Stress Indicators for Frontline COVID-19 Healthcare Workers

Team Members: Priscilla Siow, Ali Vakilimafakheri, Sarah Schlund, and Eric Wang

Mentor: Professor Franks

Sponsor: Eric J. Daza

UC SANTA BARBARA

evidation

Data Science Capstone

Introduction

❖ **Background**

➢ Using wearable, survey, and biomedical data from healthcare workers measured during the pandemic, we aim to analyze changes in sleep patterns, cognition, stress, etc. We hope that our findings will give insights into how wearable technology can help detect stress and recovery patterns.

❖ **Motivation**

➢ Our main motivation for this project was to identify which variables are most indicative of stress in our research population. Since the wellbeing of frontline healthcare workers is crucial during the worldwide pandemic, improving the observation and prediction of changes in healthcare workers through non-burdening wearable data aligns with our goal to improve the overall health of our communities.

Data

❖ **Data Collection**

➢ Wearable and survey data from healthcare workers measured during the pandemic, obtained through Synapse

- Wearable data was collected through the Oura Ring
- Survey data was collected in different time intervals through a smartphone app
- To maintain retention, a participant co-driven engagement (PCE) approach was implemented

❖ **Features**

➢ The features used primarily came from the Oura Ring, Daily Stress Measure, Post Traumatic Stress Disorder (PTSD), Adverse Childhood Events (ACE), and Demographics datasets.

Variable	Description	Range
hr_average	The average heart rate recorded during the sleep period.	> 0
restless	Restlessness of the sleep duration, measured in percentages.	(0, 100)
daily_shifts	The number of shifts worked in one day.	{0, 1, 2}
gender	Indicates the gender.	{0, 1}
latency	Detected latency from bedtime_start to the beginning of the first five minutes of persistent sleep.	> 0
total_ace	The sum of all ace scores	(0-10)

❖ **Missingness**

➢ Amount of observations that contained missing values were low; thus, we decided to drop them.

➢ Missing values could be linked to how stressed a participant was on a particular day. (high stress → no/bad survey).

❖ **Limitations**

➢ Wearable data are prone to inaccurate measurements, but we do not think the margin of error is too high.

➢ Survey responses are only as good as how the participant is feeling. Not too much concern for bias, as we believe the PCE approach mitigates the chance of half-hearted answers.

Exploratory Analysis

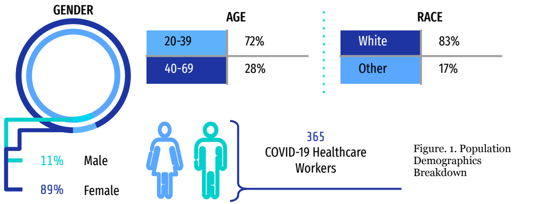

Figure. 1. Population Demographics Breakdown

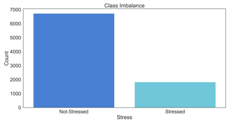

Figure. 2. Class Imbalance in binary response variable

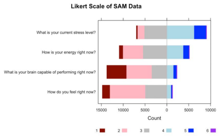

Figure. 3. Likert Scale of Survey Data

Methodology

❖ **Logistic Regression**

➢ Goal: Generalized model to predict the chance of stress given different values from Oura Ring Sleep Data and Survey Data

➢ Process: Deploy multiple logistic regression models to narrow down only significant indicators of stress

➢ Assumptions:

- Each observation is independent of one another
- No multicollinearity, that is, independent variables are not highly correlated
- Linearity of independent variables and log odds

❖ **Mixed Effects Model**

➢ Goal: Modeling to account for each individual having a different baseline for stress amongst other factors

➢ Process: Select a handful of significant variables from the results of the logistic regression model and center them to meet assumptions. Deploy a mixed effects model with a random intercept for each participant

➢ Assumptions:

- Linearity of independent variables and log odds
- No outliers
- No multicollinearity, that is, independent variables are not highly correlated

Final Results

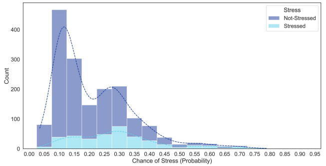

Figure. 4. Chance of being stressed compared to actually Stressed/Not-Stressed participants

Predictor	Coef.	P-Val
Intercept	-3.67	<0.05
Sleep Avg. HR	0.04	<0.05
Restless Sleep	1.86	<0.05
Daily Shifts	1.14	<0.05
Gender	-0.73	<0.05
Latency Score	-0.51	<0.05
ACE Score	0.04	<0.05

Figure. 6. Some significant predictors from Logit

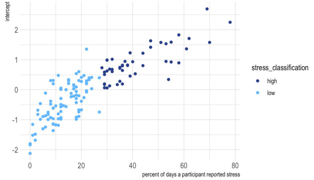

Figure. 5. Random intercepts for each participant from the mixed effects model

Predictor	Coef.	P-Val
Intercept	-1.41	<0.05
Sleep Avg. HR	0.11	<0.05
Restless Sleep	0.04	<0.05
Daily Shifts	1.24	<0.05
Gender	-0.99	<0.05
Latency Score	-0.04	<0.05
ACE Score	0.1	<0.05

Figure. 7. Same significant predictors in Mixed Effects

Conclusions

❖ We wanted to understand which variables are most indicative of stress in frontline healthcare workers

❖ Since this population is so critical in fighting the pandemic, it is important to better understand their stresses which can lead to burn out, unhappiness, etc.

❖ Wearable data can give us unique high personally insights which if analyzed correctly can help us understand our biological markers of stress

❖ Mixed effects models work well for models capturing stress within participants

References

Change in body temperature - anxiety symptom. AnxietyCentre.com. (2021, May 19). Retrieved May 25, 2022, Jin Hyun Cheong, P. D. (2022, January 7). *How to run linear mixed effects models in Python Jupyter Notebooks*. Medium. Retrieved May 25, 2022, from

Mixed Effects Logistic Regression. OARC Stats. (n.d.). Retrieved May 25, 2022,

Wilson, S. (2021, October 20). *Multiple imputation with random forests in Python*. Medium. Retrieved May 25, 2022, from