

RESEARCH ARTICLE

Machine learning-based predictive modeling of resilience to stressors in pregnant women during COVID-19: A prospective cohort study

Emily S. Nichols^{1,2}*, Harini S. Pathak³, Roberta Bgeginski⁴, Michelle F. Mottola^{4,5}, Isabelle Giroux⁶, Ryan J. Van Lieshout⁷, Yalda Mohsenzadeh^{2,3}, Emma G. Duerden^{1,2,8}

1 Applied Psychology, Faculty of Education, Western University, London, Ontario, Canada, **2** The Brain and Mind Institute, The University of Western Ontario, London, Ontario, Canada, **3** Department of Computer Science, The University of Western Ontario, London, Ontario, Canada, **4** R. Samuel McLaughlin Foundation—Exercise and Pregnancy Laboratory, School of Kinesiology, Faculty of Health Sciences, Children's Health Research Institute, Western University, London, Ontario, Canada, **5** Department of Anatomy and Cell Biology, Schulich School of Medicine and Dentistry, Western University, London, Ontario, Canada, **6** School of Nutrition Sciences, Faculty of Health Sciences, University of Ottawa, Ottawa, Ontario, Canada, **7** Department of Psychiatry and Behavioural Neurosciences, McMaster University, Hamilton, Ontario, Canada, **8** Psychiatry, Schulich School of Medicine and Dentistry, University of Western Ontario, London, Ontario, Canada

* These authors contributed equally to this work.

* enicho4@uwo.ca

OPEN ACCESS

Citation: Nichols ES, Pathak HS, Bgeginski R, Mottola MF, Giroux I, Van Lieshout RJ, et al. (2022) Machine learning-based predictive modeling of resilience to stressors in pregnant women during COVID-19: A prospective cohort study. *PLoS ONE* 17(8): e0272862. <https://doi.org/10.1371/journal.pone.0272862>

Editor: Mohd Nadhir Ab Wahab, Universiti Sains Malaysia, MALAYSIA

Received: September 15, 2021

Accepted: July 28, 2022

Published: August 11, 2022

Copyright: © 2022 Nichols et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The minimal data set can be found at <https://osf.io/5ksr6>.

Funding: This research was supported by the Canada First Research Excellence Fund (CFREF) awarded to BrainsCAN at Western University, the Whaley and Harding Fellowship-Children's Health Research Institute, and the Western Fellowship Program. The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation,

Abstract

During the COVID-19 pandemic, pregnant women have been at high risk for psychological distress. Lifestyle factors may be modifiable elements to help reduce and promote resilience to prenatal stress. We used Machine-Learning (ML) algorithms applied to questionnaire data obtained from an international cohort of 804 pregnant women to determine whether physical activity and diet were resilience factors against prenatal stress, and whether stress levels were in turn predictive of sleep classes. A support vector machine accurately classified perceived stress levels in pregnant women based on physical activity behaviours and dietary behaviours. In turn, we classified hours of sleep based on perceived stress levels. This research adds to a developing consensus concerning physical activity and diet, and the association with prenatal stress and sleep in pregnant women. Predictive modeling using ML approaches may be used as a screening tool and to promote positive health behaviours for pregnant women.

Introduction

During the COVID-19 pandemic, pregnant women have been exposed to significant stress, including infection-related fears, social isolation, loss of income, lack of access to medical care [1]. Previous studies that have examined prenatal stress during natural disasters or epidemics suggest that prenatal stress is a significant contributor to adverse maternal and neonatal health outcomes [2]. Recent evidence suggests that prenatal distress has become endemic in pregnant women during COVID-19 [3]. Given the current global pandemic, millions of pregnant

review, or approval of the manuscript; and decision to submit the manuscript for publication. There was no additional external funding received for this study.

Competing interests: The authors have declared that no competing interests exist.

women are at risk for adverse health outcomes with few guidelines currently in place to help women manage stress and mitigate its harmful effects.

Prenatal stress is a risk factor for pregnancy complications including low birth weight, pre-term labour, hypertension, and delayed neonatal development [4, 5]. Several lifestyle factors have emerged as protective factors to promote resiliency against prenatal stress in pregnant women including physical activity and diet [6]. However, due to ongoing physical distancing regulations, pregnant women may be more likely to work from home and spend leisure time online or consuming other forms of media [7], and pregnant women who reported decreases in physical activity during the pandemic exhibited higher levels of depression [8]. Another study showed that pregnant women who performed the international recommendation of 150 min of physical activity per week [9] exhibited significantly lower anxiety and depression scores compared to women who performed less [3], suggesting that physical activity may be a key resilience factor to promote maternal well-being. However, studies examining physical activity in pregnant women during the pandemic have used limited measures and focused on change in levels [3, 8]. Because the form of physical activity that many women are engaging in has likely changed due to closures and restrictions on gatherings, more research is required to determine what forms of physical activity pregnant women are engaging in, and how this relates to stress during the pandemic.

Diet is also a modifiable lifestyle factor that predicts maternal and neonatal health [10]. Psychological distress is associated with poor nutrition in healthy adults [11]; yet, stress has also been identified as a potential risk factor for emotional eating behaviours in pregnant women during the pandemic [12]. Further, a recent report indicated that households are purchasing fewer fresh items such as meat, fruits, and vegetables [13]. Limited food access and increased unhealthy eating behaviours may have consequences for perceived stress, and a better understanding of the dietary behaviours that pregnant women are currently engaging in, and their association with stress, is needed.

These changes in behaviour are likely to have downstream effects beyond the impact on stress levels. In pregnancy, sleep loss is endemic, with up to 60% of women experiencing insomnia [14], and sleep disruptions in pregnant women have only increased during the pandemic [15]. Psychological distress related to COVID-19 has been a major cause of these disruptions [15], and associations amongst sleep characteristics (i.e., sleep time, hours of sleep) and maternal distress during the early phases of the pandemic have been reported [16]. Poor sleep is also strongly predicted by stress during pregnancy, which is associated with adverse maternal outcomes [17], indicating that increased levels of pandemic-related stress may be associated with worse sleep quality in pregnant women.

Machine learning (ML) is a useful tool in examining complex relationships between many predictors. Work in pregnancy has employed ML techniques for decades; for example, a recent article reviewed 26 studies that used ML to study the optimal mode of childbirth [18]. Other work has examined risk factors in pregnancy, including for sepsis [19], ectopic pregnancy [20], and gestational diabetes (see [21] for a review).

Based on previous work, in the present study we sought to determine whether physical activity and diet are associated with resilience against prenatal stress, and whether stress is in turn associated with sleep, during the COVID-19 pandemic. We applied a series of ML algorithms to a heterogeneous dataset from an international cohort of pregnant women who completed an online survey study during May to September 2020. Our central hypothesis was that both physical activity and diet would be predictive of prenatal stress, and that in turn, stress would be predictive of sleep levels, in pregnant women, to identify modifiable lifestyle factors that are associated with reduced stress and improved sleep in pregnant women during the pandemic.

Materials and methods

Participants

A total of 1049 pregnant individuals enrolled in the study. Complete data were available for 804 after data cleaning procedures (see Section 2.4 for details). Full demographic information for the final sample is provided in Table 1. Recruitment and data collection were performed online via Prolific, Amazon's Mechanical Turk, and through advertising on social media websites, and was targeted worldwide, within a single time window. Participants were mainly recruited from the United States (31.7%), the United Kingdom (20.6%), Canada (20.5%), and India (10.7%). Women ranged in age from 18–53 years ($M = 30.67$, $SD = 5.07$), and gestational age ranged from 2 to 42 weeks ($M = 24.5$, $SD = 9.53$). This study was approved by the Western University Research Ethics Board (#115810) and all participants provided informed implied consent by clicking a button to complete the survey prior to participating.

Materials

Self-report measures used to assess maternal demographic characteristics, stress, sleep, physical activity levels, and diet were collected and are summarized below.

Table 1. Demographic information of pregnant women.

Measure	
N	804
Age, years, Median [IQR]	30 [27–33]
Gestational age, weeks, Median [IQR]	25 [17–32]
Country, % (n)	
USA	31.7% (255)
UK	20.6% (166)
Canada	20.5% (165)
India	10.7% (86)
Other	16.4% (132)
Pandemic month, % (n)	
May	27.9% (224)
June	5.0% (40)
July	61.4% (494)
August	5.7% (46)
Hours of sleep/night, Median [IQR]	7 [6–8]
Ethnicity, % (n)	
Caucasian	64.9% (522)
Asian	18.4% (148)
Black	5.6% (45)
Hispanic	3.3% (27)
Other	7.7% (62)
Education level††, Median [IQR]	18 [18–21]
<i>Perceived Stress Scale</i>	
Total score, Median [IQR]	20 [14–24]

Clinical and demographic factors IQR, interquartile range; PSS, perceived stress scale;

† Other countries included South Africa, the Philippines, Brazil, Croatia, Italy, Fiji;

†† Education level, 18: College education, 21: Graduate degree

<https://doi.org/10.1371/journal.pone.0272862.t001>

1. Demographic Characteristics

Participants reported on gender, age, location (city, country), gestational age, due date, pregnancy complications, ethnicity, education level (years), occupation, days spent in quarantine/self-isolation, ability to see their healthcare provider, diagnosed maternal health problems (diabetes, cardiovascular health, depression), weight (pre-pregnancy, pregnancy), and height.

2. Antenatal Stress

Maternal stress was assessed using the Perceived Stress Scale (PSS), a 10-item questionnaire [22] that has been validated for use in pregnant women [23]. The PSS measures the degree to which situations are appraised as stressful, and relates to experiences over the past month. The participant ranks each question on a scale of 0–4, and overall scores are obtained by reversing responses to four positively stated items and then summing across all scale items.

3. Sleep

Maternal sleep patterns were assessed using the Pittsburgh Sleep Quality Index (PSQI), a validated 10-item questionnaire designed to measure sleep quality [24] that has been validated in pregnant women [25]. Items include hours of sleep, medication use, waking during the night, and trouble falling asleep. Scores for seven domains are calculated from the participant responses, and a global PSQI is assigned based on the sum of the seven component scores. As a previous relationship has been established between hours of sleep and maternal distress [16], the present study used the total hours of sleep as our measure of sleep quality.

4. Physical Activity

Physical activity levels were assessed using the Pregnancy Physical Activity Questionnaire, a validated 36-item questionnaire [26]. Items include time spent doing housework (e.g., preparing meals), time spent in sedentary activities (e.g., watching TV), methods of transportation, and exercise. All response items are on frequency-based six-point Likert scales (e.g., ranging from “None” to “3 or more hours per week”). The nine subscales of the PPAQ were then scored according to questionnaire responses for the following types of activity: sedentary; light; moderate; vigorous; household/caregiving; occupational; sports/exercise; transportation; inactivity.

5. Diet

Maternal food and beverage intake were assessed using PrimeScreen, a validated 21-item dietary screening questionnaire [27]. This questionnaire is a semiquantitative food frequency questionnaire allowing to assess diet quality. Items inquire about intake of various food groups including fruits and vegetables, dairy products, whole grains, fish, red meat, and major contributors to saturated and trans fats. The frequency of intake of foods is categorized by participants as: less than once per week (score = 1), once per week (score = 2), 2–3 times per week (score = 3), nearly daily or daily (score = 4), or twice or more per day (score = 5).

Procedure

Data were collected online via Qualtrics, an online survey tool (Qualtrics, Provo, UT). After providing informed consent, participants completed all questionnaires. Completion time of the survey portion of the study was approximately 40 minutes.

Data cleaning and reduction

Only data from the participants who completed all relevant questionnaire items were included in the analysis, and rows with missing values were removed, eliminating 69 participants. Rows

with responses of ‘Prefer not to answer’ were also removed, eliminating 176 participants. After cleaning, a total of 804 datasets remained for analysis.

All code for feature selection and classification is available here: <https://osf.io/6n2wk>. Code for generating figures is available here: <https://osf.io/z2raq>. As there were a large number of questionnaire items, feature selection was performed in order to reduce complexity and improve the accuracy of the model [28]. Feature selection is a way of reducing the input features and strengthening the prediction results by including only relevant and meaningful features to the model [29]. Mutual information feature selection was used here to determine the optimal number of questionnaire items to maximize classification accuracy [28], and target classes for each model were determined in a hypothesis-driven manner, described below. All questionnaire responses were converted to numeric scales in order to perform feature selection.

For the first model, in which we examined whether physical activity features could accurately predict stress classes, nine physical activity-based questionnaire items were extracted for feature selection. Overall stress scores were divided into “low” and “high” target classes using a median split [30]. Each participant was then assigned to a target class of 0 (low stress levels) or 1 (high stress levels) based on their cluster membership. As can be seen in Fig 1A, classification accuracy was highest when all nine features of the PPAQ were included, and so no features were removed from further analyses [28].

In our second model, we examined whether dietary behaviours could accurately predict stress classes. We again divided overall stress scores into “low” and “high” target classes using a median split. Two questions pertaining to vitamin use were not used in the present study, resulting in 19 features being submitted to feature selection. As can be seen in Fig 1B, classification accuracy was highest when all 19 items of PrimeScreen were included.

For the third model in which we examined whether perceptions of stress could accurately classify sleep classes, feature selection again indicated that all 10 items of the PSS improved classification accuracy and were thus included in further analysis (Fig 1C). Participants’ sleep values were grouped into one of three classes based on the amount of sleep per night as follows: below 7 hours, 7–9 hours, and above 9 hours. These classes were based on prior work suggesting that pregnant women require between 7–9 hours of sleep per night [31].

Support vector machine classification

Classification was performed using a multi-class Support Vector Machine (SVM) algorithm, a supervised learning model that finds a hyperplane in an N-dimensional space (with N being the number of features) in order to classify the observations. SVM parameters (i.e., kernel, C, and gamma) were first optimized using a grid search, which performs an exhaustive search over specified parameter values. Classifiers were then trained using the optimized set of parameters (Table 2) to identify classes from the set of features assigned for each prediction (Table 3). To overcome imbalance in class membership, we randomly sampled data points in each class based on the minority class. This ensures equal number of samples in each class. Data were then split into 80% train and 20% test sets, and trained and tested multi-class SVM on the data. Each iteration involved running the classifier on a random sampling of participants for the 80:20 train/test split. This procedure was repeated for 6000 iterations and the averaged prediction performance of the classifier over these iterations was reported. Averaging a large number of iterations ensured that no random sampling biased the classifier towards higher or lower performance. To perform above chance, a mean classification accuracy above 50% must be achieved for models with two target classes, and above 33% for models with three target classes. Baseline models were also generated to which we compared the trained models’

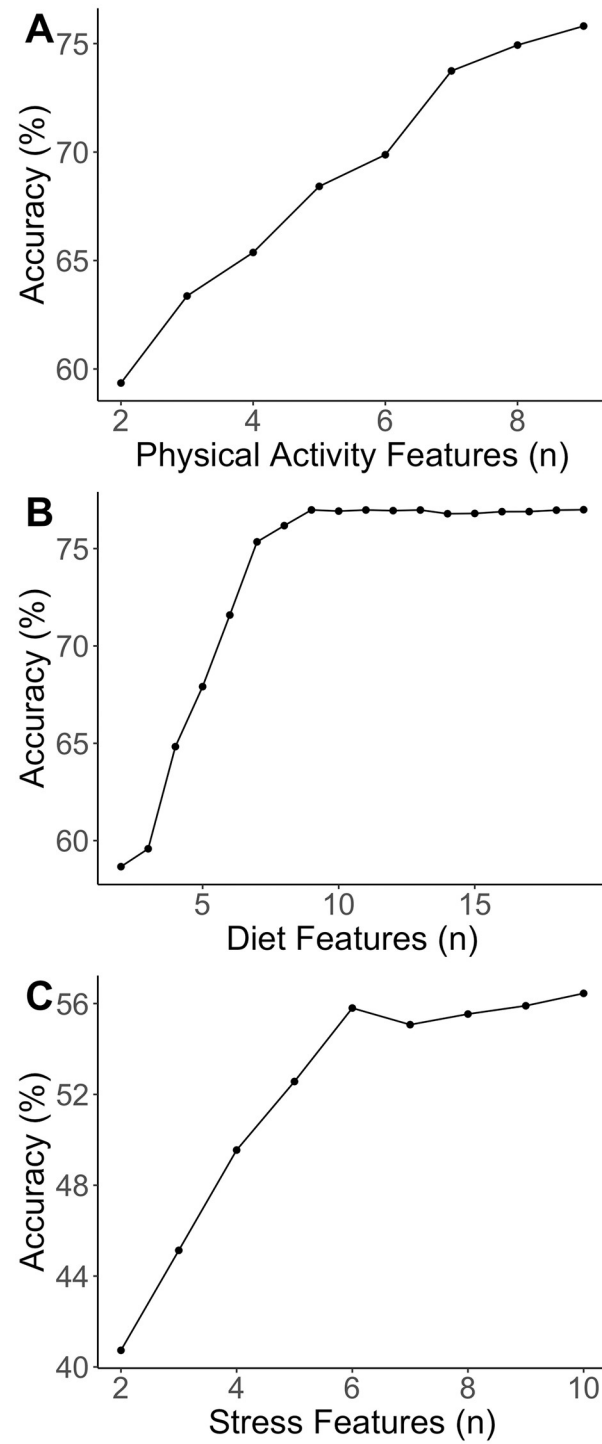


Fig 1. Accuracy values (%) during mutual information feature selection. A) physical activity predicting stress classes, B) diet predicting stress classes, C) stress predicting sleep classes.

<https://doi.org/10.1371/journal.pone.0272862.g001>

Table 2. Default and optimized parameters for each predictive model.

Parameter	Default	Optimized		
		PPAQ → Stress	Diet → Stress	Stress → Sleep
Kernel	Radial Basis Function	Radial Basis Function	Radial Basis Function	Radial Basis Function
C	1	10	1	1
Gamma	Scale	1	1	1

<https://doi.org/10.1371/journal.pone.0272862.t002>

performance. The classification procedure was repeated 6000 times, however class labels were randomly shuffled for each iteration. Baseline classification levels are reported in Table 4. To determine whether classification accuracies were significantly above chance and baseline levels, one-sample, one-sided *t*-tests were computed with μ set to chance level.

Feature ranking and target class characteristics

To determine which participant characteristics may explain target class membership, we examined the top three features that provided the most information gain. Proportions of total responses were calculated and plotted for each of the top three features, for each model, and Pearson's chi-square tests were used to test the difference in distributions between classes.

Results

Support vector machine classification

An SVM was used to determine whether, within pregnant women, physical activity could accurately predict perceived stress class. To test this, we averaged the SVM decoding accuracy across 6000 iterations; accuracy results are shown in Table 4 and precision and recall are shown in Table 5. Classification performance is shown in Fig 2. The SVM performed significantly above chance and baseline level ($t(5999) = 541.21, p < .001$), with a mean classification accuracy of 75.8% (SD = 3.7). In order to determine in turn whether dietary behaviours could accurately predict perceived stress class in pregnant women, a second SVM was run. The SVM again performed significantly above chance and baseline ($t(5999) = 659.88, p < .001$), with a mean classification accuracy of 77.0% (SD = 3.2). Finally, in order to determine in turn whether perceived stress could predict sleep class, a third SVM was run. The SVM again performed significantly above chance and baseline ($t(5999) = 207.48, p < .001$), with a mean classification accuracy of 56.4% (SD = 8.8).

Feature ranking and target class characteristics

Feature ranking for each model is shown in Fig 3, and Table 6 shows the mapping of feature index to question. The top three features for each classifier were examined in order to describe the characteristics that contributed most to classification. In the case of physical activity, the three features that contributed the most information gain to stress classification were Sport/

Table 3. Number of features and target classes for each predictive model.

Cross predictor model	No. features	No. target classes
Physical activity → Stress	8	2
Diet → Stress	19	2
Stress → Sleep	10	3

<https://doi.org/10.1371/journal.pone.0272862.t003>

Table 4. SVM mean classification accuracies and chance levels for three cross predictor models.

Cross predictor model	Accuracy	Chance level	Baseline Model
	M (SD)		
Physical activity → Stress	75.8% (3.7)	50%	50%
Diet → Stress	77.0% (3.2)	50%	50%
Stress → Sleep	56.4% (8.8)	33%	33%

<https://doi.org/10.1371/journal.pone.0272862.t004>

Exercise, Housework, and Sedentary Activity. Average responses for all features are shown in Fig 4A.

The top three dietary features that contributed the most information gain to stress classification were “How often do you add salt to food at the table?”; “How often do you eat deep fried foods (deep fried chicken, fish or seafood; French fries, onion rings)?”; and “How often do you eat beef, pork or lamb as main dish?”. Average responses for each feature are shown in Fig 4B.

Finally, the top three stress features that contributed the most information gain to sleep classification were “In the last month, how often have you felt that things were going your way?”; “In the last month, how often have you felt that you were on top of things?”; and “In the last month, how often have you been angered because of things that were outside of your control?”. Average responses for each feature are shown in Fig 4C.

Discussion

In the current study, we sought to determine whether physical activity and diet were resilience factors against prenatal stress during the pandemic, and whether perceived stress would in turn be predictive of hours of sleep, using an SVM trained on data from 804 pregnant women. Physical activity features were able to predict stress class (low/high) in pregnant women well above chance level; similarly, dietary behaviours were also able to predict stress class. When examining whether perceived stress could in turn predict sleep class in pregnant women, the classifier again performed significantly above chance level. These results suggest that physical activity and diet may be modifiable factors that could help to alleviate prenatal stress and promote maternal sleep. What choices, then, could pregnant women make in order to reduce stress levels and improve sleep?

To address this question, we examined which features provided the most explanatory power to the classifier. The top three physical activity features perhaps unsurprisingly included measures of intentional exercise such as walking, jogging, or swimming; housework; and sedentary activity. As shown in Fig 4A, while women with high stress levels reported marginally

Table 5. Evaluation of each model in terms of precision and recall.

Model	Class	Precision	Recall
PPAQ -> Stress	Low	0.77	0.75
	High	0.74	0.78
Diet -> Stress	Low	0.75	0.90
	High	0.93	0.64
Stress -> Sleep	< 7 hours	0.95	0.56
	7–9 hours	0.47	0.70
	> 10 hours	0.55	0.42

<https://doi.org/10.1371/journal.pone.0272862.t005>

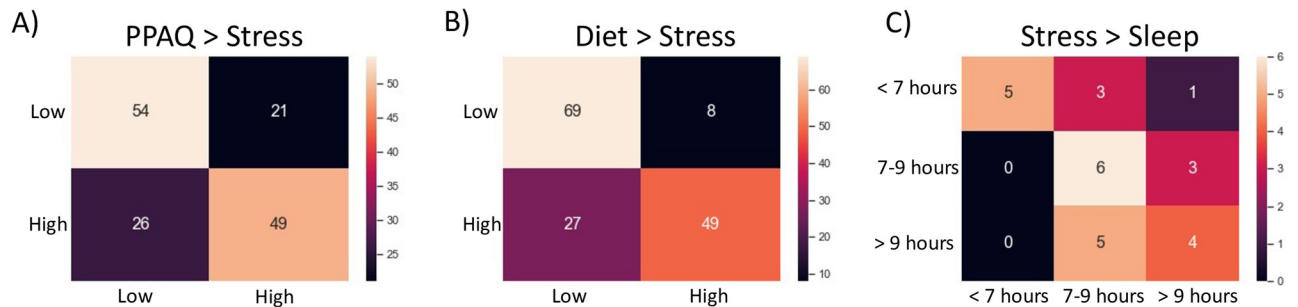


Fig 2. Confusion matrices of classifier performance. Classification models predicting A) stress class from physical activity; B) stress class from diet; and C) amount of sleep from perceived stress.

<https://doi.org/10.1371/journal.pone.0272862.g002>

higher levels of exercise, they also reported higher levels of sedentary activity and housework. In fact, across all features, high stress women reported higher levels of physical activity.

The average responses to these physical activity features suggest that while research has shown that physical activity may be protective against stress, high amounts may be detrimental. It must also be noted that being required to spend large amounts of time doing these activities could itself be related to increased stress. For example, spending some time every day doing housework and active transportation is likely to have benefits, especially over more sedentary activities such as watching television. However, spending a lot of time doing these activities may be understandably stressful, or may represent a more stressful home environment, and thus may counteract the benefits of the physical activity. It has previously been reported that pregnant women who engaged in 150 minutes of physical activity per week showed better mental health outcomes than those who performed less [3]. Importantly, in the present study increased stress levels were associated with *higher* reported frequency of these activities in pregnant women, suggesting that too much physical activity may be just as detrimental to mental health as too little.

Dietary behaviours that proved to be most predictive of stress levels were related to adding salt at the table, eating fried foods, and eating red meat. Women who fell into the high stress class responded to these items as performing them more frequently than women who fell into the low stress class. In contrast to high stress women, the majority of women in the low stress class reported performing all three of these dietary behaviours less than once a week. Taken together, healthier dietary behaviours were predictive of lower stress levels, suggesting that this may be a modifiable lifestyle factor for promoting resilience to stress during pregnancy.

The stress features that most contributed to the prediction of sleep class also revealed an interesting pattern. Although pregnant women who were achieving too little or too much sleep appeared to feel more ‘on top of things’ and that ‘things were going their way’, they more frequently reported feeling unable to control their anger because of things outside of their control than women who achieved 7–9 hours of sleep per night. This form of stress is particularly present during the pandemic, where sudden, consequential changes to circumstances have been frequent. The direct relationship between this form of stress and sleep levels suggests that we may see increased levels of poor maternal outcomes during this period.

Although we reported the top three features for each classifier, it is important to remember that in all three models, many features contributed to the SVM’s performance, indicating that it is a *set* of choices surrounding each resilience factor that predicted stress class and sleep. The lifestyle factors we examined here are extremely multi-faceted, and ML techniques are able to account for complex relationships between variables [32]. This is promising for those pregnant

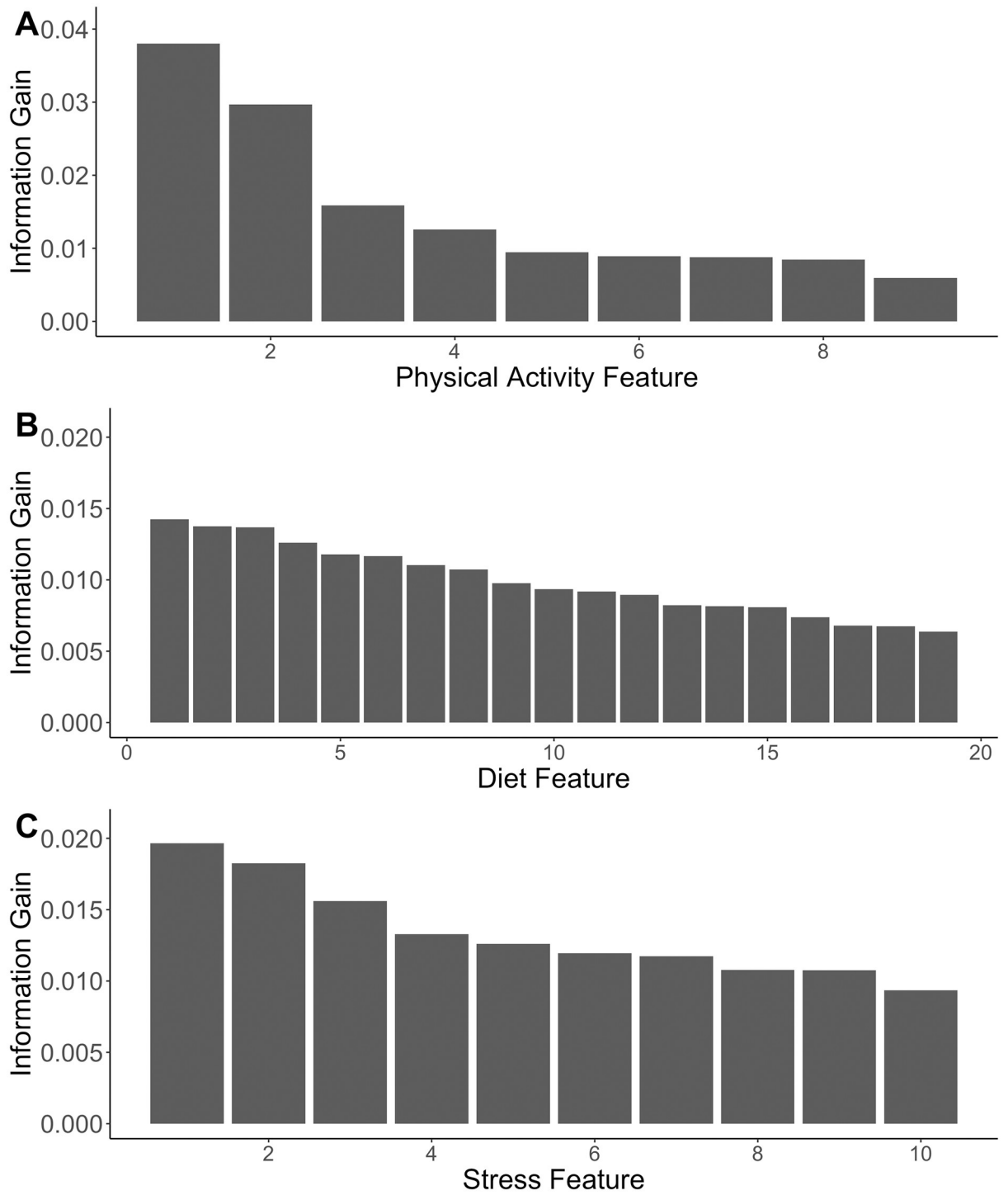


Fig 3. Features, ranked by amount of information gain they contribute. Classification models predicting A) stress class from physical activity; B) stress class from diet; and C) amount of sleep from perceived stress.

<https://doi.org/10.1371/journal.pone.0272862.g003>

Table 6. Feature corresponding to feature index in Fig 3.

Feature index	Feature
Physical Activity	
1	Sports/Exercise
2	Household/Caregiving activity
3	Sedentary activity
4	Moderate activity
5	Inactivity activity
6	Vigorous activity
7	Light activity
8	Transportation
9	Occupational activity
Diet	<i>How often do you eat/drink. . .</i>
1	Add salt to food at the table
2	Deep fried foods (deep fried chicken, fish or seafood; French fries, onion rings)
3	Beef, pork or lamb as main dish
4	Pasta, rice, noodles
5	Fish/Seafood (not fried, but broiled, baked, poached, canned)
6	Calorie-containing beverages (e.g. regular soda, fruit drinks, Nestea, Gatorade)
7	Baked products (donuts, cookies, muffins, crackers, cakes, sweet rolls, pastries)
8	Margarine (stick-type not tub)
9	Carrots
10	Broccoli, broccoli rabe, cauliflower, cabbage, brussels sprouts
11	Whole milk dairy foods (whole milk, hard cheese, butter, ice cream)
12	Processed meats (sausages, salami, bologna, hot dogs, bacon)
13	Whole eggs
14	Dark green leafy vegetables (spinach, romaine lettuce, mesclun mix, kale, turnip greens, bok choy, swiss chard)
15	Other fruits (e.g. fresh apples or pears, bananas, berries, grapes, melons)
16	Whole grain foods (e.g. whole grain breads, brown rice)
17	Citrus fruits (e.g. orange or grapefruit juice, oranges, grapefruits)
18	Low-fat milk products (e.g. low- fat/skim milk, yogurt, cottage cheese)
19	Other vegetables (e.g. peas, corn, green beans, tomatoes, squash)
Stress	<i>In the last month, how often have you felt or thought that. . .</i>
1	Things are going my way
2	I am on top of things
3	Angered by things out of my control
4	Difficulties piling up
5	Confident I can handle personal problems
6	I cannot cope with everything I have to do
7	I can control irritations in my life
8	I am upset by things happening unexpectedly
9	I cannot control the important things in my life
10	I feel nervous and stressed

<https://doi.org/10.1371/journal.pone.0272862.t006>

women who may be finding it difficult to commit to moderate physical activity every day or to limiting unhealthy foods in their diet. That is, small changes to multiple habits, for example moderately reducing salt, fried foods, and fish, as well as other unhealthy foods, may have a combined beneficial effect and are easier to accomplish than cutting them out completely.

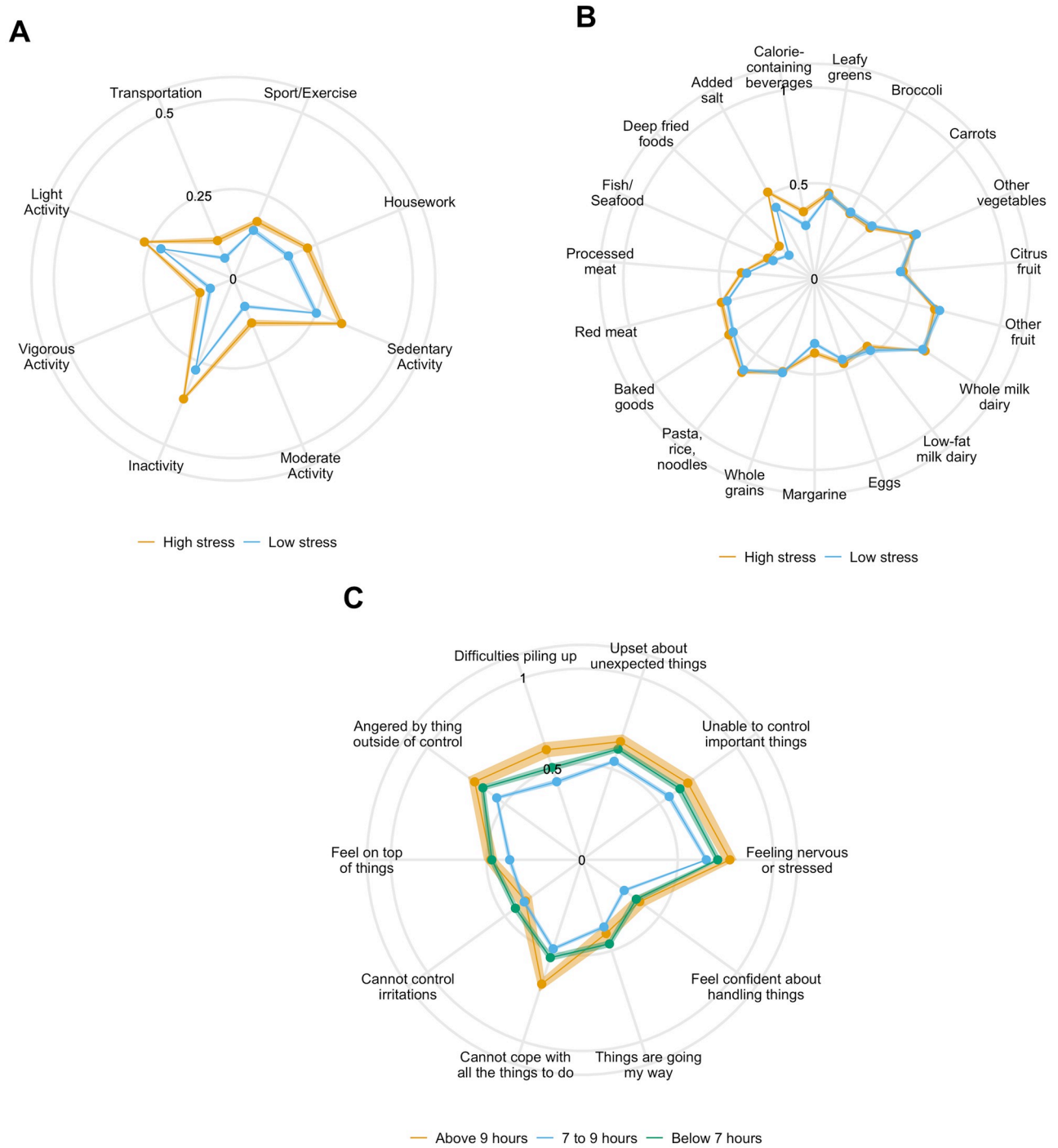


Fig 4. Distribution of mean responses by target class. A) Pregnancy Physical Activity Questionnaire features B) PrimeScreen features and C) Perceived Stress Scale features. Responses have been normalized to range from 0–1.

<https://doi.org/10.1371/journal.pone.0272862.g004>

The present results provide important insight into the relationship between diet, physical activity, stress, and sleep in pregnant women; however, several limitations must be noted. Although the sample was large and heterogeneous in terms of age and location, sampling was limited to those pregnant women who had access to an internet-connected device and saw our

study advertisement, possibly biasing the sample. Additionally, because the study was conducted online, it was not possible to confirm that participants were in fact pregnant. However, because compensation was minimal, the number of pregnancy-specific questions was large, and time to complete the study was approximately one hour, it is unlikely that non-pregnant individuals completed the entire study and therefore were included in the final analysis. Finally, the cross-sectional design limits conclusions that can be drawn about the cause and effect of our results; future work should aim to determine the causal relationship between these factors in pregnant women.

Psychological distress in pregnancy is a risk factor for adverse outcomes for maternal and neonatal health. It is critical to identify the signs, symptoms, and diagnostic thresholds that warrant prenatal intervention. This research adds to a developing consensus regarding adverse health behaviours concerning physical activity and diet, and the association with prenatal stress and the subsequent associations with sleep in pregnant women during the pandemic. Predictive modeling developed using ML-approaches may be used as a screening tool and to promote positive health behaviours for pregnant women who experience high levels of antenatal stress.

Author Contributions

Conceptualization: Emily S. Nichols, Michelle F. Mottola, Isabelle Giroux, Ryan J. Van Lieshout, Yalda Mohsenzadeh, Emma G. Duerden.

Data curation: Emily S. Nichols, Harini S. Pathak, Roberta Bgeginski.

Formal analysis: Harini S. Pathak, Roberta Bgeginski.

Funding acquisition: Emily S. Nichols, Roberta Bgeginski, Emma G. Duerden.

Investigation: Emily S. Nichols.

Methodology: Emily S. Nichols, Roberta Bgeginski, Michelle F. Mottola, Isabelle Giroux, Ryan J. Van Lieshout, Yalda Mohsenzadeh, Emma G. Duerden.

Project administration: Emily S. Nichols.

Resources: Emma G. Duerden.

Software: Harini S. Pathak.

Supervision: Emily S. Nichols, Michelle F. Mottola, Yalda Mohsenzadeh, Emma G. Duerden.

Validation: Emma G. Duerden.

Visualization: Emily S. Nichols.

Writing – original draft: Emily S. Nichols.

Writing – review & editing: Emily S. Nichols, Harini S. Pathak, Roberta Bgeginski, Michelle F. Mottola, Isabelle Giroux, Ryan J. Van Lieshout, Yalda Mohsenzadeh, Emma G. Duerden.

References

1. Profit J, Edmonds BT, Shah N, Cheyney M. The COVID-19 Pandemic as a Catalyst for More Integrated Maternity Care. *American journal of public health*. United States; 2020. pp. 1663–1665. <https://doi.org/10.2105/AJPH.2020.305935> PMID: 33026864
2. Cao-Lei L, Massart R, Suderman MJ, Machnes Z, Elgbeili G, Laplante DP, et al. DNA methylation signatures triggered by prenatal maternal stress exposure to a natural disaster: Project Ice Storm. *PLoS one*. 2014; 9: e107653. <https://doi.org/10.1371/journal.pone.0107653> PMID: 25238154

3. Davenport MH, Meyer S, Meah VL, Strynadka MC, Khurana R. Moms Are Not OK: COVID-19 and Maternal Mental Health. *Frontiers in Global Women's Health*. 2020. p. 1.
4. Cardwell MS. Stress: pregnancy considerations. *Obstetrical & gynecological survey*. 2013; 68: 119–129. <https://doi.org/10.1097/OGX.0b013e31827f2481> PMID: 23417218
5. Phelan AL, DiBenedetto MR, Paul IM, Zhu J, Kjerulff KH. Psychosocial Stress During First Pregnancy Predicts Infant Health Outcomes in the First Postnatal Year. *Maternal and child health journal*. 2015; 19: 2587–2597. <https://doi.org/10.1007/s10995-015-1777-z> PMID: 26152890
6. De Weerth C. Prenatal stress and the development of psychopathology: Lifestyle behaviors as a fundamental part of the puzzle. *Development and psychopathology*. 2018; 30: 1129–1144. <https://doi.org/10.1017/S0954579418000494> PMID: 30068418
7. Wang C, Pan R, Wan X, Tan Y, Xu L, Ho CS, et al. Immediate Psychological Responses and Associated Factors during the Initial Stage of the 2019 Coronavirus Disease (COVID-19) Epidemic among the General Population in China. *International journal of environmental research and public health*. 2020; 17: 1729. <https://doi.org/10.3390/ijerph17051729> PMID: 32155789
8. Gildner TE, Laugier EJ, Thayer ZM. Exercise routine change is associated with prenatal depression scores during the COVID-19 pandemic among pregnant women across the United States. *PloS one*. 2020; 15: e0243188. <https://doi.org/10.1371/journal.pone.0243188> PMID: 33347484
9. Physical Activity and Exercise During Pregnancy and the Postpartum Period: ACOG Committee Opinion, Number 804. *Obstetrics & Gynecology*. 2020;135.
10. Mccullough LE, Miller EE, Calderwood LE, Shivappa N, Steck SE, Forman MR, et al. Maternal inflammatory diet and adverse pregnancy outcomes: Circulating cytokines and genomic imprinting as potential regulators? *Epigenetics*. 2017; 12: 688–697. <https://doi.org/10.1080/15592294.2017.1347241> PMID: 28678596
11. Trujillo-Mayol I, Guerra-Valle M, Casas-Forero N, Sobral MMC, Viegas O, Alarcón-Enos J, et al. Western Dietary Pattern Antioxidant Intakes and Oxidative Stress: Importance During the SARS-CoV-2/ COVID-19 Pandemic. *Advances in Nutrition*. 2021. <https://doi.org/10.1093/advances/nmaa171> PMID: 33439972
12. Zhang J, Zhang Y, Huo S, Ma Y, Ke Y, Wang P, et al. Emotional Eating in Pregnant Women during the COVID-19 Pandemic and Its Association with Dietary Intake and Gestational Weight Gain. *Nutrients*. 2020; 12: 2250. <https://doi.org/10.3390/nu12082250> PMID: 32731389
13. Niles MT, Bertmann F, Morgan EH, Wentworth T, Biehl E. The Impact of Coronavirus on Vermonters Experiencing Food Insecurity. *College of Agriculture and Life Sciences Faculty Publications*. 2020; 19.
14. Sivertsen B, Hysing M, Dørheim SK, Eberhard-Gran M. Trajectories of maternal sleep problems before and after childbirth: a longitudinal population-based study. *BMC pregnancy and childbirth*. 2015; 15: 129. <https://doi.org/10.1186/s12884-015-0577-1> PMID: 26031504
15. Xie M, Wang X, Zhang J, Wang Y. Alteration in the psychologic status and family environment of pregnant women before and during the COVID-19 pandemic. *International Journal of Gynecology & Obstetrics*. 2021; 153: 71–75. <https://doi.org/10.1002/ijgo.13575> PMID: 33403679
16. Lin W, Wu B, Chen B, Lai G, Huang S, Li S, et al. Sleep Conditions Associate with Anxiety and Depression Symptoms among Pregnant Women during the Epidemic of COVID-19 in Shenzhen. *Journal of affective disorders*. 2021; 281: 567–573. <https://doi.org/10.1016/j.jad.2020.11.114> PMID: 33261931
17. Palagini L, Gemignani A, Banti S, Manconi M, Mauri M, Riemann D. Chronic sleep loss during pregnancy as a determinant of stress: impact on pregnancy outcome. *Sleep medicine*. 2014; 15: 853–859. <https://doi.org/10.1016/j.sleep.2014.02.013> PMID: 24994566
18. Islam MN, Mustafina SN, Mahmud T, Khan NI. Machine learning to predict pregnancy outcomes: a systematic review, synthesizing framework and future research agenda. *BMC Pregnancy Childbirth*. 2022; 1–19. <https://doi.org/10.1186/s12884-022-04594-2> PMID: 35546393
19. Lee J, Cohen J, Karvir H, Beim PY, Barritt J, Copperman AB. Using Machine Learning Algorithms to Identify Risk Factors Correlated With Ectopic Pregnancies at a Large IVF Program. *Fertil Steril*. 2013; 99: S4. <https://doi.org/10.1016/j.fertnstert.2013.01.005>
20. Kopanitsa G, Metsker O, Paskoshev D, Greschischeva S. Identification of Risk Factors and Prediction of Sepsis in Pregnancy Using Machine Learning Methods. *Procedia Comput Sci*. 2021; 193: 393–401. <https://doi.org/10.1016/j.procs.2021.10.040>
21. Zhang Z, Yang L, Han W, Wu Y, Zhang L, Gao C, et al. Machine Learning Prediction Models for Gestational Diabetes Mellitus: Meta-analysis. *J Med Internet Res*. 2022; 24. <https://doi.org/10.2196/26634> PMID: 35294369
22. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. *Journal of Health and Social Behavior*. US: American Sociological Assn; 1983. pp. 385–396. <https://doi.org/10.2307/2136404> PMID: 6668417

23. Solivan AE, Xiong X, Harville EW, Buekens P. Measurement of Perceived Stress Among Pregnant Women: A Comparison of Two Different Instruments. *Maternal and Child Health Journal*. 2015; 19: 1910–1915. <https://doi.org/10.1007/s10995-015-1710-5> PMID: 25652063
24. Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ, III CFR, et al. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry research*. 1989; 28: 193–213. [https://doi.org/10.1016/0165-1781\(89\)90047-4](https://doi.org/10.1016/0165-1781(89)90047-4) PMID: 2748771
25. Skouteris H, Wertheim EH, Germano C, Hons B, Paxton SJ, Milgrom J. A Study Across Two Time Points Examining the Pittsburgh Sleep Quality Index and Associations with Depressive Symptoms. *Women's Health Issues*. 2009; 19: 45–51. <https://doi.org/10.1016/j.whi.2008.10.004> PMID: 19111787
26. Chasan-Taber L, Schmidt MD, Roberts DE, Hosmer D, Markenson G, Freedson PS. Development and validation of a Pregnancy Physical Activity Questionnaire. *Medicine and science in sports and exercise*. 2004; 36: 1750–1760. <https://doi.org/10.1249/01.mss.0000142303.49306.0d> PMID: 15595297
27. Rifas-Shiman SL, Willett WC, Lobb R, Kotch J, Dart C, Gillman MW. PrimeScreen, a brief dietary screening tool: reproducibility and comparability with both a longer food frequency questionnaire and biomarkers. *Public health nutrition*. 2001; 4: 249–254. <https://doi.org/10.1079/phn200061> PMID: 11299098
28. Peng H, Long F, Ding C. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2005; 27: 1226–1238. <https://doi.org/10.1109/TPAMI.2005.159> PMID: 16119262
29. Guyon I, Elisseeff A. An Introduction to Variable and Feature Selection 1 Introduction. *Journal of Machine Learning Research*. 2003; 3: 1157–1182.
30. Pruessner JC, Hellhammer DH, Kirschbaum C. Burnout, perceived stress, and cortisol responses to awakening. *Psychosomatic medicine*. 1999; 61: 197–204. <https://doi.org/10.1097/00006842-199903000-00012> PMID: 10204973
31. Mindell JA, Ann R, Nikolovski J. Sleep patterns and sleep disturbances across pregnancy. *Sleep Medicine*. 2015; 16: 483–488. <https://doi.org/10.1016/j.sleep.2014.12.006> PMID: 25666847
32. Bzdok D, Altman N, Krzywinski M. Points of Significance: Statistics versus machine learning. *Nature Methods*. 2018; 15: 233–234. <https://doi.org/10.1038/nmeth.4642> PMID: 30100822