Predicting Maternal Stress During Pregnancy

By Harley Clifton, Sarah Schlund, Vasiliy Ostapenko, & MaShelle Gregory

I. Introduction, Scientific Goals, and Primary Questions of Interest

There is a growing body of research that demonstrates the negative effects of maternal stress during pregnancy on child developmental outcomes. This research has prompted interest in methods that can effectively decrease stress levels during pregnancy. In order to intervene during stressful moments, the ability to identify them occurring in real-time is necessary. Prior research has linked changes in heart rate, specifically changes in the interbeat interval (IBI), as a biological response to stress in humans and animals. Limited research currently exists utilizing heart rate and IBI to predict real-time stress in pregnant individuals. Data from a National Institutes of Health (NIH) funded study was used to identify instances of momentary stress from heart rate and contextual factors. Ideally, if the onset of episodes of maternal stress can be prematurely detected, intervention could potentially reduce the amount of stress the fetus encounters and consequently reduce their predisposition for negative behavior and developmental results in childhood.

There were three primary goals for this project that involved the use of contextual, biophysical, and demographic measures available in data. The first of these goals was to reliably classify the perceived stress level of an individual, which is a construct of self-reported responses calculated each time a participant completed an Ecological Momentary Assessment (EMA). The secondary goal was to quantify which times preceding stressful events are most relevant to discern how far in advance an oncoming stressful event can be detected. The final objective was to quantify between-subject differences in "predictive power" of biophysical and contextual measures (i.e. random effects).

II. Background and Related Work

Multiple studies, including one published in 2011 by Pathik Wadhwa, Sonja Entringer, Claudia Buss, and Micheal Lu, have found that early life stress can impact child neurodevelopment (Wadhwa et al.). Their findings also suggest that stress exposure can start as early as in the intrauterine environment. Fetal encounters with stress can lead to an increased potential for adverse developmental outcomes later in childhood; examples include neurodevelopmental and psychiatric disorders, cognitive impairment, anatomical and molecular neural changes, and behavioral problems.

If oncoming stressful events can be detected early enough, therapeutic intervention could potentially reduce the amount of cortisol exposure to the fetus. Consequently, this may put children at lower risk of negative developmental outcomes later in life. Thus, the aim of this study is to identify when stressful events are occurring through a low-burden method using a combination of biological, contextual, and demographic measures.

Ecological momentary assessments (EMA) are the current standard for collecting emotional, contextual, and behavioral data from participants in real time with low burden to the individuals (Shiffman et. al, 2008). A study conducted in 2021 utilized EMA data and found that people with depressive and anxiety disorders have comparatively higher affect instability – meaning they tend to experience rapid and intense mood swings that are difficult to control more often than others (Schoevers et. al). Further, EMA data can be utilized to discern mood and stress variability in a multitude of other contexts, in our case, in pregnant individuals.

In order to detect a stressful event through biophysical data, we used the root mean square of successive differences (RMSSD) of an individual's heart rate interbeat interval. As described by Minarini (2019), RMSSD of the interbeat interval is currently regarded as the leading indicator of when the parasympathetic system begins to affect the heart. It is widely known among psychologists that the parasympathetic system is activated in stressful situations; Therefore we would like to detect its activation.

Intuitively, the RMSSD calculations across consecutive IBI measurements for a single individual were highly correlated; Thus, successive RMSSD measurements resulted in multicollinearity in our model. To better manage this, a Distributed Lag model was fitted following the methodology and theory described in Joel Schwarz's paper titled *The Distributed Lag between Air Pollution and Daily Deaths* (2000).

III. Data Collection and Overview

This study was funded by a grant from the National Institutes of Health (NIH) to assess Ecological Momentary Assessment (EMA) data and cortisol levels in expectant mothers during three different time points throughout pregnancy. It was a Prospective Longitudinal Cohort Study, meaning a sample of subjects with similar characteristics (e.g. expectant mothers) was followed and observed over the course of the study. All subjects in our sample were receiving prenatal care at OB-GYN clinics in Orange County, California between 2011 and 2015. They were recruited through the University of California Irvine (UCI) Medical Center and monetarily compensated for their participation. 253 expectant mothers initially enrolled in the study, and there is available data on 246 of those mothers.

Once women were recruited, data on their demographic information was collected. Women then returned to their respective clinics for three periodic visits throughout the course of their pregnancy. At each of the three visits, the expectant mothers were sent home with an

ActiHeart Device and set up with a smartphone app to collect EMA over four consecutive days. Over the four day period for each visit, the women were prompted to complete an EMA diary approximately ten times per day during the hours they anticipated being awake. The electronic EMA diary recorded responses regarding their current stress levels, location, current activity, recent caffeine consumption, etc. Meanwhile, women were instructed to wear the ActiHeart device continuously for the duration of each four-day visit period and only to remove it when they showered. The ActiHeart device continuously monitored and recorded data on the women's heart rate (bpm), activity level (via an accelerometer), and interbeat interval (time between the peaks of two consecutive heartbeats).

There are two self-reported stress items in the dataset; the first of which is a single-item questionnaire response to the question "How stressed are you?" with responses between 0-4; where larger values represent higher levels of stress. The second – and most relevant – response of interest is the perceived stress score (PSS) which is a construct of self-reported responses (e.g. anxiety, control, confidence, irritation, anger, etc.). The PSS is also on a 0-4 scale and was transformed into a binary variable where values of 0, 1, and 2 were coded as no stress and values of 3 and 4 were indicative of a stressful event occurring.

There are a variety of covariates and biological measures available in the dataset as well. These include the age of the pregnant parent, maternal race, maternal ethnicity, visit number, day within visit, EMA number, date and time the measurements were recorded, socio-economic status index (between 0-5), pre-pregnancy BMI, healthy eating index, dietary inflammation index, and gestational age. Gestational Age, which is an estimate of time since conception, is suspected to be a potential confounder. This variable was used to create an additional "Trimester" variable based on the standard cutoffs for gestational age set by The American

College of Obstetricians and Gynecologists (2020). Physical Activity was measured via an ActiHeart device and calculated by taking the average activity level over a five-minute span prior to each EMA entry. There are also multiple indicator variables including parity (whether the mother had given birth prior to this pregnancy), whether the woman was at work at the time of the recording, whether it was a weekend or weekday, whether they were alone, having a social interaction, had consumed a caffeinated beverage within the last hour, and whether they were considered to be at Obstetric risk.

IV. Exploratory Data Analysis

Since gestational age was suspected to be a potential confounder for perceived stress score, the missing data patterns for this variable were explored. Luckily, less than 0.06% of the data on gestational age was missing, so there was abundant information to work with.

As anticipated, there is less information available on participants as the course of the study progressed. This may be due to pregnant individuals dropping out of the study, forgetting to wear their ActiHeart devices, or ignoring their EMA entry notifications. It is intriguing as to whether the lack of data may be due to increasing levels of stress as the study continued, making the subjects less likely to participate. To examine this further, we explored the proportion of responses from subjects who completed all entries to discern patterns of self-reported stress across the three visits.

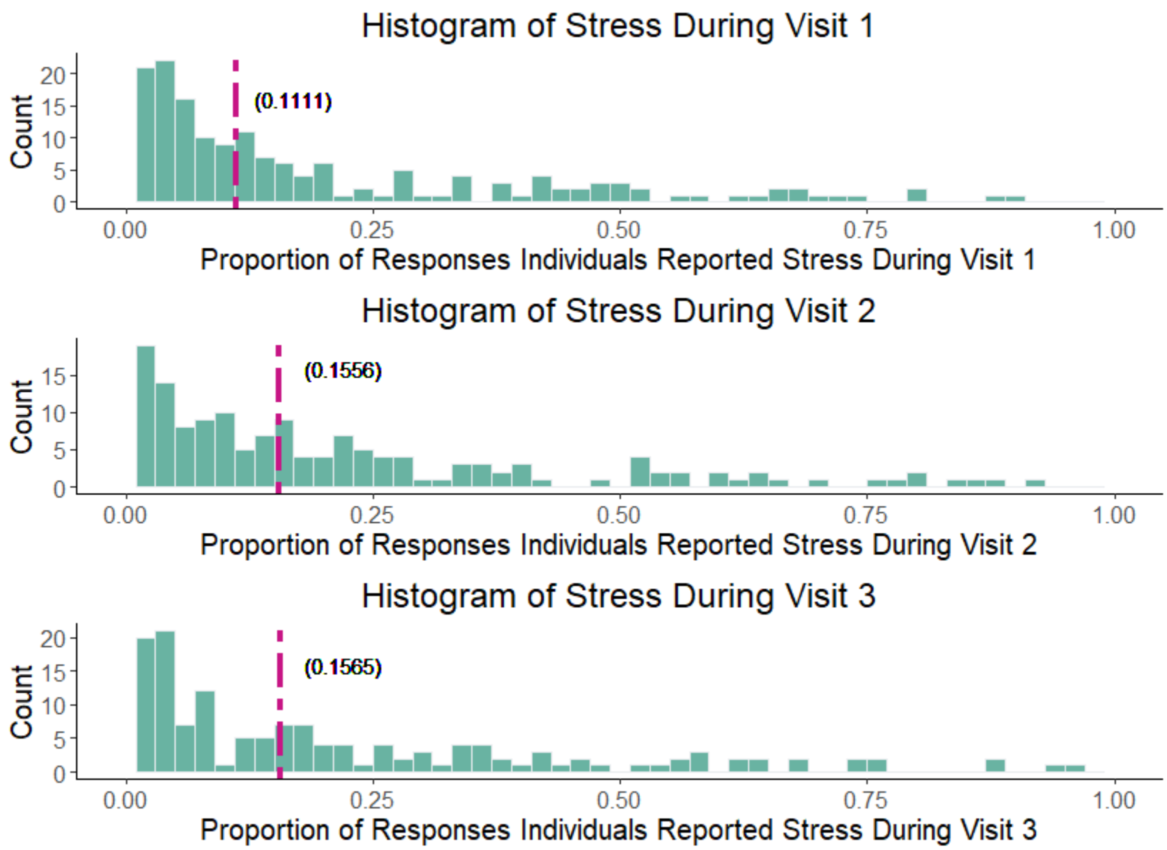

Figure 1: Histograms of Proportions of Self-Reported Stress Present in Pregnant Individuals across the three Visits.

The vertical dashed lines represent the median proportion of stress for participants within that visit. The exact median value is given in parentheses next to the lines.

We used the reported stress from the participants to find their individual proportions of the specific visits, the 1st to the 3rd, and calculated the median. From inspecting **Figure 1**, it is evident that more participants experienced lower stress and therefore, a lower stress proportion during their first visit, a higher proportion for the 2nd and 3rd visits, showing a shift and increase in stress during their pregnancies.

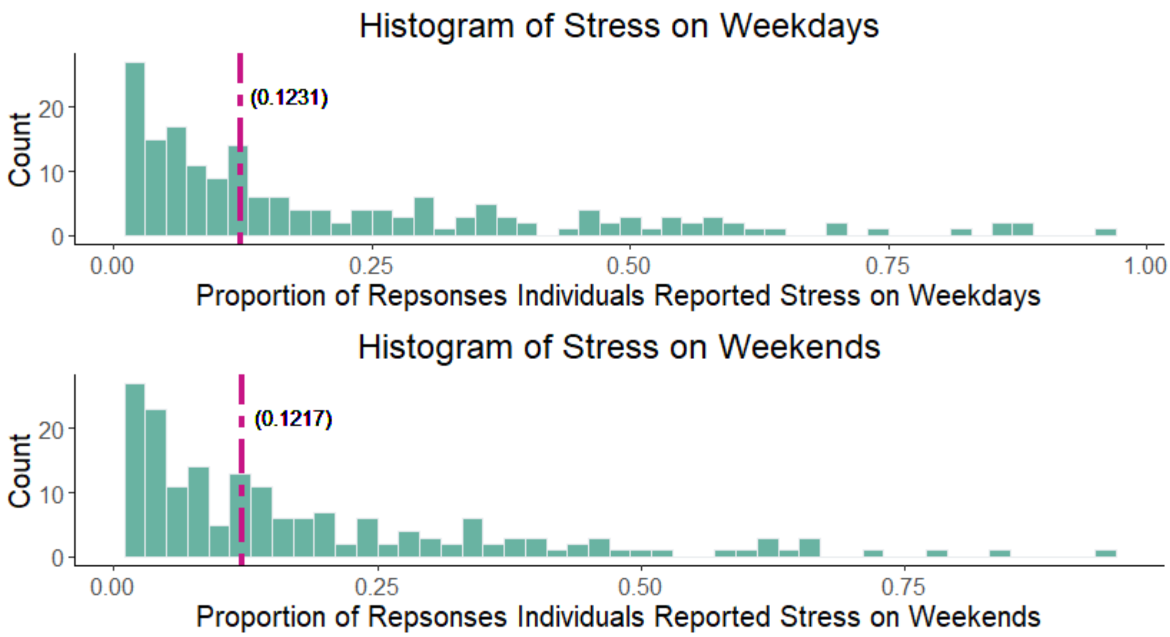

Figure 2: Histograms of Proportions of Self-Reported Stress Present in Pregnant Individuals on Weekdays compared to Weekends.

Figure 2 displays the relationship between visit number and proportion of responses that qualify for stress detection for each individual in the sample. It was expected that participants would experience stressful events more commonly on weekdays compared to on weekends. Interestingly, this relationship is present but is substantially weaker than initially anticipated.

V. Data Visualization

Since heart rate volatility, measured via RMSSD, is our main predictor of interest, we were interested in if and how the demographics of our participants varied by this volatility measure. In order to examine this, we classified participants as high and low volatility based on the mean value of their RMSSD values (recorded over the entire study period). We used the median of all participants' mean RMSSD values as the threshold to categorize whether an individual had "low" (below median) or "high" (above median) stress volatility over the course of their pregnancy.

Table 1: Maternal Demographics Stratified by Their RMSSD Volatility

	High (N=104)	Low (N=105)	Total (N=209)
Maternal Age			
Mean (SD)	26.9 (5.28)	28.2 (5.34)	27.6 (5.34)
Median [Min, Max]	27.0 [18.0, 40.0]	29.0 [18.0, 44.0]	28.0 [18.0, 44.0]
Race			
White	79 (76.0%)	81 (77.1%)	160 (76.6%)
Black	4 (3.8%)	4 (3.8%)	8 (3.8%)
Asian	9 (8.7%)	7 (6.7%)	16 (7.7%)
American Indian	7 (6.7%)	6 (5.7%)	13 (6.2%)
Pacific Islander	1 (1.0%)	0 (0%)	1 (0.5%)
Unknown	4 (3.8%)	7 (6.7%)	11 (5.3%)
Ethnicity			
Hispanic	42 (40.4%)	52 (49.5%)	94 (45.0%)
Non-Hispanic	62 (59.6%)	53 (50.5%)	115 (55.0%)
Socieconomic Status Index			
1	5 (4.8%)	5 (4.8%)	10 (4.8%)
2	29 (27.9%)	24 (22.9%)	53 (25.4%)
3	36 (34.6%)	46 (43.8%)	82 (39.2%)
4	30 (28.8%)	25 (23.8%)	55 (26.3%)
5	4 (3.8%)	5 (4.8%)	9 (4.3%)
Prior Children			
Yes	54 (51.9%)	60 (57.1%)	114 (54.5%)
No	50 (48.1%)	44 (41.9%)	94 (45.0%)
Missing	0 (0%)	1 (1.0%)	1 (0.5%)

As evident in **Table 1**, mean and median maternal age is quite similar for those who have high interbeat interval volatility compared to those who have low IBI volatility. The expected pattern was that older women would tend to have lower stress volatility on average, but this was not the case. Additionally, there is further quantitative confirmation that the original researchers recruited participants such that there was an intentional over-representation of Hispanic mothers compared to the national population during that time frame – which was approximately 17.6% according to the Pew Research Center (2017).

Table 2: Relevant Covariates and Their Relationship with Potential Effect Modifiers

	Weekend			Weekday		
	1st (N=2507)	2nd (N=5003)	3rd (N=3738)	1st (N=2873)	2nd (N=5452)	3rd (N=4167)
Reported Stress						
Stressed	389 (15.5%)	809 (16.2%)	590 (15.8%)	463 (16.1%)	994 (18.2%)	799 (19.2%)
Not Stressed	2118 (84.5%)	4194 (83.8%)	3148 (84.2%)	2410 (83.9%)	4458 (81.8%)	3368 (80.8%)
RMSSD						
Mean (SD)	62.1 (89.8)	56.5 (88.5)	39.5 (65.5)	59.1 (83.3)	55.7 (88.2)	43.7 (74.9)
Median [Min, Max]	27.8 [1.66, 654]	22.5 [1.31, 598]	16.3 [2.18, 666]	27.4 [2.24, 662]	22.3 [1.51, 772]	16.5 [1.79, 565]
Heart Rate (bpm)						
Mean (SD)	91.1 (16.9)	94.9 (17.2)	97.5 (14.7)	89.9 (15.7)	94.0 (16.3)	97.1 (14.7)
Median [Min, Max]	89.5 [51.0, 198]	93.0 [57.5, 192]	97.0 [57.5, 197]	88.5 [51.0, 192]	93.0 [56.0, 199]	96.0 [57.0, 197]
Alone?						
Yes	458 (18.3%)	884 (17.7%)	661 (17.7%)	739 (25.7%)	1447 (26.5%)	1200 (28.8%)
No	2049 (81.7%)	4119 (82.3%)	3077 (82.3%)	2134 (74.3%)	4005 (73.5%)	2967 (71.2%)
Activity						
Mean (SD)	8.02 (21.3)	9.53 (22.6)	7.87 (17.3)	8.05 (18.3)	8.42 (20.8)	7.58 (16.9)
Median [Min, Max]	0.500 [0, 617]	1.00 [0, 282]	1.00 [0, 192]	0.500 [0, 195]	1.00 [0, 609]	0.500 [0, 189]
Currently at Work?						
Yes	127 (5.1%)	275 (5.5%)	125 (3.3%)	619 (21.5%)	1214 (22.3%)	845 (20.3%)
No	2380 (94.9%)	4728 (94.5%)	3613 (96.7%)	2254 (78.5%)	4238 (77.7%)	3322 (79.7%)

As previously mentioned, both day type and trimester were suspected to be highly related to the perceived stress level of pregnant individuals. The relationships between these two predictors and other variables of interest were further explored in **Table 2**. Notably, the highest percentage of stress responses were reported during the weekdays on the third visit. This follows the expected trend of higher perceived stress on weekdays and as pregnancy perpetuates. It is also interesting that the mean beats per minute steadily increases throughout the trimesters. This is most likely due to an increase in blood volume during pregnancy which is necessary to nourish the growing baby. This subsequently causes the heart to pump more blood each minute and therefore increases overall heart rate.

Afterwards, a uniform manifold approximation and projection (UMAP) was constructed, which embedded the dataset using all features except the time lags to see if there was any clear separation between the two classes which we are asked to distinguish (binarized PSS of zero or one). UMAP is an algorithm which aligns every observation's multidimensional feature vectors onto the same plane in such a way as to group similar observations closer together by their coordinates (McInnes et al.). Thus, with clean data and quality features, the theory suggests that observations of two different classes should generally fall in two different clusters somewhat apart in the two-dimensional plane. The embedding was constructed with this in mind - using visuals to explore whether the features were sufficient for use in the classification task or if better data was needed.

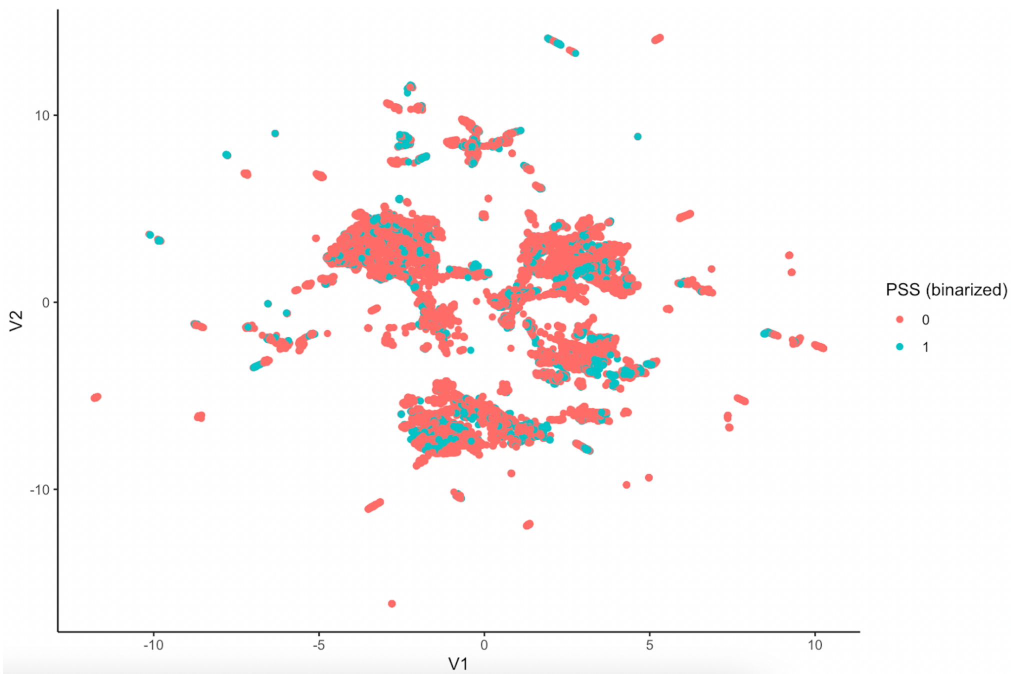

Figure 3: UMAP Feature Embedding Colored by Response Class.

Unfortunately, the UMAP embedding of the normalized features colored by response class was not satisfactory. There was not a clear separation between the two classes despite using all of our unique features and scaling the values. Therefore, these results prompt the exploration of whether functions of lagged IBI measurements would increase the ability to distinguish between the classes.

VI. Methods

RMSSD:

The Root Mean Square of Successive Differences (RMSSD) was calculated on the interbeat interval (IBI) i.e. the time between successive heartbeats in milliseconds. Each RMSSD measure was calculated for a 15-second interval using the following formula where x_i is the interbeat interval at time *i*.

$$\text{RMSSD} = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N-1}(x_{i+1}-x_i)^2}$$

Model Formulation:

To model the binary stress outcome variable, a random effect logistic regression model was used. In addition to contextual covariates, the model predictors included lags of RMSSD with a polynomial restriction on the beta coefficients. This polynomial restriction as described by Schwartz (2000) places a constraint on the model to: ensure parsimony, limit multicollinearity, and reduce variance inflation.

Since there were repeated measures within the same person over time, there was reason to believe that there was within-person correlation that needed to accounted for. To accomplish this, a generalized linear mixed effects model was used which included random intercepts for each person. This model assumes that the outcome variable follows a Bernoulli distribution, covariates are linear on a logit scale, and the random intercepts are independent following a normal distribution with mean zero and variance tau squared.

The final model formulation is as follows:

$$\begin{aligned} logit(\pi_{ij}) = \alpha_i + \beta_{activity} x_{activity_{ij}} + \beta_{work} x_{work_{ij}} + \\ \beta_{social} x_{social_{ij}} + \beta_{bpm} x_{bpm_{ij}} + \\ \beta_{weekend} x_{weekend_{ij}} + \beta_{ga} x_{ga_{ij}} + \\ \eta_0 w_{0ij} + \eta_1 w_{1ij} + \eta_2 w_{2ij} + \eta_3 w_{3ij} \end{aligned}$$

Where $\hat{\pi}$ is the probability that a pregnant individual is experiencing a perceived stress response, $x_{activity}$ is the activity level average over a five minute time span prior to an EMA entry measured via the ActiHeart device, x_{social} is an indicator variable that is 1 if the subject was having a social interaction at the time of the entry and 0 otherwise, x_{bpm} is the individual's heart rate in beats per minute, $x_{weekend}$ is an indicator variable that is 1 if the subject was recording an EMA entry on a weekend and 0 if it was recorded on a weekday, x_{ga} is the gestational age in days since conception and α is the random intercept for person i.

$$W_0 = Z_0 + Z_1 + \ldots + Z_q \qquad W_d = Z_1 + 2^d Z_2 + \ldots + q^d Z_q$$

As previously discussed, instead of using the lags of RMSSD, which are highly correlated, in the model directly, the coefficients of these variables were restricted and defined (d + 1) new variables, the W's you see above, which are weighted sums of the lagged RMSSD (Z), the desired number of lags (q) and degree of polynomial (d). These W's were then utilized in the model in place of the lagged RMSSD.

Train, Test Split Scheme:

Since both trimester and day type (weekend vs. weekday) affect stress, it was necessary that these were appropriately accounted for when splitting the data for model training and testing. As previously described, data were collected for a one visit approximately within each trimester, where participants recorded 4 days of data over two weekend days and two weekdays.

The data were split along these lines so that within each 4-day block, the model was trained on the first three days of data and tested on the fourth day. Therefore the split was stratified by trimester which ensured that model has both weekend and weekday data to train on. In rare cases when participants did not record data for all four days, the last day of their data was tested on and all previous days were designated as training data. There were even rarer cases when within a trimester, participants only recorded one day of data. Since it was impossible to implement the train/test split scheme in these very few cases, four to be exact, these participants were removed from the dataset – but only within the trimester for which they did not record more than one day of data.

VII. Model Assessment

Initial Covariate Selection:

The search space for the set of initial covariates to be included in the model was made of all biophysical and demographic features, as well as select contextual (EMA) features. In this step, derived time-lagged features of IBI (mean, RMSSD) were excluded. To select an optimal subset of features, a forward selection algorithm with AIC as the criterion was implemented. With "mean activity" as the only fixed effect in the logistic regression model with a random intercept, the algorithm tested whether the addition of any other feature as a fixed effect reduced the model's AIC. If a new minimum was found with two fixed effects, the algorithm proceeded to look for a lower AIC through adding yet another feature, and so on. The process ends if, either every feature in the search space is added as a fixed effect, or, AIC is not reduced when every combination of a certain length at some step is tried. Thus, this process resulted in a model with an optimum AIC – the model itself consisting of six features as fixed effects: “mean activity”, “work”, “social interaction”, “mean BPM”, “weekend”, and “gestational age”.

Final Model Selection:

To select a model with the optimal time-lagged features of IBI, namely choosing between lagged RMSSD or mean of IBI and the degree of polynomial used to restrict the model parameters, three-fold cross-validation was used with average accuracy over the folds as the criterion. Moreover, the folds were stratified by the response variable of interest, “pss_binary”. For the degree parameter, only two and three were tested, as greater values prevented the parameter estimation algorithm from converging. The results were very similar for the four models across the folds, which resulted in a model using time lags of RMSSD of IBI and a degree three polynomial. The average accuracy across the board was around 85%.

Model Testing:

The final model was trained on the training set and evaluated on the test set by considering classification accuracy and the ROC AUC score. We also visualized the ROC AUC curve as well as created a confusion matrix to analyze prediction error. The final model had approximately 86.6% accuracy in classifying the binarized PSS. Moreover, it scored 83.3% for ROC AUC. Finally, it is worth noting that the model is 5 times more prone to Type II error compared to Type I error.

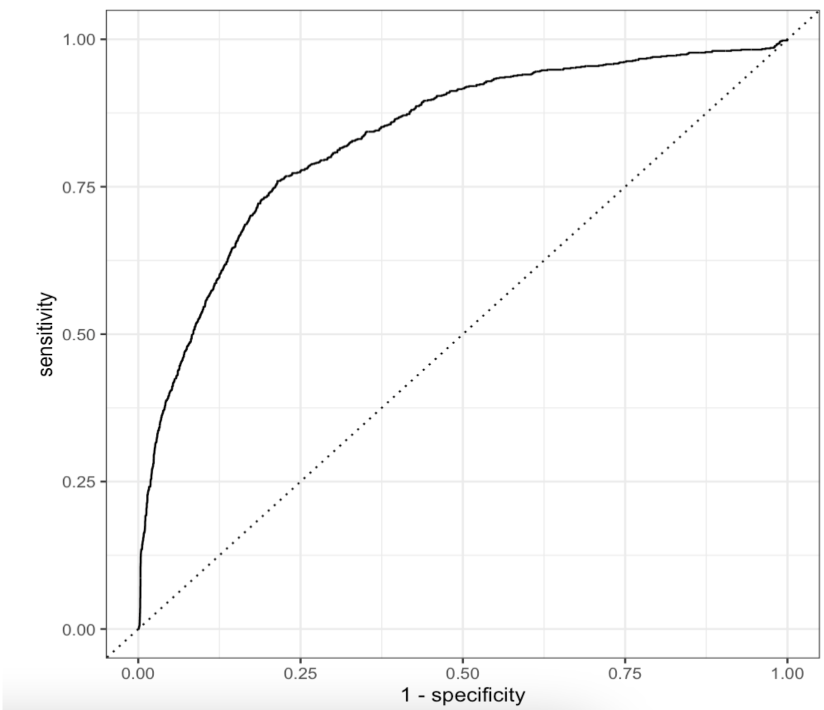

Figure 4: Receiver Operating Characteristic Curve (ROC Curve)

In this context, sensitivity, which is synonymous with the “true positive rate”, translates to how often a pregnant person who was experiencing stress was accurately predicted to be stressed. One minus the specificity, which is also called the “false positive rate”, represents when a pregnant individual is predicted to be experiencing stress but is actually not.

After creating the ROC plot (**Figure 4**), the area under the curve was calculated to quantify the model’s prediction ability. The area under the ROC curve is approximately 0.833, indicating that the model predicts perceived stress fairly well. Specifically, the probability that a randomly chosen pregnant individual who is truly experiencing stress has a higher estimated probability of perceived stress than a randomly chosen individual not actually experiencing stress is 0.833.

VIII. Interpretations & Results

Random Intercepts per Individual:

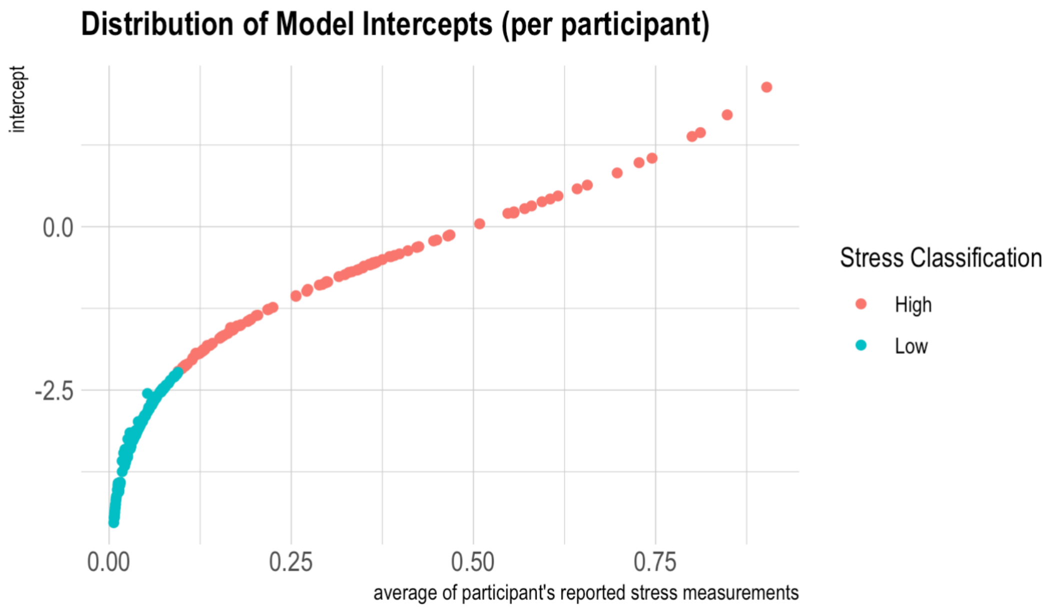

Figure 5: Distribution of Random Intercepts (per individual)
The red points represent participants who had a larger average of reported stress values than the overall average for all participants i.e. the average of all participant's averages. The blue points represent participants who had a smaller average than the overall average.

To validate that the random intercepts in the model accurately accounted for an individual's baseline stress level, a plot was created to display the random intercept for each individual vs. the average number of times they reported stress. If the model's intercepts are accurate, participants who reported stress more often should have a larger intercept than participants who rarely reported stress. As is evident in **Figure 5**, the random intercepts in the model are accurately capturing this.

In addition to validating that the random intercepts are accounting for an individual's baseline stress level, this result is important because it shows that the other variables in our model (limited to only lags of RMSSD in this plot i.e. no contextual covariates were included) are not able to account for this baseline by themselves. If they were, these intercepts would all be zero.

Throughout the modeling process, it became increasingly clear that using the individual's identifier ("id" feature) as a random intercept effect in the model was important in distinguishing between individual differences in stress baselines. For most people, the random intercept parameter takes on the highest value among all other parameters. Illustrated below is a distribution of random intercepts for all individuals on the probability scale, converted from log odds.

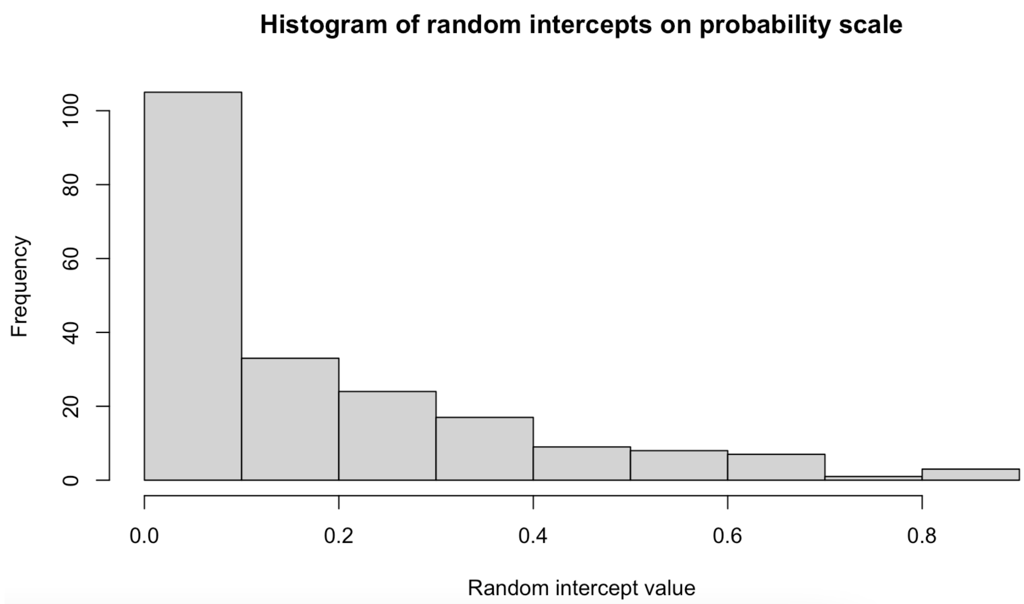

Figure 6: Histogram of Random Intercepts on the Probability Scale

The distribution is extremely skewed to the right. Through the intercept, the model gives most people the probability of an entry falling into the stressed class of about 10-15%. A few people are assigned an approximately 80% probability of a stressed entry simply due to their baseline.

Importance of Covariates:

Most of the covariates in the model turned out to be important - the respective parameter estimates were nonzero on the probability scale or nonunity on the log odds scale. Indeed, “work”, “social interaction”, “gestational age”, “mean BPM” as well as the second and third-degree terms in our decomposed time lags were significant.

```
                  Estimate Std. Error z value Pr(>|z|)    
w0                 0.05988    0.19321   0.310   0.7566    
w1                -1.04407    1.00026  -1.044   0.2966    
w2                 2.89930    1.72140   1.684   0.0921 .  
w3                -1.95528    0.90906  -2.151   0.0315 *  
mean_activity      0.03226    0.02362   1.366   0.1720    
work               0.14573    0.02428   6.002 1.94e-09 ***
social_interaction -0.11326   0.02515  -4.503 6.69e-06 ***
gestational_age    0.05192    0.02436   2.131   0.0331 *  
mean_bpm           0.05326    0.02900   1.837   0.0663 .  
weekend           -0.03770    0.02415  -1.561   0.1185    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Unexpected Results:

It was surprising to discover that the transformed time lag features would yield significant parameter estimates because they were not at all useful in increasing the predictive accuracy of the models during the cross-validation stage of model selection.

IX. Discussion

In this paper, we introduced a model to classify binarized Perceived Stress Score (PSS) from a questionnaire filled out by pregnant mothers at different points in their pregnancy. Our generalized linear mixed model, namely a logistic regression model with a random intercept, includes six covariates identified by step-AIC forward selection as well as a polynomial function of time lags of RMSSD of inter-beat interval (IBI). We demonstrate that out-of-sample model error stays consistent on validation data and show that around 85% accuracy on this classification task can be achieved, which is in line with some non-linear model approaches. Furthermore, looking at the estimated model parameters for the different coefficients, we can assess the relative importance of the covariates by using the estimate and the variance-covariance matrix of model coefficients to create confidence intervals. Looking at the distribution of the random intercepts for the different individuals also gives us insight into how the stress baseline varies between people. In summary, we demonstrate that it is possible to classify stress episodes based on biophysical, contextual, and demographic data, as well as that feature importance can be analyzed to quantify between-subject differences. Some further avenues for research include quantifying which time periods are most relevant to stress prediction, as well as determining how long prior to a stressful event can we detect it.

We would now also like to discuss some potential limitations of our model, along with interpretations of our results. The apparent cap on predictive accuracy of the different models we tried may be linked to the absence of some important features without which we cannot accurately perform the prediction task. Despite including lagged terms and varying the degree of polynomial restriction imposed, there were no major changes in accuracy. Moreover, using a non-linear model architecture, namely a random forest, there was also not an increase in predictive accuracy. These observations, along with the original UMAP embedding demonstrating a lack of separation between the two classes, leads us to believe our model is failing to account for some factors. To aid further research in this area, it would be helpful to gather more biophysical and contextual data per individual over time. Concretely, it would be interesting to see if cortisol measurements as well as sleep duration and quality data could be helpful in predicting an individual's stress. The former is a useful feature because it is the main biochemical driver of stress in the body. Changes in levels of the hormone cortisol are associated with changes in levels of stress and anxiety in the human subject (Kev et al.). However, the widespread method of measuring cortisol through saliva makes it difficult to attach a time stamp on the measurement, while cortisol during sleep is not measured at all. The latter feature's strong association with stress also makes it a valuable piece of information to know. Lack of quality sleep has a strong impact on the mental state of the subject, with stress, anxiety, and memory impacted (Kim et al.). However, sleep data could be inaccurate if collected through subject self-assessment, as opposed to electronic means.

Additionally, as with all studies, there are limitations when it comes to the generalizability of results due to the sample's composition. First, we can only generalize these results to people who can get pregnant and receive health care. There is no representation of individuals with any health irregularities like cognitive disorders, heart defects or conditions such as Arrhythmia, or those taking any known medications. Individuals that fall into these categories were excluded from the sample to begin with. Secondly, the sample only consists of pregnant individuals receiving care at clinics in Orange County, California, and therefore the results are further limited in their generalizability. Third, since this was a longitudinal cohort study, it is implied that it is observational in nature, and thus no causal inferences between variables can be drawn.

Finally, it would have been preferable to have data on other variables in addition to those available to us. These include information on aspects such as body temperature, since thermal stress can affect heart rate. It is also known that hydration level can impact heart rate; Lower hydration levels mean the heart has to beat more to circulate the necessary oxygen around the body. Recent or impending illness can result in an elevated heart rate as well. In future studies, the collection of information regarding these variables would be suggested.

References

The American College of Obstetricians and Gynecologists (2020). “How Your Fetus Grows During Pregnancy: How Is the Length of My Pregnancy Measured?” *ACOG*. www.acog.org/womens-health/faqs/how-your-fetus-grows-during-pregnancy#:%7E:text=Here’s%20how%20the%20trimesters%20are,of%20rapid%20growth%20and%20development.

Flores, A., López, G., & Radford, J. (2017). 2015, Hispanic population in the United States Statistical Portrait. Pew Research Center's Hispanic Trends Project. https://www.pewresearch.org/hispanic/2017/09/18/2015-statistical-information-on-hispanics-in-united-states-trend-data/.

Vedhara, K., et al. (2003). An Investigation into the Relationship Between Salivary Cortisol, Stress, Anxiety and Depression. *Biological Psychology*; 62(2): 89-96, ISSN 0301-0511, https://doi.org/10.1016/S0301-0511(02)00128-X.

Kim, E,, & Dimsdale, J. (2007). The Effect of Psychosocial Stress on Sleep: A Review of Polysomnographic Evidence. *Behavioral Sleep Medicine*; 5(4):256-78. https://doi.org/10.1080/15402000701557383. PMID: 17937582; PMCID: PMC4266573.

Minarini, G. (2020). "Root Mean Square of the Successive Differences as Marker of the Parasympathetic System and Difference in the Outcome after ANS Stimulation". Autonomic Nervous System Monitoring - Heart Rate Variability, edited by Theodoros Aslanidis, IntechOpen. https://doi.org/10.5772/intechopen.89827.

McInnes, L., Healy, J., & Melville, J. (2018). “UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction”. https://doi.org/10.48550/arXiv.1802.03426.

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Rich, B. (2021). The table1 Package: Tables of Descriptive Statistics in HTML. Version 1.4.2, https://CRAN.R-project.org/package=table1.

Schoevers, R. A., van Borkulo, C. D., Lamers, F., Servaas, M. N., Bastiaansen, J. A., Beekman, A., van Hemert, A. M., Smit, J. H., Penninx, B., & Riese, H. (2021). "Affect fluctuations examined with ecological momentary assessment in patients with current or remitted depression and anxiety disorders". *Psychological Medicine*, 51(11), 1906–1915. https://doi.org/10.1017/S0033291720000689.

Schwartz, J. (2000). "The Distributed Lag between Air Pollution and Daily Deaths". *Epidemiology (Cambridge, Mass.)*; 11(3):320-326. https://doi.org/10.1097/00001648-200005000-00016.

Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological Momentary Assessment. *Annual Review of Clinical Psychology*, 4, 1–32. https://doi.org/10.1146/annurev.clinpsy.3.022806.091415

Wadhwa, P., Entringer, S., Buss, C., & Lu, M. (2011). "The Contribution of Maternal Stress to Preterm Birth: Issues and Considerations". *Clinics in Pernatology*: 38(3):351-384. https://doi.org/10.1016/j.clp.2011.06.007.

Wickham, H. (2016). The ggplot2 Package: Elegant Graphics for Data Analysis. Springer-Verlag New York. https://ggplot2.tidyverse.org.