 2002).

A preprocessing and classification pipeline to differentiate pain vs. not-pain was developed utilizing Python's scikit-learn library. The pipeline applied standard scaling (Z-normalization) as well as dimensionality reduction via principal component analysis (PCA) before modeling. Five-fold cross-validation was implemented to evaluate the average performance of the algorithm on the train set. Logistic regression classifier was chosen to differentiate CPT (pain) vs. baseline (not-pain).

Pain Classifier Implementation and Validation Study

To validate the performance of the pain classifier within the fighter pilot community, the classifier was integrated with a flight-approved smartwatch (Garmin fēnix 5) and mobile application. The Python developed algorithm was translated to Kotlin and Swift for integration with Android operating system and Apple operating system, respectively. Field data collection is currently underway with the 56th Training Squadron at Luke Air Force Base (AFB) to determine classification accuracy and ruggedness in operational environments. Each participant was given a mobile application for download in addition to their flight-approved smartwatch and were instructed on the use of the application in conjunction with the smartwatch. The application measures pain levels and provides participants with pain alerts once their unique thresholds are crossed. Participants are then prompted to fill out pain surveys which provide additional context of the pain event to assess algorithm accuracy. The end user has the ability to select a specific location on a 3-D model in a specific region. Once the region is selected, the end

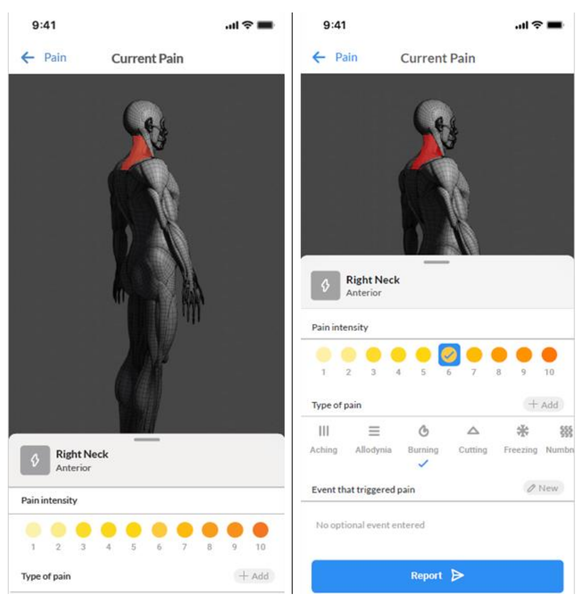

Figure 4. Conceptual mobile application designs for self-reporting pain functions.

user then goes through a series of questions identifying the pain intensity, type of pain, and provide a brief description on the event that caused the pain (Figure 4). Results will allow us to validate the accuracy between cardiovascular metrics derived using our algorithm and the gold standard approach.

RESULTS

ECG RR intervals were used for analysis, feature generation, and initial model selection in an effort to mimic a real-world collection scenario from a smartwatch that can be deployed within the USAF pilot community. This initial modeling approach with the RR intervals was reliable in creating an algorithm capable of classifying pain at an accuracy level of 79.6% with 81% sensitivity and 78% specificity. This initial model was generated utilizing all features (N=46). After analyzing the permutation importance of all the features , it was clear that the HRV time domain features heavily influenced the model. The permutation importance is an intuitive, model-agnostic method to estimate the feature importance for classifier and regression models. Further inspection using boxplots was done to visualize the difference in features for the BL and CPT classes. Boxplots are a standardized way of displaying the distribution of data by showing the five-number summary of a set of data: including the minimum score, first (lower) quartile, median, third (upper) quartile, and maximum score. The boxplots of the top four contributing features showed a clear difference between the features for the BL and CPT classes (Figure 5). These top 4 contributing features were all in the HRV time-domain, and specifically included the maximum and minimum heart rate and RR interval.

Since the HRV time-domain features influenced the model heavily, only the HRV time-domain features (n=18) were utilized with the initial classification pipeline. This optimized algorithm was capable of classifying pain at an accuracy level of 78.7% with 85% sensitivity and 73% specificity. Finding the optimal balance between model complexity and performance ensures that the model can run efficiently when deployed and minimizes the potential for overfitting.

The pain algorithm was then integrated with a flight-approved smartwatch (Garmin fēnix 5) and mobile application for classification in real-world environments. Field data collection is currently underway with the 56th Training Squadron at Luke Air Force Base (AFB) to determine classification accuracy and ruggedness in operational environments. Results will allow us to validate the accuracy between cardiovascular metrics derived using our algorithm and the gold standard approach.

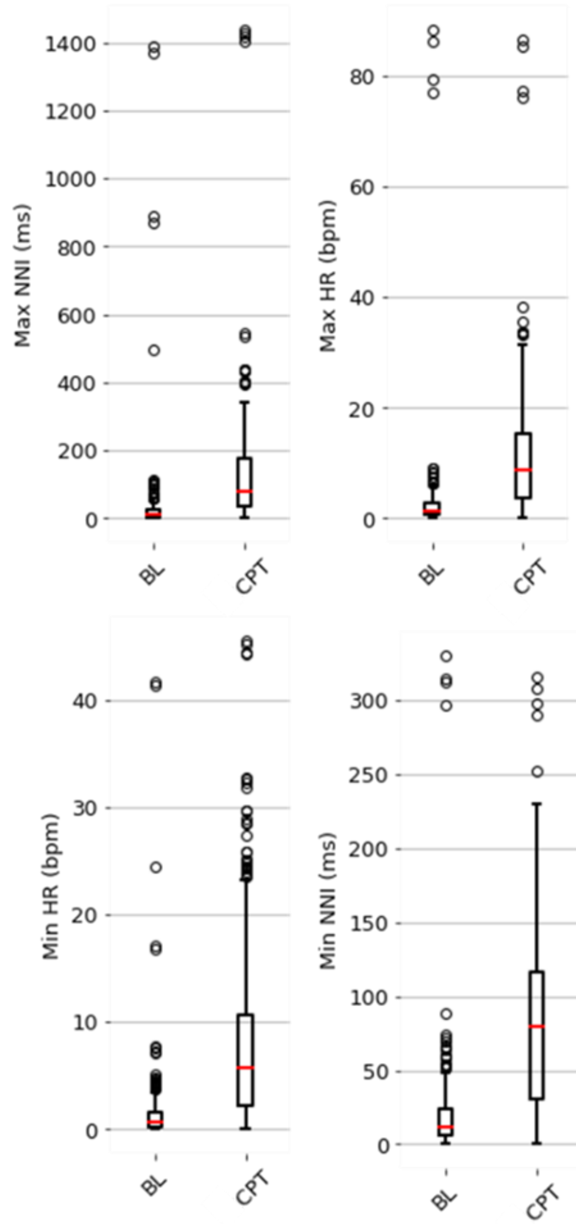

Figure 5. These four plots represent the features contributing most to the model. Each boxplot represents all of the subject's epochs for the feature present.

DISCUSSION

The standard method for prediction of the absence and presence of pain has long been self-report. The current study shows the feasibility of creating an individualized, physiological classifier of pain with a high degree of accuracy and compatible with a flight approved smartwatch. The use of such an algorithm in the fighter pilot community may allow for early pain detection to help augment rehabilitation and enhance pilots' treatment decisions. Given the large number of individuals in the fighter pilot community that experience musculoskeletal pain and the unmet need for detecting pain, such an algorithm can have the potential to augment treatment. Current approaches to provide preventative care for pilots for early management of neck and back pain has been centered around increasing physical therapy efforts and wellness personnel. There is a unique opportunity to implement an algorithm that can collect, analyze, store and present objective pain data.

A number of previous groups have reported physiology-based approaches to pain classification, but such approaches generally require the use of laboratory equipment or controlled settings. For instance, one group leveraged fMRI data and support vector machines (SVMs) to develop an algorithm of pain with 81% accuracy (Brown et al. 2011). Another group reported an electroencephalography (EEG)-based algorithm of pain by analyzing 64 channels of EEG data with an accuracy of 80% (Huang et al. 2013). Another group leveraged blood volume pulse (BVP), ECG, and skin conductance level (SCL) to create a pain classifier with a 75% accuracy (Chu et al. 2017). The current effort was able to achieve comparable accuracy by leveraging cardiovascular sensors from emerging wearable fitness devices. This also represents the first effort to provide a fully fieldable solution which does not require controlled laboratory conditions.

Classification of pain was based on features gathered from a user group that underwent a pain induction task (CPT), which is considered a reliable experimental method for controlled pain induction (Mitchell, MacDonald, and Brodie 2004) and is currently being validated with USAF pilots to determine classification accuracy and ruggedness in operational environments. Pain classification was based on cardiovascular inputs and variance due to individual differences was addressed by individual baseline normalization. The current classifier is a binary output signaling if an individual is in pain or not in pain. Future work will focus on reporting pain on scale to show the level of pain that an individual is experiencing.

Future research efforts will be to integrate the pain classification algorithm with a broader class of algorithms such as stress and fatigue that can be utilized within the fighter pilot community. Raw physiological data, including blood volume pulse, electrodermal activity, temperature, and movement data, can be processed in real time to present objective stress of a fighter pilot (Chadderdon et al. 2014; Winslow et al. 2016). This stress algorithm can be used to keep tabs on the mental strain faced by fighter pilots during training, operations, and outside of work, thereby contributing to a holistic wellness approach. Additionally, by utilizing a combination of sleep (actigraphy) data from wearable sensors, self-reported information such as caffeine intake, and a game-based cognitive assessment, there is the capability to predict an individual's current and future mental and physical fatigue over the next 72 hours (Winslow, Nguyen, and Venta 2017). A valid assessment of current and future mental/physical fatigue could be used to adjust fighter pilot training and mission schedules, inform optimal times for fatigue interventions, and support a more personalized rehabilitation regimen (Strahler et al. 2016). Combining pain, stress, fatigue algorithms allows for a more holistic, individualized approach to pilot wellness.

Pilot health and readiness are critical to success of the squadron and to meet the overall mission of the unit. Within the aviation community, pilots are either in an "UP" status and able to fly or in a "Medically Down" status and unable to fly due to injuries or illness. There is a concern of pilot acceptance to wearing a physiological sensor that provides an output on the level of pain they are experiencing that is accessible by their medical staff. To combat this, the data must be utilized as an identifier to enhance the individual pilot's physical training and physical therapy program and not utilized as a tool to downgrade or bypass the pilot's own assessment of their physical state. Utilizing physiological sensors to monitor a pilot's level of pain can positively impact the fighter pilot community by providing valuable data to the medical staff and leadership. The capability to identify trends associated with pilots showing variable signs of pain throughout the training curriculum, down to specific maneuvers or duration in flight, can provide highly valuable data to the medical staff and human performance teams. This data can allow the training staff to modify their physical therapy routines and adjust their strength and conditioning programs for the individual pilot.

ACKNOWLEDGEMENTS

This work was supported by grants from the Office of Naval Research (N00014-12-G-0427) and the US Air Force (FA864920P0199).

REFERENCES

Bent, Brinnae, Benjamin A. Goldstein, Warren A. Kibbe, and Jessilyn P. Dunn. 2020. “Investigating Sources of Inaccuracy in Wearable Optical Heart Rate Sensors.” *npj Digital Medicine* 3(1): 1–9. https://doi.org/10.1038/s41746-020-0226-6 (June 29, 2021).

Billman, George E. 2011. “Heart Rate Variability - A Historical Perspective.” *Frontiers in Physiology* 2 NOV: 86. www.frontiersin.org (June 29, 2021).

Brown, Justin E., Neil Chatterjee, Jarred Younger, and Sean Mackey. 2011. “Towards a Physiology-Based Measure of Pain: Patterns of Human Brain Activity Distinguish Painful from Non-Painful Thermal Stimulation.” *PLoS ONE* 6(9). https://pubmed.ncbi.nlm.nih.gov/21931652/ (June 29, 2021).

Brownlee, Jason. 2016. “Overfitting and Underfitting With Machine Learning Algorithms.” https://machinelearningmastery.com/overfitting-and-underfitting-with-machine-learning-algorithms/ (June 29, 2021).

CardioPoint, BTL. 2014. *BTL CardioPoint-Poincaré Graph Poincaré Graph Complete ECG Record in One Sight*.

Chadderdon, George L., Brent D. Winslow, Sara J. Dechmerowski, and David Jones. 2014. “System, Method, and Computer Program Product for the Real-Time Mobile Evaluation of Physiological Stress.”

Charlton, Peter H. et al. 2018. “Breathing Rate Estimation from the Electrocardiogram and Photoplethysmogram: A Review.” *IEEE Reviews in Biomedical Engineering* 11: 2–20. https://pubmed.ncbi.nlm.nih.gov/29990026/ (June 29, 2021).

Chawla, Nitesh V., Kevin W. Bowyer, Lawrence O. Hall, and W. Philip Kegelmeyer. 2002. “SMOTE: Synthetic Minority over-Sampling Technique.” *Journal of Artificial Intelligence Research* 16: 321–57. https://www.jair.org/index.php/jair/article/view/10302 (June 29, 2021).

Chu, Yaqi, Xingang Zhao, Jianda Han, and Yang Su. 2017. “Physiological Signal-Based Method for Measurement of Pain Intensity.” *Frontiers in Neuroscience* 11(MAY): 279. www.frontiersin.org (June 29, 2021).

Cowen, R., M. K. Stasiowska, H. Laycock, and C. Bantel. 2015. “Assessing Pain Objectively: The Use of Physiological Markers.” *Anaesthesia* 70(7): 828–47. https://pubmed.ncbi.nlm.nih.gov/25772783/ (June 29, 2021).

Gomes, Pedro. 2021. “PyHRV - OpenSource Python Toolbox for Heart Rate Variability Documentation.” https://pyhrv.readthedocs.io/en/latest/ (June 29, 2021).

Honkanen, Tuomas et al. 2019. “Cross-Sectional Area of the Paraspinal Muscles and Its Association with Muscle Strength among Fighter Pilots: A 5-Year Follow-Up.” *BMC Musculoskeletal Disorders* 20(1): 1–8. https://doi.org/10.1186/s12891-019-2551-y (June 29, 2021).

Huang, G. et al. 2013. “A Novel Approach to Predict Subjective Pain Perception from Single-Trial Laser-Evoked Potentials.” *NeuroImage* 81: 283–93. https://pubmed.ncbi.nlm.nih.gov/23684861/ (June 29, 2021).

Jafari, Hassan et al. 2017. “Pain and Respiration: A Systematic Review.” *Pain* 158(6): 995–1006. https://pubmed.ncbi.nlm.nih.gov/28240995/ (June 29, 2021).

Kikukawa, A, S Tachibana, and S Yagura. 1995. “G-Related Musculoskeletal Spine Symptoms in Japan Air Self Defense Force F-15 Pilots.” *Aviat Space Environ Med* 66(3): 269–72.

Knudson, R., D. McMillan, D. Doucette, and M. Seidel. 1988. “A Comparative Study of G-Induced Neck Injury in Pilots of the F/A-18, A-7, and A-4.” *Aviation Space and Environmental Medicine* 59(8): 758–60. https://europepmc.org/article/med/3178626 (June 29, 2021).

Koenig, J. et al. 2014. “Heart Rate Variability and Experimentally Induced Pain in Healthy Adults: A Systematic Review.” *European Journal of Pain (United Kingdom)* 18(3): 301–14. https://pubmed.ncbi.nlm.nih.gov/23922336/ (June 29, 2021).

Ledowski, Thomas et al. 2007. “The Assessment of Postoperative Pain by Monitoring Skin Conductance: Results of a Prospective Study.” *Anaesthesia* 62(10): 989–93. https://pubmed.ncbi.nlm.nih.gov/17845649/ (June 29, 2021).

Loggia, Marco L., Mylne Juneau, and M. Catherine Bushnell. 2011. “Autonomic Responses to Heat Pain: Heart Rate, Skin Conductance, and Their Relation to Verbal Ratings and Stimulus Intensity.” *Pain* 152(3): 592–98. https://pubmed.ncbi.nlm.nih.gov/21215519/ (June 29, 2021).

Lowery, Daniel, Roger B. Fillingim, and Rex A. Wright. 2003. “Sex Differences and Incentive Effects on Perceptual and Cardiovascular Responses to Cold Pressor Pain.” *Psychosomatic Medicine* 65(2): 284–91. https://pubmed.ncbi.nlm.nih.gov/12651996/ (June 29, 2021).

Mccrary, Brian F, Daniel L. V A N Syoc Mccrary Bf, and D L Van Svoc. 1995. 73 Aviat Space Environ Med *Permanent Flying Disqualifications of USAF Pilots and Navigators (1995-1999).* https://www.researchgate.net/publication/11034708 (June 29, 2021).

Mitchell, Laura A., Raymond A.R. MacDonald, and Eric E. Brodie. 2004. "Temperature and the Cold Pressor Test." *Journal of Pain* 5(4): 233–37. https://pubmed.ncbi.nlm.nih.gov/15162346/ (June 29, 2021).

Murphy, Kevin P. 2012. *Machine Learning A Probabilistic Perspective*.

Peltola, Mirja A. 2012. "Role of Editing of R-R Intervals in the Analysis of Heart Rate Variability." *Frontiers in Physiology* 3 MAY. /pmc/articles/PMC3358711/ (June 29, 2021).

Raja, Srinivasa N. et al. 2020. "The Revised International Association for the Study of Pain Definition of Pain: Concepts, Challenges, and Compromises." *Pain* 161(9): 1976–82. https://pubmed.ncbi.nlm.nih.gov/32694387/ (June 29, 2021).

Rintala, Harri, Arja Häkkinen, Simo Siitonen, and Heikki Kyröläinen. 2015. "Relationships between Physical Fitness, Demands of Flight Duty, and Musculoskeletal Symptoms among Military Pilots." *Military Medicine* 180(12): 1233–38. https://academic.oup.com/milmed/article/180/12/1233/4160603 (June 29, 2021).

Schneider, Stefan, Doerte U. Junghaenel, Masakatsu Ono, and Arthur A. Stone. 2018. "Temporal Dynamics of Pain: An Application of Regime-Switching Models to Ecological Momentary Assessments in Patients with Rheumatic Diseases." *Pain* 159(7): 1346–58. https://pubmed.ncbi.nlm.nih.gov/29557930/ (June 29, 2021).

Shaffer, Fred, and J. P. Ginsberg. 2017. "An Overview of Heart Rate Variability Metrics and Norms." *Frontiers in Public Health* 5: 258. /pmc/articles/PMC5624990/ (June 29, 2021).

Strahler, Jana, Nadine Skoluda, Nicolas Rohleder, and Urs M. Nater. 2016. "Dysregulated Stress Signal Sensitivity and Inflammatory Disinhibition as a Pathophysiological Mechanism of Stress-Related Chronic Fatigue." *Neuroscience and Biobehavioral Reviews* 68: 298–318. https://pubmed.ncbi.nlm.nih.gov/27208412/ (June 30, 2021).

Streff, Anouk, Linn K. Kuehl, Gilles Michaux, and Fernand Anton. 2010. "Differential Physiological Effects during Tonic Painful Hand Immersion Tests Using Hot and Ice Water." *European Journal of Pain* 14(3): 266–72. https://pubmed.ncbi.nlm.nih.gov/19540783/ (June 29, 2021).

Tayel, Mazhar B., and Eslam I. AlSaba. 2015. "Poincaré Plot for Heart Rate Variability." https://zenodo.org/record/1109321 (June 29, 2021).

Winslow, Brent D. et al. 2016. "Development and Clinical Evaluation of an MHealth Application for Stress Management." *Frontiers in Psychiatry* 7(JUL): 130. /pmc/articles/PMC4960497/ (June 30, 2021).

Winslow, Brent D., Nam Nguyen, and Kimberly E. Venta. 2017. "Improved Mental Acuity Forecasting with an Individualized Quantitative Sleep Model." *Frontiers in Neurology* 8(APR): 160. www.frontiersin.org (June 30, 2021).

Yasuma, Fumihiko, and Jun Ichiro Hayano. 2004. "Respiratory Sinus Arrhythmia: Why Does the Heartbeat Synchronize with Respiratory Rhythm?" *Chest* 125(2): 683–90.