Development and Evaluation of Biosensing Apparel for Monitoring Fighter Pilot Physiological Episodes

Nichola Lubold, Tor Finseth
Honeywell International
Phoenix, AZ
nichola.lubold@honeywell.com,
tor.finseth@honeywell.com

Bridget Rinkel, Raisa Marshall
Naval Air Warfare Center Aircraft Division
Patuxent River, MD
bridget.k.rinkel.civ@us.navy.mil,
raisa.e.marshall.civ@us.navy.mil

ABSTRACT

As aircraft become more sophisticated, there are increasing demands on pilots to monitor and assess aircraft operations. Simultaneously, a pilot's ability to monitor and assess operations may be hindered by physiological episodes or a loss in performance due to fatigue, stress, hypoxia, hypocapnia, atelectasis, hypothermia, and G-force. This is particularly relevant for fast jet aircraft where a physiological episode may spell disaster. Measuring and recording physiological parameters that directly assess human health and performance in the flight environment is critical to understanding, supporting, and training pilots. However, gathering high quality data is challenging. Sensor accuracy can be impacted by sensor location, sampling rates, movement artifacts, and more. Layers of gear can limit viability of sensors and impose constraints on practicality, signal quality, and user acceptability. This work explores whether biosensing garments can address these challenges, among others, through three research questions: (1) how effective are biosensing garments in capturing physiological signals for both males and females in the environments that fighter pilots operate, (2) how acceptable are biosensing garments to fighter pilots for capturing physiological signals, and (3) how can biosensing garments detect states that may indicate fighter pilot physiological episodes. To answer these questions, seven within-subject evaluations were conducted where biosensing garments suited to both male and female physiology were assessed using a reduced oxygen breathing device (ROBD), altitude chamber, thermal chamber, cycle ergometer, centrifuge, and aircraft. The apparel was compared to well-established, clinical grade sensors, and subjective feedback was collected. Results suggest the apparel has potential for effectively capturing heart rate, breathing rate, and breathing depth in the extreme environments fighter pilots operate and highlight the challenge of assessing physiological monitoring in flight. Assessment of pilot acceptance criteria provides insight into the best applications, potential limitations, and optimal areas of use for biosensing apparel.

ABOUT THE AUTHORS

Dr. Nichola Lubold is a lead research scientist at Honeywell Aerospace in human-centered systems. She is a principle investigator on a broad-range of multi-disciplinary research efforts focusing on modeling physiological and cognitive states and identifying the contributions and complications that result from individual differences through application of unique sensors, natural language processing, signal processing, and machine learning.

Dr. Tor Finseth is an R&D human factors scientist at Honeywell Aerospace. He earned his PhD at Iowa State University in Aerospace Engineering and Human-Computer Interaction. His research interests include the study of human stress response in context of occupational training and advanced technologies such as virtual reality, time-series machine learning, stress detection, adaptive systems.

Raisa E. Marshall is a Human Systems Engineer at the Naval Air Warfare Center Aircraft Division (NAWCAD). She earned her B.S in Biology from Morgan State University. She currently completing her Master's in Human Systems Integration. She is part of the Physiological Monitoring Team that is focused on developing devices to monitor the warfighter in-flight to prevent physiological episodes and improve performance.

Bridget K. Rinkel is a Biomedical Engineer at the Naval Air Warfare Center Aircraft Division (NAWCAD). She earned her B.S in Biomedical Engineering and M.S. in Neuroscience from Worcester Polytechnic Institute. She is part of the Physiological Monitoring Team that is focused on developing devices to monitor the warfighter in-flight to prevent physiological episodes and improve performance.

Development and Evaluation of Biosensing Apparel for Monitoring Fighter Pilot Physiological Episodes

Nichola Lubold, Tor Finseth
Honeywell International
Phoenix, AZ
nichola.lubold@honeywell.com, tor.finseth@honeywell.com

Bridget Rinkel, Raisa Marshall
Naval Air Warfare Center Aircraft Division
Patuxent River, MD (NAWCAD)
bridget.k.rinkel.civ@us.navy.mil, raisa.e.marshall.civ@us.navy.mil

INTRODUCTION

Fighter pilots experience extreme physical and mental stressors ranging from high G-forces and rapid altitude changes to prolonged periods of intense concentration. Ensuring the health and well-being of fighter pilots in this demanding and high-stakes environment is paramount to mission success and overall operational readiness. Historically, health monitoring of fighter pilots has relied heavily on pre- and post-mission evaluations, which, while valuable, do not provide real-time data on the pilot's physiological state during flight. Pilots can experience a range of physiological issues during flight caused by respiratory, cardiovascular, neurological, and environmental factors (Shaw & Harrell, 2023). As the complexity of aerial missions continues to grow, so does the necessity for real-time, advanced pilot health monitoring systems to ensure peak performance and safety.

Presented in the form of apparel that an individual can wear on an everyday basis, biosensing garments have the potential to revolutionize the way we monitor and support pilots' physiological state during flight operations. Biosensing apparel are garments that have physiological sensors integrated into the fabric and by an individual wearing the garment, the sensors can make direct contact with the skin and record biosignals directly at their source, moving with the body and maintaining contact through natural movements. Biosensing apparel has seen significant advancements over the past few years (Li et al., 2023; Schauss, Arquilla, & Anderson, 2022; Azeem et al, 2024). The majority of applications for biosensing apparel have been in the medical community or in the sports health community; however, the integration of sensors and smart textiles could support the detection and management of harmful physiological states for fighter pilots, such as G-induced loss of consciousness (G-LOC), hypoxia, hypercapnia, hypocapnia, fatigue, and stress. In addition, biosensing apparel has the potential to support and accelerate pilot training, enhance pilot situation awareness, and facilitate decision-making through real-time insights, as seen in other domains (Ghahari, et al., 2018; Bourdon, et al., 2017).

Despite the potential of biosensing apparel, there are several challenges when considering the use of this technology for fighter pilots. First, biosensing garments have to strike a delicate balance between compression and comfort. Common strategies such as the use of synthetic fabrics that can offer high degrees of compression to keep sensors in contact with the skin can contribute to uncomfortable garments that restrict movement, offer poor regulation of body heat and moisture, and may not be an ideal material for the environment of fast-jet aircraft (Di Domenico, Hoffman, & Collins, 2022). Additionally, the interface between electronics and textiles can require physical hardware that can be uncomfortable and may interfere with other flight apparel a pilot needs to wear. Second, reliability and consistency of measurement over time can present a challenge. Biosensing garments require the integration of rigid electronics with garments. The integration has to survive repeated use and remain tolerant to water, detergents, and high levels of mechanical agitation. Third, the environment under which fighter pilots operate is unique, as studies have found that psychophysiological responses differ between real and simulated flights (Fuentes-García et al., 2021). Lastly, the design of flight suits and protective gear has in the past been predominantly based on male anatomical and anthropometric data, which can result in suboptimal fit and performance for female pilots (Engel, 2021). This disparity not only affects pilot comfort and mobility but can also compromise the effectiveness of the gear. The consequences of designs based only on male physiology has led to growing recognition of the need for designs that consider both male and female anatomy and physiology.

Several studies have attempted to develop biosensing technology for fighter pilots. Kim et al. (2017) developed a system that used photoplethysmogram (PPG) from a finger-based sensor and electromyogram (EMG) sensors located on pilots' calves to detect G-LOC. They validated the system using a centrifugal simulator. This work demonstrated the potential of wearable systems to detect and provide context-awareness to pilots, but the sensor system is not easy to put on and results in additional required equipment for the pilot. Rice et al. (2016) used a portable triaxial accelerometer and Zephyr BioPatch sensor (Zephyr Technology, Annapolis, MD) to measure heart rate, respiratory rate, and temperature while F-18 pilots flew low and high G maneuvers. While the Zephyr was beneficial in the collection of data, the author notes that only a fraction of physiological episodes would have been detected and more reliable biosensors are needed.

This paper describes the extent to which a novel biosensing apparel design can monitor physiological responses of fighter pilots. We conducted a series of human-centered evaluations to answer three research questions:

1. Given the challenge of balancing comfort with compression, how effective are biosensing garments in capturing physiological signals for both males and females in the environments that fighter pilots operate?
2. How acceptable are biosensing garments to fighter pilots for capturing physiological signals?
3. Can biosensing garments be used to detect states that may indicate fighter pilot physiological episodes?

To answer these questions, seven within-subjects evaluations were performed in varying environments using a reduced oxygen breathing device (ROBD), altitude chamber, thermal chamber, cycle ergometer, centrifuge, and aircraft. Sensor data from the apparel were compared to well-established, clinical grade sensors, and subjective feedback was collected to evaluate the apparel. For the aircraft environment, a flight squadron was asked to wear the apparel over several months. Subjective assessments from the pilots, evaluations of the collected data, and measurements of the change in the apparel over time were used to assess how reliably and consistently the apparel performed while the squadron was using it.

BIOSENSING APPAREL DESIGN

To be used in everyday military applications, the apparel was required to be commensurate to regular clothing. Therefore, we focused on designs of apparel in the form of t-shirts for men and sports bras for women, that could be used in conjunction with a flight suit and accommodate male and female pilots (see Figure 1).

Military standards required the biosensing apparel to be robust, with fabric certified as no-melt, no-drip and Berry compliant (meaning the product is made and sourced in the United States). Comparisons of fabric comfort, burst strength, weight, liquid evaporation rate, and thermal insulation resulted in the selection of a fabric composition consisting of 55% cotton, 37% nylon, and 8% spandex.

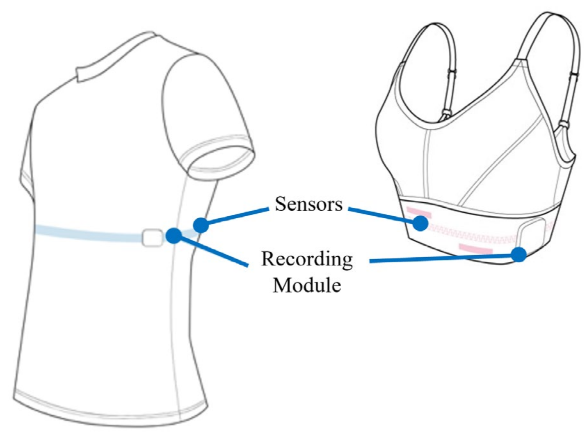

Figure 1. Biosensing apparel is in the form of a t-shirt for men and a sports bra for women with sensors embedded in the garment.

For this work, three kinds of biosensors were integrated into the apparel to measure cardiovascular effects during high G-force maneuvers: electrocardiogram (ECG) at 250 Hz, respiration at 25 Hz, and acceleration motion at 50 Hz. These biosignals were the focus as they hold promise for early detection of physiological episodes (Kumagai, et al, 2023). Each garment integrates electronics featuring the biosensors, conductive traces to transmit sensed signals to a hardware processing module, and a connector where the soft materials of the garment meet hard materials of the hardware module. The hardware module, also known as a “puck”, converts analog signals from ECG and breathing sensors embedded in the garment to digital data. Up to 24 hours of data can be stored on the recording module or the data can be transmitted real-time via Bluetooth Low Energy. The removable hardware module connects to the garment via mechanical conductive fasteners. Raw ECG, respiration, and movement signals are converted into heart rate, breathing rate, breathing depth, and G-force by a series of algorithms that filter and process the raw data.

METHOD

Seven within-subject evaluations were conducted to explore the performance of the biosensing apparel in different task environments, presented in Table 1. For each participant, the physiological signals from the biosensing apparel were compared to reference sensors. Reference sensors were placed on the participants in all evaluations except for the "in-the-wild" flight squadron evaluation. The reference sensors differed depending on the evaluation due to availability of sensor systems and the ability to place reference sensors on participants. High-level descriptions of the evaluations, participant sample sizes, reference sensors, and other gear participants wore are provided in Table 1. For all evaluations, research protocols specified applicable termination criteria, and Institutional Review Board approval was obtained by the Environmental Physiology Human Performance Laboratory through the U.S. Navy.

Table 1. Descriptions of the different evaluations used to assess the biosensing apparel.

Evaluation	Description	Participants	Reference Sensors	Other Gear
Reduced Oxygen Breathing Device	Mixed oxygen & nitrogen simulating three different altitude profiles ranging from 0 ft to 18,000 ft	3 male 1 female	Mindray ECG, Masimo Forehead Sensor, pulse oximetry pod	Flight suit and helmet
Altitude Chamber	Simulated air pressure conditions replicating changes in air pressure at altitudes from 10,000 ft to 17,500 ft in two different conditions	3 male	Nonin WristOx	Flight suit and helmet, life support equipment
Thermal Chamber	Thermal exercise heat stress test (work/rest exercise routine) in 95° Fahrenheit and 60% humidity	5 male 4 female	Mindray ECG	Flight suit and helmet
Cycle Ergometer	Low and high metabolic demand in a work/rest exercise routine, interspaced with performing specific breathing maneuvers	25 male 11 female	Pulse oximetry pod	No other gear
Centrifuge	Simulated gravitational forces up to 9Gz. Six acceleration profiles ranged from 0.1G/s up to 6 G/s	6 male 2 female	ECG (unspecified) Nonin WristOx	Flight suit and helmet, anti-G suit, life support equipment
Flight Testing	Dedicated in-flight evaluation with specific high-G aircraft maneuvers	9 male 1 female	Nonin WristOx	Flight suit and helmet, anti-G suit, life support equipment, other flight gear.
Flight Squadron	Tag-along flight (in-the-wild use context); apparel worn as a part of regular gear.	15 male 2 female	None	Flight suit and helmet, anti-G suit, life support equipment, other flight gear.

Task Environments

Reduced Oxygen Breathing Device (ROBD)

ROBDs can mix oxygen and nitrogen to generate breathing environments that pilots may be exposed to at varying altitudes. In this evaluation, an ROBD was used to mix oxygen and nitrogen reflective of three different altitude profiles ranging from 0 ft. to 18,000 ft. In the first profile, which was a gradual onset rate (GOR) profile, participants experienced 3 minute stepped plateaus simulating 8,000, 10,000, 14,000, and 17,500 ft. altitudes at an ascent rate of 5,000 feet per minute. They then descended at 5,000 feet per minute to ground level and continued to breathe attached to the ROBD at 21% oxygen for a 15 minute recovery period. After a 30 minute rest, participants then experienced a second GOR profile consisting once again of 3 minute stepped plateaus simulating 8,000, 10,000, 14,000, and 17,500 ft. altitudes at an ascent rate of 5,000 feet per minute. This was followed by 3 minute step-down plateaus simulating 14,000, 10,000, and 8,000 ft. at a descent rate of 5,000 feet per minute to ground level. Participants then continued to breathe attached to the ROBD at 21% oxygen for a 15 minute recovery period. Finally, after another 30 minute rest, participants experienced the last profile, a rapid onset rate (ROR), consisting of 3 minute stepped plateaus simulating 8,000, 10,000, 14,000, and 17,500 ft. at an ascent rate of 1,000 feet per second. This was followed by a descent to ground level at 1,000 feet per second and a 15 minute recovery period at 21% oxygen. Prior to the three profiles, each participant completed a baseline test in which oxygen levels represented normal ratios. Three out of four participants completed all three scenarios successfully. Due to time limitations, one participant was unable to complete the rapid onset rate profile.

Altitude Chamber

Altitude chambers can produce pressure conditions that emulate ascending and descending through various altitudes. For this evaluation, a human-rated hypobaric chamber on the Brooks City-Base complex in San Antonio, Texas, was used. Three male participants were exposed to two conditions: one with a total simulated altitude of 17,500 ft. with 21% O_2 and one with a total simulated altitude of 25,000 ft. with 21% O_2. For the first condition, participants were exposed to three altitude levels: 10,000 ft. for 10 minutes, 14,000 ft. for 10 minutes, and 17,500 ft. for 20 minutes. In the second condition, participants were exposed to two altitude levels: 10,000 ft. for 10 minutes and 25,000 ft. for 20 minutes. Prior to the 25,000 ft. exposure, participants pre-breathed 100% oxygen for 30 minutes to “wash out” tissue nitrogen. Participants were exposed to the two conditions on separate days.

Thermal Chamber

Exercise plus heat stress can push physiological systems to the limit by forcing the body to simultaneously support competing metabolic and thermoregulatory demands. Five males and four females participated in a work/rest exercise routine in a thermal chamber set to 95° Fahrenheit and 60% humidity. Participants used a cycle ergometer in the chamber, cycling 1.5 minutes and resting for 1 minute at 25% VO_2 max for 60 minutes.

Cycle Ergometer

Metabolic state was manipulated through a similar work/rest exercise routine as to the thermal chamber. However, in this evaluation, participants were not exposed to heat stress in the thermal chamber but were instead directed to perform specific breathing maneuvers on the cycle ergometer to systematically collect breathing responses and patterns when at rest and under high metabolic demand. Breathing maneuvers included normal breathing patterns, coughing, hyperventilation, holding one’s breath, shallow breathing, and speaking.

Participants first received a familiarization trial on the different breathing behaviors, followed by a baseline trial where normal breathing patterns were elicited, and participants could practice the different breathing behaviors. Each breath type with the exception of normal breathing patterns was elicited at three different intensities, where participants performed the breathing maneuver three times from a highest to lowest subjective intensity. To facilitate performance of the breath types, participants were provided with audible and visual cues.

After the baseline trial, participants performed each breath type separated by 2 minutes of normal breathing while at-rest. Then participants performed each breath type in between bouts of heavy exercise. The exercise protocol was designed to increase overall metabolic demand, thus eliciting high heart rate and ventilation values (Whipp & Rossiter, 2013). Participants exercised for two minutes, performed a breathing maneuver, and then rested for 1 minute. The exercise intensity was set to approximately 75% of the participants predicted maximal heart rate (Tanaka et al., 2001), which was determined by step-incremental testing (e.g., 30 second stages of constant-load exercise) until the target heart rate value is achieved. The exercise intensity was adjusted throughout to always elicit approximately 75% of the individuals maximal heart rate. Seventy-five percent was chosen as this usually corresponds to physiological responses like ventilation and metabolic disturbance associated with the heavy exercise intensity domain (Iannetta et al., 2020) but can be maintained for an extended duration, especially when performed in an intermittent format (Davies et al., 2017; Skiba et al., 2014).

Centrifuge

In the centrifuge evaluations, participants were exposed to physical stress by simulating accelerations and the resulting gravitational forces. The centrifuge on the Brooks City-Base complex in San Antonio, Texas, was used for this evaluation. Participants experienced six acceleration profiles that included a GOR, five ROR, and a profile that mimics what an individual might experience during multiple maneuvers in a fast-jet aircraft called the “Woody” profile. The GOR consisted of accelerating up to a force of 9G at 0.1 G/s. The five ROR consisted of accelerating at 6 G/s up to: 5G two times, 7G two times, and up to 9G once. The “Woody” profile consisted of accelerating at 6 G/s to four plateaus maintained for 10 seconds each. The plateaus consisted of 7.5G, 7.0G, 6.0G, and 6.5G respectively. Participants had 140 seconds of rest at 0.2 G/s in between the 10 second plateaus.

Flight Testing

Pilots wore the biosensing apparel for ten dedicated flight tests, where they performed a variety of planned maneuvers (e.g., set of high-G turns, rolls, and high-angle of attack maneuvers) while wearing the apparel. Maneuvers involved a series of high-G twists, turns, and rolls. Average time to perform the maneuvers was 20.6 minutes (SD = ±1.7).

Flight Squadron

Seventeen pilots from a flight squadron were fitted with two biosensing garments each and asked to wear one garment at least once when they were assigned to fly over a period of 12 weeks. In total, 60 flight events (95.3 hours) were conducted with pilots wearing the apparel. Of these 60 events, three events were executed by females, and six events failed to capture or record data, potentially due to lack of full connectivity between the puck and garment.

Procedure

For all evaluations, participants must have completed the appropriate training to qualify as a participant, or they were screened prior to participating to ensure they had the physical capability to complete an activity using the "Physical Activity Readiness Questionnaire for Everyone" (Warburton et al., 2011). Individuals did not participate in multiple evaluations. Prior to each evaluation, participants were briefed on the task environment and signed a consent form. They were then given a familiarization session specific to the type of evaluation. The familiarization sessions gave participants an opportunity to become familiar with the environment, the setup, and what they would experience. After familiarization, participants donned the biosensing apparel, reference sensor system, and additional gear for that evaluation (see Table 1).

Dependent Variables

The raw ECG and respiration data from the biosensing apparel were aggregated into three biometrics: heart rate, breathing rate, and breathing depth. These three biometrics were compared to third-party reference sensor data to evaluate the apparel's effectiveness in different environments. These sensors had been previously established in the target environments and included the Mindray-BeneVision ECG Telemetry Monitoring System, Nonin WristOx, the Masimo Forehead Sensor Headband, and a fingertip pulse oximeter (e.g., iHealth PO3). For the centrifuge, a generic, medical grade ECG that is a part of the centrifuge training system was used to collect electrocardiogram data. Not all sensors were used in all evaluations (see Table 1). In order to have a common denominator, we report correlations to two biometrics, peripheral oxygen saturation (SpO_2) and heart rate as collected from the sensor systems available in each evaluation. The reference sensors for SpO_2 and heart rate are specified in the results.

Subjective metrics on comfort and acceptability were also collected in all evaluations, and semi-structured interviews were conducted with participants. Participants were asked to provide general thoughts and comments. The following six questions were posed to participants across all evaluations:

1. Describe the general fit and length of the garment.
2. Describe the feel of the fabric.
3. Describe the ease of movement and ability to put on/take off the garment.
4. Can you feel the sensors or connectors?
5. Are there any uncomfortable zones?

For flight squadron testing, a series of additional questions were posed to the participants. These questions included:

6. Are you willing to use biosensing garments in test, training, and/or fleet environments?
7. Did you notice the indicator (a double vibration) that the puck was recording data?
8. Was there any impact to pre/post flight timeline?
9. Was there any impact during the flight?
10. How satisfied were you with the overall fit of the biosensing garment?
11. What do you like about the biosensing garment? What do you not like about the biosensing garment?

State Detection Algorithms

In addition to the dependent variables, a convolution neural network was used to assess whether the apparel can be used to detect states indicative of physiological episodes. States of interest included normal breathing patterns, coughing, hyperventilation, holding one's breath (apnea), shallow breathing, big breaths, and speaking. Not all of these states are indicative of a potential physiological episode, but modeling changes in breathing rate and breathing depth across different contexts is critical to a holistic image of pilot lung functioning. The data collected during the cycle ergometer evaluation was the main dataset used for this assessment. Individual breaths from the participants in that evaluation were manually reviewed and labeled by an expert in electrophysiology. A standard breath detection algorithm was used to "cut" the recording session into individually detected breaths and labels were manually applied, reviewed, and adjusted as needed.

The labeled data were separated into training and validation (participants = 28, number breaths = 61,323) and test (participants = 8, number breaths = 10,272). The training and validation data were used to train the neural network using cross-validation; the test set was withheld to evaluate the model's performance on unseen data. For each breath, the neural net fed an input image that was a vector with 784 pixel values representing the normalized "image" of the individual breath (amplitude vs. time), the average breathing rates of the previous breath, last five breaths, and last ten breaths prior to the current breath, the average depth of the last three breaths prior to the current breath, and whether the participant was male or female. The breathing depth was normalized based on the average depth of the whole session. The model was a convolutional neural network, with two 3x3 convolutional layers and one 2x2 max pooling layer. The model was trained for up to 100 epochs with early stopping criteria. Dropout was introduced after each convolutional layer. Other hyperparameters were fixed: learning rate = .001 and $\beta = .9$.

Data Analysis

Inconsistent signal coverage is a major concern of any biosensor, but with the biosensing apparel, the balance between comfort and compression can potentially lead to less coverage in favor of comfort. ECG quality can be assessed using interbeat interval (RR) coverage, where RR is the time between two consecutive R waves. An initial analysis of the signal quality of the biosensing apparel was performed using the RR coverage metric. RR coverage was calculated using the time series signal for heart rate variability based on the RR interval (Steinberg et al., 2019). Higher values indicate greater consistency in signal coverage.

We determined the degree of similarity between the three biometrics from the biosensing apparel and the reference systems in each evaluation by computing Pearson's correlation coefficient in a two-tailed t-test. We consider results with $p < 0.01$ to be statistically significant and results with $p < 0.05$ to approach significance. For the flight testing and flight squadron evaluations, data captured were also compared to results from other evaluations (e.g., centrifuge). These comparisons were used to determine if there were any changes in quality from the simulated environments on the ground to the aircraft environment. For the state detection, analysis was performed by assessing the model against a with-held test set using accuracy, precision, and F1 scores.

RESULTS

Effectiveness of Biosensing Garments

Across all evaluations, RR coverage exceeded 85%. The lowest coverage was in the flight squadron testing with 88.3%. The overall average across all evaluations was 94.5 (SD = ± 4.8). Results are shown in Table 2.

Table 2. Average RR coverage, a calculation based on the time series for heart rate variability, indicates ECG quality.

	ROBD	Altitude Chamber	Thermal Chamber	Cycle Ergometer	Centrifuge	Flight Testing	Flight Squadron
Average RR Coverage %	98.0 (2.78)	98.1 (1.9)	93.1 (3.9)	92.0 (6.5)	95.8 (4.3)	96.5 (2.6)	88.3 (11.8)

Table 3 provides the correlations for each evaluation to the reference systems which measured heart rate and SpO_2. For comparison to the Mindray ECG and the centrifuge ECG, we observe strong positive and significant correlations between the biosensing apparel heart rate and the heart rates detected by the reference systems. The heart rates recorded by the three systems were on average within 4.45 beats per minute (bpm) 81% of the time with a mean difference of 0.48 bpm and standard deviation of 8.45 bpm. The pulse rate from the Nonin WristOx resulted in weaker correlations to the heart rate detected by the biosensing apparel.

For the reference pulse oximeter systems, which included the Nonin WristOx and pulse oximetry pod, the results indicate significant negative correlations between the SpO_2 detected by the reference systems and the heart rate detected by the biosensing garment, with the exception of the cycle ergometer and flight testing. There were also significant negative correlations between SpO_2 and breathing depth across all evaluations except for the cycle ergometer. In contrast, SpO_2 detected by the reference systems was positively correlated with the breathing rate detected by the biosensing garment.

Table 3. Correlations (*r*) between reference systems and biosensing apparel biometrics. (**) indicates significantly correlated at p < .001, (*) indicates correlated at p < .01, (+) indicates correlated at p < .05. N/A indicates that there was not an available reference sensor.

	Biosensing Garment	Heart Rate		Breathing Depth		Breathing Rate	
Evaluation	**Reference Sensor(s)**	**HR**	**SpO_2**	**HR**	**SpO_2**	**HR**	**SpO_2**
ROBD	HR: Mindray ECG SpO_2: Pulse oximetry pod	.95**	-.61**	.23	-.25*	-.19	.26**
Altitude Chamber	HR: Nonin WristOx (pulse rate) SpO_2: Nonin WristOx	.48*	-.79**	.19	-.35**	-.23	.39**
Thermal Chamber	HR: Mindray ECG	.99**	*N/A*	.22	*N/A*	.67*	*N/A*
Cycle Ergometer	SpO_2: Pulse oximetry pod	*N/A*	.41	*N/A*	.09	*N/A*	.40
Centrifuge	HR: (unspecified) ECG SpO_2: Nonin WristOx	.87**	-.36*	.31**	-.67**	.16**	.19
Flight Testing	HR: Nonin WristOx (pulse rate) SpO_2: Nonin WristOx	.30**	.26**	.12**	-.11**	.14**	.10**
Flight Squadron	None	*N/A*	*N/A*	*N/A*	*N/A*	*N/A*	*N/A*

In a deep dive of the signals for individual participants, the data collected from the apparel reflect temporal dynamics. Figure 2 provides a visual depiction of the gradual onset rate acceleration up to 9G for one participant. For this participant, the GOR began just prior to 18:14:30 and proceeded until 18:15:45. During this period, the participant experienced increasing acceleration forces up to 9G followed by a deceleration, as depicted in the last graph in Figure 2. As the acceleration increased, the participant's heart rate and breathing rate correspondingly increased while their breathing depth shrank. Once the acceleration began to decrease at 18:15:40, the subject's breathing depth and heart rate began to recover.

Comparing data from the simulated ground environments (e.g., centrifuge, altitude chamber) with the flight testing and flight squadron evaluations allows us to explore if there were any changes in physiological data between the different environments. There were clear corresponding changes in heart rate and breathing rate in the aircraft environment when pilots experienced high-G forces that resulted in patterns similar to those observed in the centrifuge evaluations (Figure 3). In comparing changes in heart rates and breathing rates between different pilots, we found that pilots' physiological responses differed significantly under high-G in flight. Unfortunately, baseline resting heart rates and breathing rates were not tracked in the flight testing and flight squadron evaluations, and there were no reference systems for comparison in the flight squadron evaluation. Future work should consider this in order to facilitate further understanding regarding the accuracy of the apparel when exploring potential individual differences in physiological responses to high-G.

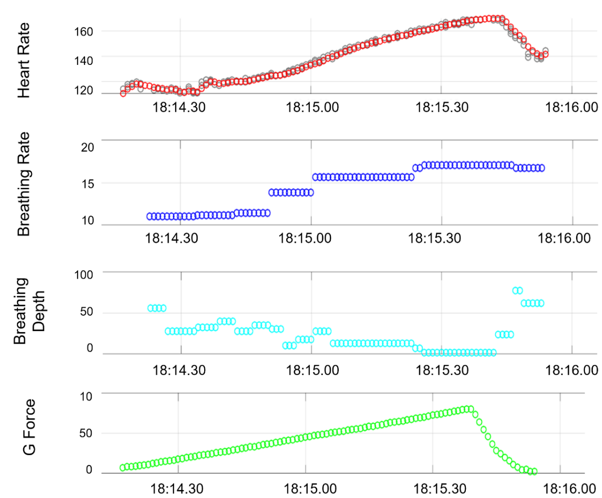

Figure 2. Heart rate, breathing rate, and breathing depth from biosensing apparel for one participant as accelerating up to 9 G_z

The "in-the-wild" flight squadron evaluation allowed us to look at how the apparel performed over time, both in terms of sensing and in durability. Regarding sensing, seven of the 17 pilots who participated in the evaluation flew four or more flights while wearing the garments. For these participants, deviations in heart rate and breathing rate can provide insight into potential changes in the apparel and/or in the physiological data captured across flights. The standard deviation for changes in average heart rate ranged from 5.9 beats per minute to 15.8 beats per minute.

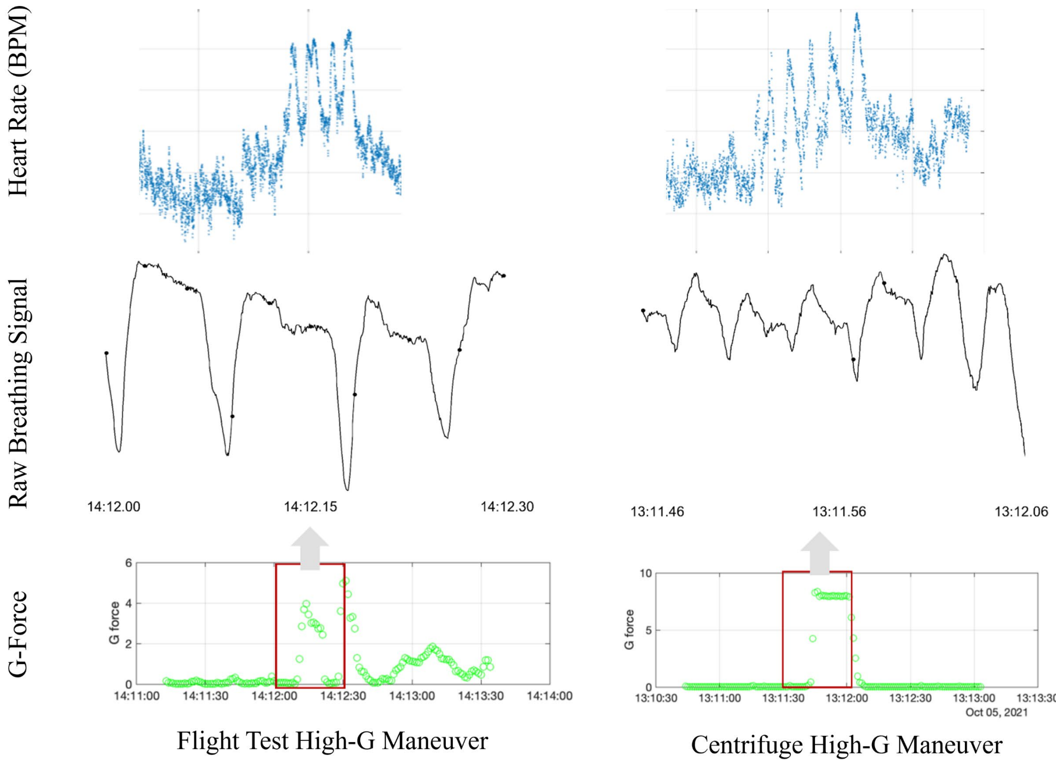

Figure 3. G-force in aircraft versus centrifuge and captured breathing signal and heart rate.

Regarding the durability of the apparel over time, flight squadron participants washed the biosensing garments themselves, following a protocol that involved no-heat drying or line-drying. Based on the measurements taken, it was observed the fabric shrank in the lengthwise direction by 1.2%, with negligible effects on fit or sizing. The fabric shrink 11.2% along the crosswise direction after approximately 25 washes. This means the garments fit more snugly over time. This did not appear to impact signal quality, and may have actually enhanced it, as a tighter fit usually means more solid connection with the sensors. However, the change in size could impact comfort. Interestingly, participants reported that they felt the biosensing apparel became more comfortable over time.

Acceptability of Biosensing Apparel

Participants filled out subjective assessments regarding the apparel. Across all evaluations, approximately 20% of the participants commented that initially they felt that the garment was "tight" and that they could feel the sensors, but over time, they no longer noticed the sensors and the garments "felt good" and were "comfortable". Some comments included that they "didn't notice the shirt unless looking for it," "the shirt and box are not even really noticeable," "the sensor is appropriately located and unnoticeable in flight," and "when thinking about it, I know it's there but otherwise it's not noticeable."

Eleven participants from the flight squadron evaluation were interviewed further regarding their perspective on the apparel. All eleven participants felt it could be useful in a flight test environment, and eight of eleven felt that it would be useful in a training environment. With regard to training, several felt concerns that the biosensing apparel could create additional demand on students, which suggests participants have potential procedural concerns regarding training applications of the biosensing apparel. Two of the participants noted a slight impact due discomfort from the module when executing high-G maneuvers. The other nine participants did not note any impact. When considering the overall comfort of the biosensing garments, seven participants were completely or mostly satisfied, one was neutral, and two were somewhat unsatisfied, with one individual choosing not to respond. Perspectives on the puck placement appear to be divided as some pilots commented that the puck has a "low profile," is "unintrusive," and has "good ergonomics" while others commented that they felt the placement of the puck was "uncomfortable." The different perspectives did not appear to be associated with different demographics.

State Detection

Two models were built using a convolution neural network to measure the extent to which the biosensing apparel can detect breathing states that can be related to physiological episodes. The breathing states included normal breathing patterns, holding one's breath, large breaths, coughing, hyperventilation, shallow breathing, and speaking. The two models explored were: (a) a model with hyperventilation and shallow breathing treated as two separate behaviors, and (b) a model with hyperventilation and shallow breathing combined into one behavior. Results are reported from three perspectives: (1) user confidence - where all breaths are recognized simultaneously and the breath type was chosen based on the user defined confidence threshold of 90%, (2) best guess - where all breaths are recognized simultaneously and the breath type chosen was the one with the highest confidence, and (3) breath type - recognition of each breath independently of all other breaths, where type was chosen only if it was above 90%. Results for each model are given in Table 4.

Table 4. Test results of CNN for detecting breathing behaviors of interest.

	Model A Shallow & Hyperventilation Separate			**Model B** Shallow & Hyperventilation Combined		
	Accuracy	**Precision**	**F1**	**Accuracy**	**Precision**	**F1**
User Confidence	61.6	87.4	76.2	82.4	94.3	90.4
Best Guess	77.0	77.0	87.0	89.0	89.0	94.2
Breath Type	94.7	87.2	72.3	97.2	94.3	88.0

We also explored individual breath recognition. These results are given in Table 5. From these results, there is evidence that recognizing different breaths using the data from the apparel is possible. However, some breath types present issues, as demonstrated by the results for the overall models where combining shallow breaths and breaths characteristic of hyperventilation resulted in greater model performance.

Table 5. Test results of CNN for detecting breathing behaviors of interest.

	Accuracy	**Precision**	**F1**
Normal breath	92.8	95.9	75.65
Breath hold	98.95	100	71.6
Big breath	95.5	90.65	83.1
Cough	97.5	76.4	72.2
Hyper/shallow	95.45	93.85	93.75
Talking	97.85	77.7	90.4

DISCUSSION

We posed three research questions regarding how biosensing apparel may be used to monitor physiological responses of fighter pilots. Our first question considered how effective biosensing garments may be in capturing physiological signals for both males and females in the environments that fighter pilots operate. To answer this question, we conducted seven human-centered, within-subject evaluations. We compared how the biosensing apparel performed against well-established reference systems. These comparisons indicate that, for the most part, the biosensing apparel aligned in expected directions with the reference systems, particularly for heart rate. The heart rate detected by the biosensing apparel was strongly positively correlated with clinical-grade ECG systems. The one reference system for which we did not observe strong alignment was the pulse rate detected by the Nonin WristOx. However, there is evidence to suggest that pulse oximeters may be less accurate under intense exercise. Iyriboz et al (1991) found that for individuals experiencing high heart rates (> 155 beats) while exercising, two different oximeters underestimated the pulse rate by up to 16 beats per minute. The strong relationship to the clinical grade ECG suggests that the apparel is effective in measuring heart rate across different environments, and we observed that this was true of the apparel for both males and females.

The alignment between the apparel and the reference system measures of SpO_2 was more complex and less clear. The biosensing garment heart rate and breathing depth tended to increase as SpO_2 decreased while breathing rate increased when SpO_2 increased. As participants experienced increasingly physically challenging environments, they were breathing faster and heart rate was increasing but they did not always take deeper breaths. SpO_2 or the amount of oxygen in the peripheral blood vessels frequently decreased in the more challenging environments where participants also tended to take shallow but faster breaths. Similar results between heart rate and SpO_2 have been found in other oxygen-deprived environments (Shin et al., 2020), suggesting that these results are accurate. There were two exceptions, however. In the first exception, the cycle ergometer, heart rate and breathing depth increased with increases in SpO_2. The cycle ergometer was a very different environment than the ROBD, altitude chamber, and centrifuge evaluations. Under intense exercise in a normal oxygen-rich environment, healthy individuals may maintain SpO_2 (Radak et al., 2014), which could result in a positive, though not strong correlation with heart rate and breathing depth, as we observed. In the second exception, flight testing, heart rate increased with SpO_2 while breathing depth was still negatively correlated, though the correlation was weak if significant. During the flight test evaluations, participants wore the apparel on average for two hours and only experienced high-G maneuvers for approximately 16% of the time they were wearing the apparel while up to 50% of the time they were on the ground. This may explain the positive correlation with heart rate in that environment in comparison to the other evaluations.

We did not have a reference system for the flight squadron evaluation but comparing the data from the flight squadron evaluation to the other evaluations (e.g., flight testing and centrifuge) suggests that the apparel appeared to capture heart rate accurately at altitude in the aircraft environment. We did observe changes within participants in the flight squadron evaluation over time that could be reflective either of inconsistencies in the apparel or of variability within individuals. Previous work found that the resting heart rate of individuals can change on a day-to-day basis up to 10bpm (Quer et al., 2020). The apparel indicated changes up to 15bpm; however, it is important to note the flight squadron evaluation did not capture a pure resting heart rate but an inferred. The participants were mostly physically active after donning the apparel, and greater variability from sample to sample might be expected due to the type of activity participants were engaging in. Together with the other evaluations, these findings highlight the challenge of physiological monitoring in the flight environment. Future work would benefit from a focused longitudinal study on physiological changes in fighter pilots over time, both on the ground and in the air.

Our second research question explored how acceptable biosensing garments are to fighter pilots for capturing physiological signals. Surveys and interviews were conducted in the seven evaluations. In interviews, the majority of participants were mostly satisfied with the apparel as a garment. In terms of accepting the biosensing apparel for physiological monitoring, responses focused primarily on its use for flight testing, where one could argue pilots face the most risk and physiological monitoring may be most applicable. However, most interviewees also felt they could accept the apparel in training how individuals handle high-G forces. A focus of this work was to ensure that the design of the apparel was acceptable to both males and females, considering that flight gear in the past has been predominantly based on male anatomical and anthropometric data. Females were represented in all evaluations except for the altitude chamber. Excluding the altitude chamber, female representation ranged from 5% (flight squadron) to 44% (thermal chamber). Evaluation of the apparel did not indicate any differences in quality of physiological data captured between males and females in any of the environments, and subjective feedback in surveys and interviews indicated that both males and females found the apparel acceptable to wear long-term.

Our third research question focused on whether the apparel has potential for detecting states that may indicate fighter pilot physiological episodes. We explored whether the apparel could be used to detect different breathing states related to potential physiological episodes. We were able to achieve an F1 score of 94.2, but differentiating between shallow breathing and hyperventilation was challenging. This could be due in part to how breaths were collected. While there are differences in how shallow breaths and breaths resulting from hyperventilation occur and participants received training on each, this was done in a controlled environment and the training may have been insufficient to truly elicit underlying differences. We may also need more data to build a model that can reliably detect the small differences inherent in these two breath types. The apparel offers a unique form factor for visualizing and detecting information related to breathing rate and depth, such as different breath types. This capability may help with monitoring physiological states where respiration is key. However, precision appears to be challenging. Future work should explore how more data and other sensors may improve state modeling.

There were some limitations observed regarding the apparel. First, in adhering to safety requirements for the biosensing garments to be no melt/no drip and Berry compliant, the garments consist of greater than 50% cotton. While all fabric was pre-washed prior to assembling the prototypes, the fabric is susceptible to change over time and use. The flight squadron evaluation did not suggest significant degradation in the signal quality for those pilots who used the apparel more frequently, but changes in the fabric were observed and there is still potential for that to result in degraded signals. Continued iteration on the fabric is needed to address this. Secondly, the required hardware to collect and store data is a limitation of the apparel. The puck connects to the garment, and this creates an opportunity for issues to occur when used in the wild. If the puck is not full connected to the garment, this can lead to poor quality data or no data. This was observed in six flight squadron events where poor signal quality, or a complete lack of data recording occurred. Future work should explore how to reduce the potential for such issues.

CONCLUSION

Biosensing apparel has received considerable attention in the medical and sports health communities. In this work, we present a novel biosensing apparel design that has potential for effectively capturing heart rate, breathing rate, and breathing depth in the extreme environments fighter pilots operate. The apparel was acceptable to both males and females, particularly for use in flight testing and training. The development of breath detection algorithms suggests that the apparel also has potential for detecting relevant states associated with physiological episodes. Future work will continue the development of biosensing systems like the apparel that can detect physiological episodes. More work is also needed to explore and adequately represent both individual differences in physiological state and how pilot physiological state may change over time in the flight environment, and systems like biosensing apparel may be useful for this as well.

ACKNOWLEDGEMENTS

Thank you to Michael Dillard and Santosh Mathan for their contributions and leadership in this work, and a very special thank you to Frederic Chanay, Annie Brisson, Eric Reiher, James Duverger, and others from OmSignal for your contributions in data collection and dedication to the biosensing garment technology. Thank you to all the volunteers, the NAWCAD Human Systems Engineering Department and NAVAIR PMA, 202 Aircrew Systems.

REFERENCES

Azeem, M., Shahid, M., Masin, I., & Petru, M. (2024). Design and development of textile-based wearable sensors for real-time biomedical monitoring; a review. *The Journal of The Textile Institute*, 1-16.

Bourdon, P. C., Cardinale, M., Murray, A., Gastin, P., Kellmann, M., Varley, M. C., ... & Cable, N. T. (2017). Monitoring athlete training loads: consensus statement. *International journal of sports physiology and performance*, *12*(s2), S2-161.

Davies, M. J., Benson, A. P., Cannon, D. T., Rossiter, H. B., Ferguson, C., Marwood, S., & Kemp, G. J. (2017). *Dissociating external power from intramuscular exercise intensity during intermittent bilateral knee-extension in humans*. *21*, 6673–6686. https://doi.org/10.1113/JP274589.

Di Domenico, I., Hoffmann, S. M., & Collins, P. K. (2022). The role of sports clothing in thermoregulation, comfort, and performance during exercise in the heat: a narrative review. *Sports Medicine-Open*, *8*(1), 58.

Engel, A. C. (2021). *I just want to do my job: The experience of female fighter pilots in the United States Air Force*. University of Montana.

Fuentes-García, J. P., Clemente-Suárez, V. J., Marazuela-Martínez, M. Á., Tornero-Aguilera, J. F., & Villafaina, S. (2021). Impact of real and simulated flights on psychophysiological response of military pilots. *International Journal of Environmental Research and Public Health*, *18*(2), 787. https://doi.org/10.3390/ijerph18020787.

Ghahari, R. R., Holden, R. J., Flanagan, M. E., Wagner, S., Martin, E., Ahmed, R., ... & Toscos, T. (2018). Using cardiac implantable electronic device data to facilitate health decision making: a design study. *International Journal of Industrial Ergonomics*, *64*, 143-154.

Iannetta, D., Inglis, E., Mattu, A., Fontana, F., Pogliaghi, S., Keir, D., & Murias, J. (2020). A Critical Evaluation of

Current Methods for Exercise Prescription in Women and Men. *Medicine & Science in Sports & Exercise*, *52*(2), 466–473. https://doi.org/10.1249/MSS.0000000000002147.

Iyriboz, Y., Powers, S., Morrow, J., Ayers, D., & Landry, G. (1991). Accuracy of pulse oximeters in estimating heart rate at rest and during exercise. *British journal of sports medicine*, *25*(3), 162-164.

Kim, S., Choi, B., Cho, T., Lee, Y., Koo, H., & Kim, D. (2017, July). Wearable bio signal monitoring system applied to aviation safety. In *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 2349-2352). IEEE.

Kumagai, K., Maruyama, S., Imamura, T., Iwamoto, T., Kanamaru, Y., Mine, M., ... & Wada, K. (2023). Effective monitoring for early detection of hypoxia in fighter pilots. *The International Journal of Aerospace Psychology*, *33*(3), 174-181.

Li, S., Li, H., Lu, Y., Zhou, M., Jiang, S., Du, X., & Guo, C. (2023). Advanced textile-based wearable biosensors for healthcare monitoring. *Biosensors*, *13*(10), 909.

Quer, G., Gouda, P., Galarnyk, M., Topol, E. J., & Steinhubl, S. R. (2020). Inter-and intraindividual variability in daily resting heart rate and its associations with age, sex, sleep, BMI, and time of year: Retrospective, longitudinal cohort study of 92,457 adults. *Plos one*, *15*(2), e0227709.

Rice, G. M., Snider, D., Moore, J. L., Lavan, J. T., Folga, R., & VanBrunt, T. B. (2016). Evidence for− Gz Adaptation Observed with Wearable Biosensors During High Performance Jet Flight. *Aerospace Medicine and Human Performance*, *87*(12), 996-1003. https://doi.org/10.3357/AMHP.4609.2016

Radak, Z., Zhao, Z., Koltai, E., Ohno, H., & Atalay, M. (2013). Oxygen consumption and usage during physical exercise: the balance between oxidative stress and ROS-dependent adaptive signaling. *Antioxidants & redox signaling*, *18*(10), 1208-1246.

Schauss, G., Arquilla, K., & Anderson, A. (2022, April). ARGONAUT: An inclusive design process for wearable health monitoring systems. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (pp. 1-12).

Shin, S., Yasukochi, Y., Wakabayashi, H., & Maeda, T. (2020). Effects of acute hypobaric hypoxia on thermoregulatory and circulatory responses during cold air exposure. *Journal of physiological anthropology*, *39*, 1-6.

Shaw, D. M., & Harrell, J. W. (2023). Integrating physiological monitoring systems in military aviation: a brief narrative review of its importance, opportunities, and risks. *Ergonomics*, *66*(12), 2242-2254.

Skiba, P. F., Jackman, S., Clarke, D., Vanhatalo, A., & Jones, A. M. (2014). Effect of Work and Recovery Durations on W’ Reconstitution during Intermittent Exercise. *Med. Sci. Sports Exerc.*, *46*(22), 1433–1440. https://doi.org/10.1249/MSS.0000000000000226

Steinberg, C., Philippon, F., Sanchez, M., Fortier-Poisson, P., O’Hara, G., Molin, F., ... & Champagne, J. (2019). A novel wearable device for continuous ambulatory ECG recording: proof of concept and assessment of signal quality. *Biosensors*, *9*(1), 17.

Tanaka, H., Monahan, K. D., & Seals, D. R. (2001). Age-predicted maximal heart rate revisited. *Journal of the American College of Cardiology*, *37*(1), 153-156. https://doi.org/10.1016/S0735-1097(00)01054-8

Whipp, B. J., & Rossiter, H. B. (2013). The kinetics of oxygen uptake: Physiological inferences from the parameters. In *Oxygen Uptake Kinetics in Sport, Exercise and Medicine*. https://doi.org/10.4324/9780203613771.