JAMIA Open, 6(2), 2023, ooad029
https://doi.org/10.1093/jamiaopen/ooad029
Research and Applications

Research and Applications

A machine learning approach to determine resilience utilizing wearable device data: analysis of an observational cohort

Robert P. Hirten[1,2], Maria Suprun[3], Matteo Danieletto[2,4], Micol Zweig[2,4], Eddye Golden[2,4], Renata Pyzik[5], Sparshdeep Kaur[2], Drew Helmus[1], Anthony Biello[1], Kyle Landell[2], Jovita Rodrigues[2], Erwin P. Bottinger[2], Laurie Keefer[1], Dennis Charney[6,7], Girish N. Nadkarni[2,8,9], Mayte Suarez-Farinas[3,4], and Zahi A. Fayad[5,10]

[1]The Dr. Henry D. Janowitz Division of Gastroenterology, Icahn School of Medicine at Mount Sinai, New York, New York, USA, [2]The Hasso Plattner Institute for Digital Health at the Mount Sinai, New York, New York, USA, [3]Department of Population Health Science and Policy, Center for Biostatistics, Icahn School of Medicine at Mount Sinai, New York, New York, USA, [4]Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, New York, USA, [5]The BioMedical Engineering and Imaging Institute, Icahn School of Medicine at Mount Sinai, New York, New York, USA, [6]Office of the Dean, Icahn School of Medicine at Mount Sinai, New York, New York, USA, [7]Nash Family Department of Neuroscience, Icahn School of Medicine at Mount Sinai, New York, New York, USA, [8]The Department of Medicine, Icahn School of Medicine at Mount Sinai, New York, New York, USA, [9]The Charles Bronfman Institute for Personalized Medicine, Icahn School of Medicine at Mount Sinai, New York, New York, USA and[10]Department of Diagnostic, Molecular and Interventional Radiology, Icahn School of Medicine at Mount Sinai, New York, New York, USA

Corresponding Author: Robert P. Hirten, MD, The Dr. Henry D. Janowitz Division of Gastroenterology, Icahn School of Medicine at Mount Sinai, 1468 Madison Avenue, Annenberg Building RM 5-12, New York, NY 10029, USA; robert.hirten@mountsinai.org

Received 4 January 2023; Revised 22 March 2023; Editorial Decision 29 March 2023; Accepted 6 April 2023

ABSTRACT

Objective: To assess whether an individual's degree of psychological resilience can be determined from physiological metrics passively collected from a wearable device.

Materials and Methods: Data were analyzed in this secondary analysis of the Warrior Watch Study dataset, a prospective cohort of healthcare workers enrolled across 7 hospitals in New York City. Subjects wore an Apple Watch for the duration of their participation. Surveys were collected measuring resilience, optimism, and emotional support at baseline.

Results: We evaluated data from 329 subjects (mean age 37.4 years, 37.1% male). Across all testing sets, gradient-boosting machines (GBM) and extreme gradient-boosting models performed best for high- versus low-resilience prediction, stratified on a median Connor-Davidson Resilience Scale-2 score of 6 (interquartile range = 5–7), with an AUC of 0.60. When predicting resilience as a continuous variable, multivariate linear models had a correlation of 0.24 ($P = .029$) and RMSE of 1.37 in the testing data. A positive psychological construct, comprised of resilience, optimism, and emotional support was also evaluated. The oblique random forest method performed best in estimating high- versus low-composite scores stratified on a median of 32.5, with an AUC of 0.65, a sensitivity of 0.60, and a specificity of 0.70.

Discussion: In a *post hoc* analysis, machine learning models applied to physiological metrics collected from wearable devices had some predictive ability in identifying resilience states and a positive psychological construct.

Conclusions: These findings support the further assessment of psychological characteristics from passively collected wearable data in dedicated studies.

Key words: machine learning, wearable device, mental health, resilience, well-being

LAY SUMMARY

Mental health issues are common however resources for their evaluation and treatment are limited. Digital technologies, such as wearable devices, provide a possible means to automate mental health assessments. Resilience, or an individual's ability to cope with adversity, is an important psychological feature which can improve stress and psychological well-being. The goal of this study is to see whether we are able to predict a person's degree of resilience, and other psychological features, using the information collected from wearable devices. Using machine learning algorithms, we evaluated the changes in the time between each heartbeat, or heart rate variability, which is collected from wearable devices. Heart rate variability reflects the body's nervous system activity and its physical response to stress. We found that it is feasible to determine a person's degree of resilience, as well as a combination of his or her resilience, optimism, and emotional support, based on an individual's heart rate variability. Further studies are needed to build on this work and further evaluate these findings.

BACKGROUND AND SIGNIFICANCE

Mental health issues are common and account for 13% of the burden of global disease, with a quarter of the population at some point suffering from psychological illness. The World Health Organization considers psychological disorders to be the leading cause of disability in the world, constituting the largest single source of global health economic burden.[1,2] However, mental health resources are limited with wide disparities in access across geography and socioeconomic status.[3] Access is further limited by the need for in-person assessment or the completion of validated mental health surveys. A better understanding of who is at psychological risk is needed. The growth of digital technology presents an opportunity to improve access to mental health services through telemedicine, smartphone applications that monitor well-being, and wearable technology.[4]

Wearable devices collect continuous physiological data remotely and without active input from users.[5] Interrogation of these data through machine learning techniques offers a novel means to automate mental health assessments. This approach has been explored in a limited number of psychological conditions, including anxiety, emotional state transitions, and loneliness, in often small studies that include both wearable and mobile phone data.[6–10] Resilience, or an individual's ability to cope with and recover from adversity, is an area of growing interest, especially as the coronavirus disease 2019 (COVID-19) pandemic has increased psychological distress.[11] Resilience mitigates stress and improves psychological well-being. It impacts physical health, particularly in chronic disease states, where it reduces morbidity and healthcare utilization.[12,13] Resilience monitoring and modification have been employed in the patient-centered medical home model of care and by resilience-building health companies. Thus, resilience is an attractive parameter to monitor passively and remotely. Complementary to resilience as a sole parameter, is the increasing emphasis placed on positive psychological constructs as a buffer against negative psychological effects.[14,15] These constructs foster a resilient mindset, or resilient cognitions that facilitate creating a proactive way of addressing challenges.[16] Other psychological features, such as optimism and emotional support, which are complementary to resilience have direct links to healthy behaviors and can be used to create a positive psychological composite of the resilient mindset.[15,17] Optimism, defined as positive expectations about the future, is modifiable and is associated with improved emotional regulation, mood, confidence, adaptive coping mechanisms, and higher resilience.[18–21] Emotional support, or an individual's support through social ties with others, enhances psychological well-being, reduces depression, improves stressor response, and fosters resilience.[22] Thus, the ability to passively evaluate positive psychological constructs using wearable devices is also of interest in light of their important effects on well-being.

The autonomic nervous system (ANS), a primary component of the physiological stress response, can be assessed by measuring heart rate variability (HRV), the small time differences between each heartbeat.[23] Prior studies have demonstrated that psychological features such as resilience, impact the physiological stress response, with high vagally mediated HRV associated with higher resilience scores.[24–27] Similarly, optimism and emotional support have been shown to impact ANS function and increase parasympathetic tone.[28] Our group extended these observations using wearable device data to demonstrate that average circadian features of HRV, across the study population, were significantly different based on one's degree of resilience and emotional support.[29] Building on these observations that physiological parameters differ according to psychological state, we sought to examine whether machine learning algorithms can determine an individual's resilience by examining physiological markers collected from Apple Watches. This extends our prior population-level findings to the individual person, and using machine learning algorithms lays the framework for personalized psychological care and a novel approach to evaluate an individual's psychological well-being that can be applied in the clinical setting.

OBJECTIVES

We present here a secondary analysis of the Warrior Watch Study dataset.[29–31] The primary aim of this secondary analysis is to determine whether an individual's degree of resilience can be ascertained from physiological metrics passively collected from a wearable device. Our secondary exploratory aim used the same approach. It evaluated whether a positive psychological construct comprised of a composite of optimism, emotional support, and resilience assessments, can be determined. This was explored given the ability of such positive constructs to buffer against negative psychological effects, as described above.

MATERIALS AND METHODS

Study design and procedures

The Warrior Watch Study was an observational cohort study that enrolled healthcare workers (HCWs) across 7 hospitals in New York City (The Mount Sinai Hospital, Morningside Hospital, Mount Sinai West, Mount Sinai Beth Israel, Mount Sinai Queens, New York Eye and Ear Infirmary, and Mount Sinai Brooklyn). Participants had to be an employee at one of the hospitals, have an iPhone Series 6 or higher, have an Apple Watch Series 4 or 5, and be ≥18 years of age to enroll. Employees with underlying chronic diseases or medications that impact ANS function were excluded from enrollment in the study. The study was approved by the Mount Sinai institutional review board.

The initial objective of the Warrior Watch Study was to determine if wearable-based physiological signatures can identify and predict severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infections and to evaluate longitudinal stress in HCWs. Participants downloaded our Warrior Watch app to their smartphones and used it to complete surveys at enrollment and longitudinally. Subjects were recommended to wear their Apple Watch for at least 8 h per day for the duration of the study. Validated surveys that assessed psychological well-being were completed at baseline. The Connor-Davidson Resilience Scale-2 (CD-RISC-2) is a 2-question survey that measures resilience, with higher scores reflecting higher resilience.[32–34] The CD-RISC scales represent the most validated resilience scales.[35] The 2-item version includes item 1 and item 5 of the CD-RISC-10, which each has a 5-point Likert scale, and represents the 2 concepts felt to best define resilience-adaptability and the ability to “bounce back.” The Patient-Reported Outcomes Measurement Information System (PROMIS) emotional support questionnaires are validated surveys which assess individuals’ emotional support by inquiring about whether a person has someone who listens to them and with who they can discuss their feelings.[36,37] The 2-item version is scored from 2 to 10 points, with higher scores correlating with higher support. The Life Orientation Test assesses optimism and is composed of 6 questions, graded on a 5-point Likert scale, that ask subjects whether he or she expects things to go wrong, whether more good is expected compared to bad, and whether the best is expected in uncertain times.[18,38,39] It is considered the gold standard for assessing optimism and has a high score of 24, which reflects greater optimism.[40] Based upon our observation that circadian features of HRV differ based on the degree of resilience and emotional support, we performed a *post hoc* analysis of the data collected from the Warrior Watch Study to evaluate the study objectives listed above.[29]

Wearable device

The Apple Watch Series 4 or 5 was worn by subjects capturing heartrate and HRV, which is a measure of the small-time difference between each heartbeat. Values were transmitted via Bluetooth to a subject’s iPhone. Apple Watch’s photoplethysmogram sensor contains a green light-emitting diode with a photodiode creating time series peaks.[41] During 60-s recording periods, the Apple Health App can calculate HRV, using these time series peaks as the standard deviation of NN intervals (SDNN). This is a time domain index reflecting sympathetic and parasympathetic nervous activity.[23] Each SDNN datapoint is calculated from inter-beat intervals. Apple’s proprietary algorithms incorporate photoplethysmography and accelerometer information to clean the data prior to processing for the calculation of SDNN measurements. HRV measurements are obtained sporadically throughout the day by the Apple Watch and are collected through the ehive app. The Apple Watch has been used for the study of health and disease, with several studies validating the heart rate and HRV measurements derived from the device.[42–44]

Statistical analysis

Study population

The cohort’s demographic and clinical history was summarized as frequencies and proportions. Participants were split into 2 groups based on the median cutoff of resilience, defined as low versus high resilience. Demographics were compared with the Chi-squared test for categorical and independent variables and the *t*-test for continuous variables.

HRV modeling

The first 14 days of HRV data recorded from the Apple Watch after enrollment were included for analysis. For participants with a documented positive SARS-CoV-2 infection, data points at least 14 days prior to infection were used. As described previously in detail,[30] several parameters that characterize HRV patterns were extracted using a *cosinor* model. In brief, for each participant, a *cosinor.lm* function from the *cosinor* R package was applied to model daily circadian rhythm over a 24-h period.[45] Several parameters were then calculated: MESOR (M, the midline statistics of rhythm), Amplitude (Amp, a measure of half of the extent of variation within a day), Acrophase (Acr, a measure of the time of day at which overall high values recur), and a time of day at which the single highest Amplitude was observed (T_Amp).

A cosinor modeling approach was used to evaluate the relationship between HRV parameters and resilience on a population level using *fit.cosinor.mixed* function from the *cosinoRmixedeffects* R package.[46] In this mixed-effects model, HRV was an outcome, and dependent variables were body mass index (BMI), age, and an interaction term of biological sex (male vs female) with resilience (high vs low, based on the median cutoff), with a random intercept for each participant.

Machine learning

HRV variables (M, Amp, Acr, T_amp), resting heart rate (RHR) (minimum, maximum, mean, standard deviation), and basic demographics (age, sex, BMI) were used to determine resilience. To create a positive psychological construct, resilience was combined with the emotional support (“ReEm”) and optimism (“OReEm”) metrics to represent a new composite outcome. As part of our feasibility exploration, this composite index followed an implicit approach, combining each scales raw scores. Resilience (CD-RISC-2) and OReEm were considered as continuous or binary (split by the median) values. For all modeling, data were split 3:1 into training and testing sets. Model selection and optimization were done using only data from the training set and then evaluated on the “unseen” test set.

For the binary outcome, first, features were selected using either BORUTA or Recursive Feature Elimination (RFE) method.[47] BORUTA iteratively removes features that are statistically less predictive than random noise, referred to as a “shadow.” RFE first fits a model using all predictors, which are then iteratively eliminated, and the model is re-evaluated until the best combination is found. Since, for both algorithms, the original set of predictors can produce different results; we have evaluated feature selection using HRV and demographics, with or without RHR variables. Then, for each set of predictors, 12 distinct algorithms were trained (*glmnet*, *bayesglm*, *rf*, *ORFridge*, *xgbTree*, *rpart*, *knn*, *lda2*, *gamSpline*, *gamboost*, *nnet*, *gbm*). The training was done using 10-fold cross-validation for hyperparameter tuning, maximizing the area under the curve-receiver operating characteristic curve (AUC-ROC). Models' performance was evaluated based on Accuracy, AUC, Sensitivity (recall), Specificity, Negative and Positive Predictive Value (NPV, PPV, respectively), Kappa, F1, and Brier metrics.

For the continuous resilience outcomes, 3 sets of predictors were considered: HRV alone, HRV with demographics, and HRV with demographics and RHR. For each set, 26 algorithms were trained (*lm*, *ridge*, *lasso*, *bayesglm*, *glmboost*, *gaussprLinear*, *svmLinear*, *kernelpls*, *pls*, *spls*, *bridge*, *BstLm*, *cubist*, *enet*, *icr*, *lars*, *M5Rules*, *rqnc*, *nnls*, *penalized*, *pcr*, *rqlasso*, *relaxo*, *rvmLinear*, *spikeslab*, *brnn*), using 10-fold cross-validation and minimizing the root mean square error (RMSE). Model performance was evaluated by RMSE, R-squared, and a correlation of the observed and predicted outcome.

RESULTS

While our manuscript assessing psychological resilience and HRV changes included 361 subjects, a total of 329 subjects were enrolled between April 29 and September 29, 2020, in this observational study analysis, in whom baseline psychological assessments and 14 days of HRV data were available. The mean age of participants was 37.4 years, with 122 (37.1%) subjects being male (Table 1). The median CD-RISC-2 score was 6 (interquartile range [IQR] = 5–7). The median emotional support score was 8 (IQR = 6–8), and the median optimism score was 20 (IQR = 17–23). Based on the median CD-RISC-2 score, participants were stratified into "low" ($n = 241$) and "high" ($n = 88$) resilience groups (Table 1). Subjects on average had 67 HRV measurements available over the study period (95% confidence interval [CI] 59–75). This corresponds to an average of 7.6 HRV readings per day (95% CI 6.7–8.3).

A *cosinor* mixed-effects model evaluating the relationship of HRV (SDNN) features with resilience demonstrated significantly higher acrophase in the low-resilience group ($P < .001$). When stratified by sex, a significantly lower acrophase of the circadian pattern of SDNN was noted for both males ($P = .038$) and females ($P < .001$) with high compared to low resilience (Figure 1). In males, a significant difference in the amplitude ($P = .006$) was observed in low compared to high resilience.

Determination of high and low resilience

The data were split such that 75% ($n = 247$) of the data were included in the training set, and 25% ($n = 82$) was included in the testing set. This produced a more realistic estimate of model performance in the "unseen" set of participants. Several different machine learning models were explored to identify whether a subject has high compared to low resilience, defined based on the median CD-RISC-2. Twelve models (Supplementary Table S1) were assessed, with gradient-boosting machines (GBM) and extreme gradient boosting (xgbTree) performing best (Figure 2A and B). While a high AUC was observed in training (0.87) for xgbTree, in testing, it dropped to 0.60. The GBM model had more comparable but only moderate performance in training and testing, with AUCs of 0.67 and 0.60, respectively. Both models had higher Specificity/Positive Predictive Value than Sensitivity/Negative Predictive Value (ie, better at predicting high resilience). Both models had 2 features in common, including sex and HRV amplitude, while xgbTree also had age and BMI.

Determination of resilience as a continuous variable

Resilience was evaluated as a continuous variable and explored using 26 machine learning models to understand the correlation of predicted compared to actual resilience in training and testing sets. Interestingly, all models had consistent performance (Supplementary Table S2), with all of them having low-performance estimates. For example, a multivariable linear model (LM) in training had a correlation of 0.33 ($P < .001$) and RMSE of 1.34, which dropped to a correlation of 0.24 ($P = .029$) and RMSE of 1.37 in the testing data (Figure 2C). The features utilized by the best-performing LM included sex, age, BMI, HRV mesor, and amplitude (Figure 2D).

Based on our observation of an inflection point at a CD-RISC-2 score of 5 in the continuous predictive analysis (Figure 2C), we further evaluated the predictive ability to characterize high and low resilience based on this value instead of the median. Twelve models were explored (Supplementary Table S3). Again, GBM and xgbTree were best at determining high versus low resilience. However, it was not better than using the median CD-RISC-2 score (Figure 2E and F).

Determination of a positive psychological construct

To evaluate a positive psychological construct and increase the numerical spread of values for analysis, we further assessed whether the predictive ability of a machine learning algorithm is improved with the addition of other baseline scores assessing psychological features. Optimism and emotional support assessments were added to resilience scores (OReEm). This cumulative score had a median and mean of 32.5 and 32.1 (SD = 5.57), respectively (Figure 3A). Twelve machine learning models were assessed (Supplementary Table S4). GBM and oblique random forest method (ORFridge) were best at determining high versus low scores in training, based on the median (Figure 3B and C). However, the results were poor and did not translate in testing, with GBM producing an AUC of 0.56, an accuracy of 0.54, a sensitivity of 0.45, and a specificity of 0.67. The ORFridge model was moderately effective and produced an AUC of 0.65, with an accuracy of 0.64 and sensitivity and specificity of 0.60 and 0.70, respectively. Both models included 7 features: sex, age, BMI, HRV mesor, acrophase, amplitude, and time of highest amplitude.

DISCUSSION

This study aimed to determine if machine learning models applied to wearable device outputs can be trained to differentiate an individual's degree of resilience. To the best of our knowledge, this is the first study to evaluate this approach and was motivated by our previous observation of differences in resilience tertials and HRV metrics.[29] This *post hoc* analysis of the Warrior Watch dataset produced multiple models with AUC's that ranged from low to fair in their predictive ability. However, this demonstrates that it is feasible to assess these psychological features via passively collected wearable metrics and that dedicated studies are warranted to further explore this approach.

Table 1. Demographic and medical history

	Overall	Low resilience (≤median)	High resilience (>median)	*P*-value
Sample size	329	241	88	
Age, mean (SD)	37.4 (9.6)	37.3 (9.9)	37.6 (8.5)	.814
Male sex (%)	122 (37.1)	71 (29.5)	51 (58.0)	<.001
BMI, mean (SD)	25.6 (4.9)	25.8 (5.2)	25.0 (3.8)	.182
BMI category (%)				.330
Underweight	3 (0.9)	3 (1.2)	0 (0.0)	
Normal	177 (53.8)	126 (52.3)	51 (58.0)	
Overweight	91 (27.7)	65 (27.0)	26 (29.5)	
Obese	58 (17.6)	47 (19.5)	11 (12.5)	
Race (%)				
Asian	70 (21.3)	49 (20.3)	21 (23.9)	.589
South Asian	27 (8.2)	22 (9.1)	5 (5.7)	.435
Black	38 (11.6)	29 (12.0)	9 (10.2)	.796
Middle Eastern	10 (3.0)	8 (3.3)	2 (2.3)	.899
Native American/Pacific Islander	4 (1.2)	2 (0.8)	2 (2.3)	.625
White	119 (36.2)	82 (34.0)	37 (42.0)	.226
Hispanic or Latino (%)	63 (19.1)	51 (21.2)	12 (13.6)	.168
Smoking—never/rarely (%)	278 (84.5)	203 (84.2)	75 (85.2)	.961
CD-RISC-2, mean (SD)	5.7 (1.4)	5.0 (1.1)	7.4 (0.5)	<.001
Emotional support, mean (SD)	6.9 (1.5)	6.7 (1.6)	7.5 (1.2)	<.001
Optimism, mean (SD)	19.5 (4.2)	18.8 (4.1)	21.4 (3.8)	<.001

SD, standard deviation.

There is growing interest in leveraging wearable technology for the management of health and disease. Limited access to mental health services makes psychological and well-being assessment an attractive area for passive monitoring. There have been limited studies evaluating wearable or mobile data for the assessment of mental health. Shaukat-Jali et al[9] demonstrated that subclinical social anxiety could be detected in 12 subjects using physiological data collected from wearable devices during speech tasks meant to provoke anxiety. Sultana et al[6] leveraged a publicly available data set comprised of mobile and wearable data from 18 individuals to demonstrate that transitions in emotional states can be detected. In a larger study of 160 subjects, Doryab et al[7] used Fitbit and mobile phone data to assess loneliness in college students. Data collected include sleep and activity data, as well as smartphone metrics such as screen status, call logs, and location data. They found that machine learning models reliably detected loneliness. Along similar lines of exploration, Sano et al[8] evaluated the survey, diary, wearable, and mobile phone usage data in 201 college students. Machine learning algorithms applied to this multimodal data were able to identify objective features that could classify self-reported stress and mental health groups. Sükei et al[10] evaluated mobile phone data and wearable data, if available, from 943 people in a previously collected dataset. They demonstrated the feasibility of using machine learning models for predicting emotional states.

While these prior works relied on multimodal data sources, our approach was unique in our interrogation of only passively collected watch data. The ability to determine psychological states from a single data source that does not rely on a second device (ie, Cellphone) can expand the applicability of these findings to those who do not regularly use smartphones or those who do not wish to answer questions about their psychological well-being. Through our exploration of machine learning models, we found GBM and xgbTree best at determining resilience with an AUC of 0.60. Both models had higher specificity and positive predictive value, which was 73% and 82% for GBM, highlighting the ability to identify high resilience. The Warrior Watch Study was designed to derive machine learning models to predict SARS-CoV-2 infections and understand longitudinal psychological stress in HCWs. Its assessment of resilience relied on the CD-RISC-2 survey, which is a commonly used 2 question survey for assessing this feature. Our secondary analysis evaluated whether positive psychological constructs, in this case comprised of optimism, emotional support, and resilience could be passively determined from wearable device data. This resulted in a slight improvement in our predictive ability, with the ORFridge model being moderately effective at determining composite resilience, optimism, and emotional support with an AUC of 0.65. The composite measures slight improvement in AUC may mean that these additional psychological features should be explored in greater detail using similar approaches.

Resilience is a complex psychological metric to assess, with differing definitions that encompass dimensional conceptualizations that evaluate varying degrees of resilience, as well as longitudinal viewpoints.[48] The CD-RISC surveys, from which our 2-question version was derived, have the best psychometric ratings.[49] However, assessments of other aspects of this construct need to be explored as well. Building upon our observation that it is feasible to use physiological metrics to determine an individual's degree of resilience, employment of these other and more detailed scales is needed and may result in superior outcomes. Through the exploration of different characterizations of resilience, improved physiological mapping with wearables may be possible. Furthermore, our demonstration that positive psychological constructs and the resilient mindset can be evaluated via passively collected wearable data, may mean that our approach can be applied more broadly to psychological assessment and the evaluation of multidimensional psychological traits. This is important in light of the ability of such positive constructs to improve mental and physical health and their direct link to health behavior.

Our findings extend prior observations of how the physiological stress response or ANS reflects resilience state. While prior work

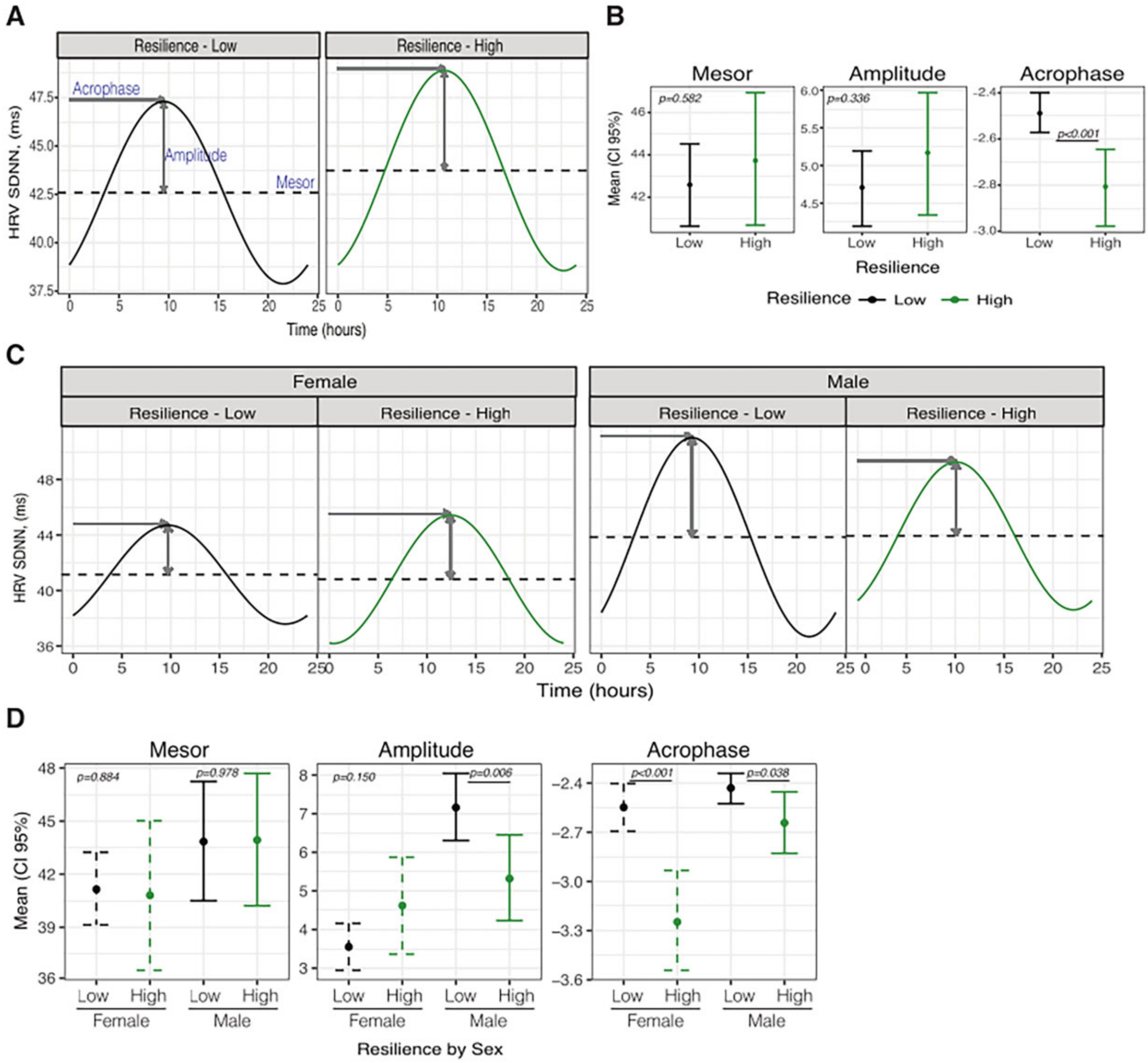


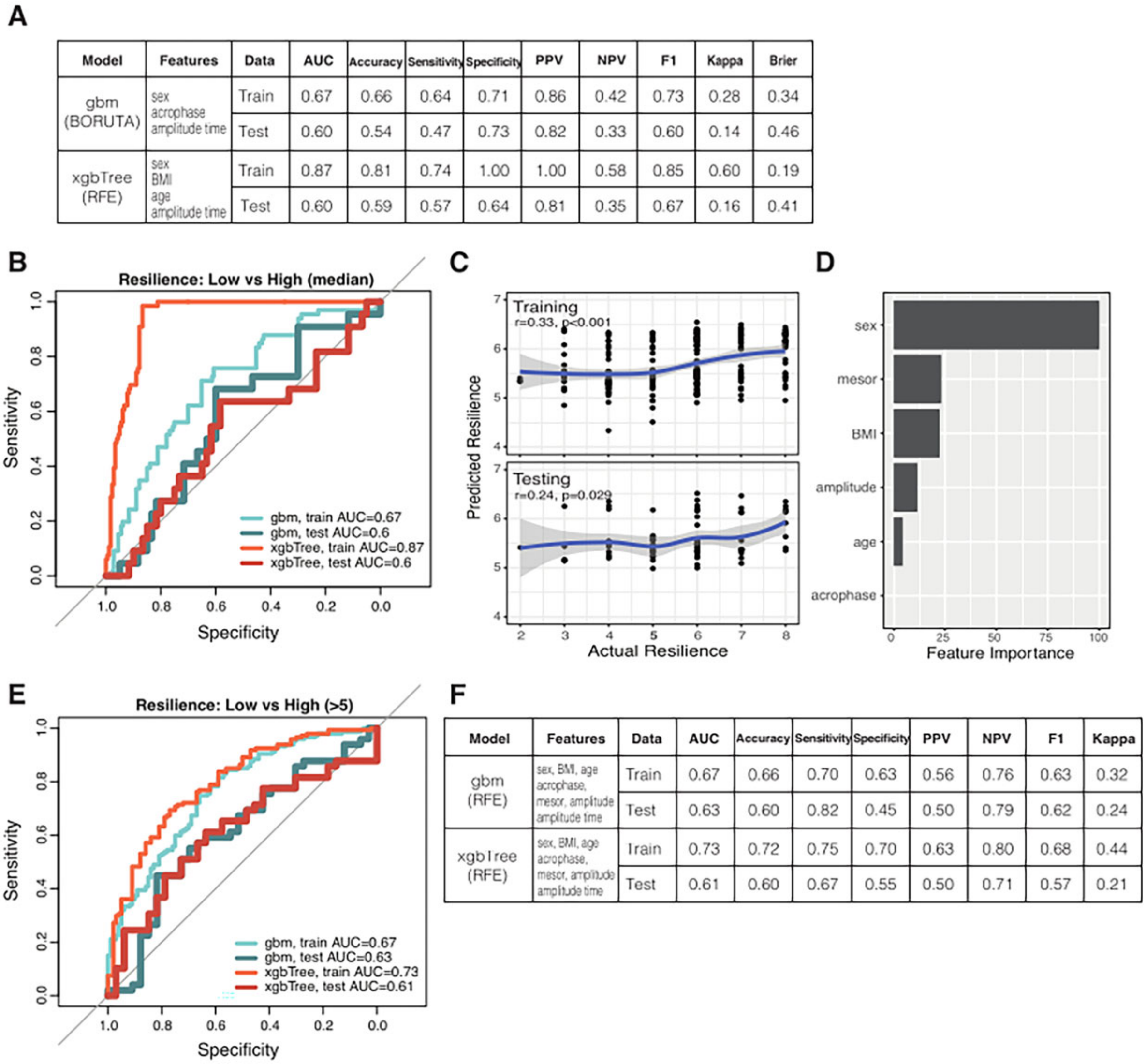
Figure 1. Population-level relationship between HRV parameters and resilience. Plots (A) and (B) show univariate association with resilience, while (C) and (D) show resilience by sex interaction. (A) and (C) represent an average daily circadian HRV rhythm, while (B) and (D) show a pairwise comparison between groups for each HRV parameter.

demonstrated that high vagally mediated HRV had been associated with higher resilience scores, this has been frequently based on single time point assessments of autonomic function. Wearable technology has enabled the frequent characterization of HRV and evaluation of the ANS's circadian pattern.[50] Our group has demonstrated that changes in ANS circadian function can be modeled to identify and predict physiological events.[29–31] To the best of our knowledge, our characterization of how circadian changes in autonomic function relate to resilient state is the first exploration of this relationship, with this current analysis further extending our prior findings to the individual level.[29] By evaluating changes in circadian HRV measurements we are able to better understand how the ANS circadian pattern is being altered, providing a more detailed characterization of an individual's physiological stress response and its relationship to resilience. Additionally, this approach can mitigate the impact that acute stressors have on one-time HRV readings. Further evaluation of this approach in the resilience literature is needed.

The goal of this research is to utilize wearable devices in patient and psychological care, allowing the remote and passive assessment of resilient or positive psychological states. Additionally, it could be utilized to monitor an individuals' response to psychological interventions. However, studies are first needed to further evaluate this concept. Understanding of how physiological metrics relate to other characterizations of resilience or other positive psychological constructs should be explored. Additionally, the reproducibility of determining resilience state over time and in response to resilience-building interventions would be important to delineate. Understanding how dynamic changes in ANS circadian features relate to changes in resilience needs to be further explored. Through such evaluations it may be possible to employ this approach in clinical practice.

Limitations

There are several limitations to our analysis. The primary limitation is that the Warrior Watch dataset was not designed with the intent of applying a machine learning algorithm to assess psychological well-being from wearable parameters. The validated but brief



Model	Features	Data	AUC	Accuracy	Sensitivity	Specificity	PPV	NPV	F1	Kappa	Brier
gbm (BORUTA)	sex acrophase amplitude time	Train	0.67	0.66	0.64	0.71	0.86	0.42	0.73	0.28	0.34
		Test	0.60	0.54	0.47	0.73	0.82	0.33	0.60	0.14	0.46
xgbTree (RFE)	sex BMI age amplitude time	Train	0.87	0.81	0.74	1.00	1.00	0.58	0.85	0.60	0.19
		Test	0.60	0.59	0.57	0.64	0.81	0.35	0.67	0.16	0.41

Model	Features	Data	AUC	Accuracy	Sensitivity	Specificity	PPV	NPV	F1	Kappa
gbm (RFE)	sex, BMI, age acrophase, mesor, amplitude amplitude time	Train	0.67	0.66	0.70	0.63	0.56	0.76	0.63	0.32
		Test	0.63	0.60	0.82	0.45	0.50	0.79	0.62	0.24
xgbTree (RFE)	sex, BMI, age acrophase, mesor, amplitude amplitude time	Train	0.73	0.72	0.75	0.70	0.63	0.80	0.68	0.44
		Test	0.61	0.60	0.67	0.55	0.50	0.71	0.57	0.21

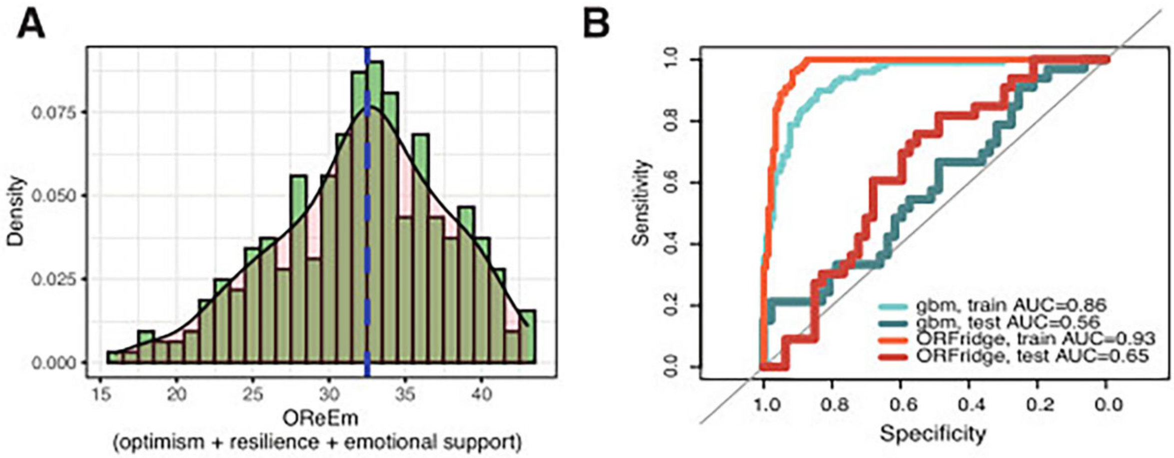
Figure 2. Machine learning to predict resilience from HRV and demographic data. (A) Performance statistics in training and testing of top 2 models predicting resilience using the median cut of and corresponding ROC plots (B). (C) Predicting resilience as a continuous outcome using linear regression and its feature importance (D). (E) Corresponding ROC plots and performance statistics in training and testing of top 2 models predicting resilience using 5 as a cutoff (F).

surveys utilized were chosen to help improve compliance rather than evaluate the endpoint of this *post hoc* analysis. However, a more detailed evaluation of resilience and other psychological features might better delineate subjects' psychological states and improve the predictive ability of the algorithms. This limitation is important, as the resilience and emotional support surveys contained only 2 questions each, while the optimism assessment contained only 6 questions. However, this secondary analysis was performed to provide proof of concept and demonstrate feasibility.

Additionally, limitations within the Warrior Watch Study dataset include the Apple Watch providing HRV only in one domain (SDNN), which precludes us from evaluating other HRV features in our machine learning. Furthermore, the HRV metrics are obtained sporadically throughout the day. Although our statistical modeling takes this into account, more frequent readings may strengthen the analysis. It is important to recognize that HRV can be impacted by different life events, which can confound the associations. To mitigate this, we evaluated the ANS function of each individual over a long observation window (14 days). This window was also informed by the relative stability of resilience and CD-RISC scores over at least a 4-week period.[51] Thus, by choosing a 14-day HRV observation window we ensure it falls within the 4-week period after CD-RISC assessment. Additionally, our characterization and use of 24-h circadian features of ANS activity further mitigate the impact that discrete transient stressors have on our assessment.

CONCLUSIONS

We derived machine learning models applied to the physiological metrics collected from wearable devices, which demonstrated some determinative ability in identifying resilience state. Considering the limitations of this *post hoc* analysis, our results provide insight into the feasibility of assessing psychological characteristics from passively collected wearable data. Further evaluation of the assessment



C

Model	Features	Data	AUC	Accuracy	Sensitivity	Specificity	PPV	NPV	F1	Kappa	Brier
gbm (RFE)	sex, BMI, age acrophase, mesor, amplitude amplitude time	Train	0.86	0.80	0.80	0.91	0.93	0.76	0.86	0.69	0.15
		Test	0.56	0.45	0.45	0.67	0.66	0.46	0.53	0.11	0.46
ORFridge (RFE)	sex, BMI, age acrophase, mesor, amplitude amplitude time	Train	0.93	0.85	0.85	1.00	1.00	0.83	0.92	0.83	0.09
		Test	0.65	0.60	0.60	0.70	0.74	0.55	0.66	0.28	0.36

Figure 3. Machine learning to predict combined optimism, resilience, and emotional support from HRV and demographic data. (A) Histogram showing the distribution of the new OReEm score. (B) ROC plot and (C) performance statistics in training and testing of top 2 models predicting OReEm using the median score.

of resilience and other psychological features is warranted in dedicated studies.

FUNDING

Support for this study was provided by the Ehrenkranz Lab For Human Resilience, the BioMedical Engineering and Imaging Institute, The Hasso Plattner Institute for Digital Health at Mount Sinai, The Mount Sinai Clinical Intelligence Center, The Dr. Henry D. Janowitz Division of Gastroenterology and by K23DK129835 (RPH).

AUTHOR CONTRIBUTIONS

RPH, MS, GNN, MS-F, and ZAF made substantial contributions to the conception or design of the work. RPH, MS, MD, LK, GNN, MS-F, and ZAF contributed to the acquisition, analysis, or interpretation of the data for the work. RPH, MS, MD, EG, MZ, SK, DH, AB, RP, DC, KL, JR, EPB, LK, GNN, MS-F, and ZAF made contributions to the drafting of the work or revising it critically for important intellectual content. RPH, MS, MD, EG, MZ, SK, DH, AB, RP, DC, KL, JR, EPB, LK, GNN, MS-F, and ZAF provided final approval of the version to be published and agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are addressed.

SUPPLEMENTARY MATERIAL

Supplementary material is available at *JAMIA Open* online.

CONFLICT OF INTEREST STATEMENT

DC is a coinventor on patents filed by the Icahn School of Medicine at Mount Sinai (ISMMS) relating to the treatment for treatment-resistant depression, suicidal ideation, and other disorders. ISMMS has entered into a licensing agreement with Janssen Pharmaceuticals, Inc, and it has received and will receive payments from Janssen under the license agreement related to these patents for the treatment of treatment-resistant depression and suicidal ideation. Consistent with the ISMMS Faculty Handbook (the medical school policy), DC is entitled to a portion of the payments received by the ISMMS. Because SPRAVATO has received regulatory approval for treatment-resistant depression, through the ISMMS, DC will be entitled to additional payments beyond those already received under the license agreement. DC is a named coinventor on several patents filed by ISMMS for a cognitive training intervention to treat depression and related psychiatric disorders. The ISMMS has entered into a licensing agreement with Click Therapeutics, Inc and has received and will receive payments related to the use of this cognitive training intervention for the treatment of psychiatric disorders. In accordance with the ISMMS Faculty Handbook, DC has received a portion of these payments and is entitled to a portion of any additional payments that the medical school may receive from this license with Click Therapeutics. DC is a named coinventor on a patent application filed by the ISMMS for the use of intranasally administered Neuropeptide Y for the treatment of mood and anxiety disorders. This intellectual property has not been licensed. DC is a named coinventor on a patent application in the United States and several issued patents outside the United States filed by the ISMMS related to the use of ketamine for the treatment of posttraumatic stress disorder. This intellectual property has not been licensed. EPB reports consultancy agreements with Deloitte and Roland Berger; ownership interest in Digital Medicine E. Böttinger GmbH, EBCW GmbH, and Ontomics, Inc; receiving honoraria from Bayer, Bosch Health Campus, Sanofi, and Siemens; and serving as a scientific advisor or member of Bosch Health Campus and Seer Biosciences Inc. LK declares research funding from Abbvie and Pfizer, consulting for Abbvie and Pfizer, and equity ownership/stock options in MetaMe Health and Trellus Health. MS-F declares research support from Novartis and Allergenis. GNN reports employment with, consultancy agreements with, and ownership interest in Pensieve Health and Renalytix AI; receiving consulting fees from AstraZeneca, BioVie, GLG Consulting, and Reata; and serving as a scientific advisor or member of Pensieve Health and Renalytix AI. ZAF discloses consulting fees from Alexion, GlaxoSmithKline, and Trained Therapeutix Discovery and research funding from Daiichi Sankyo, Amgen, Bristol Myers Squibb, and Siemens Healthineers. ZAF receives financial compensation as a board member and advisor to Trained Therapeutix Discovery and owns equity in Trained Therapeutix Discovery as a cofounder. The remaining authors declare no conflicts of interest.

DATA AVAILABILITY

The database that supports this study's findings includes information about healthcare workers. Due to privacy issues, researchers interested in gaining access to the data can contact the corresponding author.

REFERENCES

1. Bucci S, Schwannauer M, Berry N. The digital revolution and its impact on mental health care. *Psychol Psychother* 2019; 92 (2): 277–97.
2. Mental Health. https://www.who.int/news-room/facts-in-pictures/detail/mental-health. Accessed July 15 2022.
3. Coombs NC, Meriwether WE, Caringi J, Newcomer SR. Barriers to healthcare access among U.S. adults with mental health challenges: a population-based study. *SSM Popul Health* 2021; 15: 100847.
4. Kaonga NN, Morgan J. Common themes and emerging trends for the use of technology to support mental health and psychosocial well-being in limited resource settings: a review of the literature. *Psychiatry Res* 2019; 281: 112594.
5. Insel TR. Digital phenotyping: a global tool for psychiatry. *World Psychiatry* 2018; 17 (3): 276–7.
6. Sultana M, Al-Jefri M, Lee J. Using machine learning and smartphone and smartwatch data to detect emotional states and transitions: exploratory study. *JMIR Mhealth Uhealth* 2020; 8 (9): e17818.
7. Doryab A, Villalba DK, Chikersal P, *et al.* Identifying behavioral phenotypes of loneliness and social isolation with passive sensing: statistical analysis, data mining and machine learning of smartphone and Fitbit data. *JMIR Mhealth Uhealth* 2019; 7 (7): e13209.
8. Sano A, Taylor S, McHill AW, *et al.* Identifying objective physiological markers and modifiable behaviors for self-reported stress and mental health status using wearable sensors and mobile phones: observational study. *J Med Internet Res* 2018; 20 (6): e210.
9. Shaukat-Jali R, van Zalk N, Boyle DE. Detecting subclinical social anxiety using physiological data from a wrist-worn wearable: small-scale feasibility study. *JMIR Form Res* 2021; 5 (10): e32656.
10. Sükei E, Norbury A, Perez-Rodriguez MM, Olmos PM, Artés A. Predicting emotional states using behavioral markers derived from passively sensed data: data-driven machine learning approach. *JMIR Mhealth Uhealth* 2021; 9 (3): e24465.
11. Herrman H, Stewart DE, Diaz-Granados N, Berger EL, Jackson B, Yuen T. What is resilience? *Can J Psychiatry* 2011; 56 (5): 258–65.
12. Keefer L, Gorbenko K, Siganporia T, *et al.* Resilience-based integrated IBD care is associated with reductions in health care use and opioids. *Clin Gastroenterol Hepatol* 2022; 20 (8): 1831–8.
13. Martin CM. What matters in "multimorbidity"? Arguably resilience and personal health experience are central to quality of life and optimizing survival. *J Eval Clin Pract* 2018; 24 (6): 1282–4.
14. Sheldon KM, King L. Why positive psychology is necessary. *Am Psychol* 2001; 56 (3): 216–7.
15. Seligman ME, Csikszentmihalyi M. Positive psychology. An introduction. *Am Psychol* 2000; 55 (1): 5–14.
16. Wong PTP, Mayer CH, Arslan G. Editorial: COVID-19 and existential positive psychology (PP2.0): the new science of self-transcendence. *Front Psychol* 2021; 12: 800308.
17. Bartley EJ, Palit S, Fillingim RB, Robinson ME. Multisystem resiliency as a predictor of physical and psychological functioning in older adults with chronic low back pain. *Front Psychol* 2019; 10: 1932.
18. Scheier MF, Carver CS, Bridges MW. Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): a reevaluation of the Life Orientation Test. *J Pers Soc Psychol* 1994; 67 (6): 1063–78.
19. Chu Y, Zhang Y, Wang S, Dai H. Resilience mediates the influence of hope, optimism, social support, and stress on anxiety severity among Chinese patients with cervical spondylosis. *Front Psychiatry* 2022; 13: 997541.
20. Souri H, Hasanirad T. Relationship between resilience, optimism and psychological well-being in students of medicine. *Proc Soc Behav Sci* 2011; 30: 1541–4.
21. Tusaie-Mumford K. Psychosocial resilience in rural adolescents: optimism. Perceived social support and gender differences. *Am J Psychiatry* 2001; 160: 496–503.
22. Lin N, Simeone RS, Ensel WM, Kuo W. Social support, stressful life events, and illness: a model and an empirical test. *J Health Soc Behav* 1979; 20 (2): 108–19.
23. Shaffer F, Ginsberg JP. An overview of heart rate variability metrics and norms. *Front Public Health* 2017; 5: 258.
24. Souza GG, Magalhães LN, Cruz TA, *et al.* Resting vagal control and resilience as predictors of cardiovascular allostasis in peacekeepers. *Stress* 2013; 16 (4): 377–83.
25. Souza GG, Mendonça-de-Souza AC, Barros EM, *et al.* Resilience and vagal tone predict cardiac recovery from acute social stress. *Stress* 2007; 10 (4): 368–74.
26. Agorastos A, Boel JA, Heppner PS, *et al.* Diminished vagal activity and blunted diurnal variation of heart rate dynamics in posttraumatic stress disorder. *Stress* 2013; 16 (3): 300–10.
27. Carnevali L, Koenig J, Sgoifo A, Ottaviani C. Autonomic and brain morphological predictors of stress resilience. *Front Neurosci* 2018; 12: 228.
28. Maunder RG, Nolan RP, Hunter JJ, Lancee WJ, Steinhart AH, Greenberg GR. Relationship between social support and autonomic function during a stress protocol in ulcerative colitis patients in remission. *Inflamm Bowel Dis* 2012; 18 (4): 737–42.
29. Hirten RP, Danieletto M, Tomalin L, *et al.* Factors associated with longitudinal psychological and physiological stress in health care workers during the COVID-19 pandemic: observational study using Apple Watch data. *J Med Internet Res* 2021; 23 (9): e31295.
30. Hirten RP, Danieletto M, Tomalin L, *et al.* Use of physiological data from a wearable device to identify SARS-CoV-2 infection and symptoms and predict COVID-19 diagnosis: observational study. *J Med Internet Res* 2021; 23 (2): e26107.
31. Hirten RP, Tomalin L, Danieletto M, *et al.* Evaluation of a machine learning approach utilizing wearable data for prediction of SARS-CoV-2 infection in healthcare workers. *JAMIA Open* 2022; 5 (2): ooac041.
32. Vaishnavi S, Connor K, Davidson JR. An abbreviated version of the Connor-Davidson Resilience Scale (CD-RISC), the CD-RISC2: psychometric properties and applications in psychopharmacological trials. *Psychiatry Res* 2007; 152 (2–3): 293–7.

33. Wollny AI, Jacobs I. Validity and reliability of the German versions of the CD-RISC-10 and CD-RISC-2. *Curr Psychol* 2023; 42 (5): 3437–48.
34. Waddimba AC, Baker BM, Pogue JR, *et al.* Psychometric validity and reliability of the 10- and 2-item Connor-Davidson resilience scales among a national sample of Americans responding to the Covid-19 pandemic: an item response theory analysis. *Qual Life Res* 2022; 31 (9): 2819–36.
35. Ahern NR, Kiehl EM, Sole ML, Byers J. A review of instruments measuring resilience. *Issues Compr Pediatr Nurs* 2006; 29 (2): 103–25.
36. Hahn EA, Devellis RF, Bode RK, *et al.*; PROMIS Cooperative Group. Measuring social health in the patient-reported outcomes measurement information system (PROMIS): item bank development and testing. *Qual Life Res* 2010; 19 (7): 1035–44.
37. Cella D, Riley W, Stone A, *et al.*; PROMIS Cooperative Group. The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005-2008. *J Clin Epidemiol* 2010; 63 (11): 1179–94.
38. Connor KM, Davidson JR. Development of a new resilience scale: the Connor-Davidson Resilience Scale (CD-RISC). *Depress Anxiety* 2003; 18 (2): 76–82.
39. Hays RD, Schalet BD, Spritzer KL, Cella D. Two-item PROMIS® global physical and mental health scales. *J Patient Rep Outcomes* 2017; 1 (1): 2.
40. Cano-García FJ, Sanduvete-Chaves S, Chacón-Moscoso S, *et al.* Factor structure of the Spanish version of the Life Orientation Test-Revised (LOT-R): Testing several models. *Int J Clin Health Psychol* 2015; 15 (2): 139–48.
41. Monitor your heart rate with Apple Watch. https://support.apple.com/en-us/HT204666. Accessed October 22, 2020.
42. Abt G, Bray J, Benson AC. The validity and inter-device variability of the Apple Watch™ for measuring maximal heart rate. *J Sports Sci* 2018; 36 (13): 1447–52.
43. Bai Y, Hibbing P, Mantis C, Welk GJ. Comparative evaluation of heart rate-based monitors: Apple Watch vs Fitbit Charge HR. *J Sports Sci* 2018; 36 (15): 1734–41.
44. Hernando D, Roca S, Sancho J, Alesanco Á, Bailón R. Validation of the Apple Watch for heart rate variability measurements during relax and mental stress in healthy subjects. *Sensors (Basel)* 2018; 18 (8): 2619.
45. Sachs M. cosinor: tools for estimating and predicting the cosinor model. R package version 1.1. 2014. https://CRAN.R-project.org/package=cosinor.
46. Hou R, Tomalin LE, Suárez-Fariñas M. cosinoRmixedeffects: an R package for mixed-effects cosinor models. *BMC Bioinformatics* 2021; 22 (1): 553.
47. Kursa MB, Rudnicki WR. Feature selection with the Boruta package. *J Stat Soft* 2010; 36 (11): 1–13.
48. Horn SR, Charney DS, Feder A. Understanding resilience: new approaches for preventing and treating PTSD. *Exp Neurol* 2016; 284 (Pt B): 119–32.
49. Windle G, Bennett KM, Noyes J. A methodological review of resilience measurement scales. *Health Qual Life Outcomes* 2011; 9: 8.
50. Riganello F, Prada V, Soddu A, di Perri C, Sannita WG. Circadian rhythms and measures of CNS/autonomic interaction. *Int J Environ Res Public Health* 2019; 16 (13): 2336.
51. Steinhardt M, Dolbier C. Evaluation of a resilience intervention to enhance coping strategies and protective factors and decrease symptomatology. *J Am Coll Health* 2008; 56 (4): 445–53.