Deleterious and Protective Psychosocial and Stress-Related Factors Predict Risk of Spontaneous Preterm Birth

Martin Becker, PhD1,* Jonathan A. Mayo, PhD2,* Nisha K. Phogat, MS3 Cecele C. Quaintance, MS2 Ana Laborde, BA2 Lucy King, BA4 Ian H. Gotlib, PhD4 Brice Gaudilliere, MD, PhD1 Martin S. Angst, MD1 Gary M. Shaw, PhD2,5 David K. Stevenson, MD2 Nima Aghaeepour, PhD1 Firdaus S. Dhabhar, PhD3,5

1 Department of Anesthesiology, Perioperative and Pain Medicine, Stanford University School of Medicine, Stanford, California 2 Department of Pediatrics, Stanford University School of Medicine, Stanford, California 3 Department of Psychiatry & Behavioral Sciences and Sylvester Comprehensive Cancer Center, Miller School of Medicine, University of Miami, Miami, Florida 4 Department of Psychology, Stanford University, Stanford, California 5 Department of Microbiology & Immunology, Miller School of Medicine, Univ. of Miami, Miami, Florida

Am J Perinatol

Abstract

Objectives The aim of the study was to: (1) Identify (early in pregnancy) psychosocial and stress-related factors that predict risk of spontaneous preterm birth (PTB, gestational age <37 weeks); (2) Investigate whether “protective” factors (e.g., happiness/social support) decrease risk; (3) Use the Dhabhar Quick-Assessment Questionnaire for Stress and Psychosocial FactorsTM (DQAQ-SPF™) to rapidly quantify harmful or protective factors that predict increased or decreased risk respectively, of PTB.

Study Design This is a prospective cohort study. Relative risk (RR) analyses investigated association between individual factors and PTB. Machine learning-based interdependency analysis (IDPA) identified factor clusters, strength, and direction of association with PTB. A nonlinear model based on support vector machines was built for predicting PTB and identifying factors that most strongly predicted PTB.

Results Higher levels of deleterious factors were associated with increased RR for PTB: General anxiety (RR = 8.9; 95% confidence interval or CI = 2.0,39.6), pain (RR = 5.7; CI = 1.7,17.0); tiredness/fatigue (RR = 3.7; CI = 1.0,13.5); perceived risk of birth complications (RR = 4; CI = 1.6,10.01); self-rated health current (RR = 2.6; CI = 1.0,6.7) and previous 3 years (RR = 2.9; CI = 1.1,7.7); and divorce (RR = 2.9; CI = 1.1,7.8). Lower levels of protective factors were also associated with increased RR for PTB: low happiness (RR = 9.1; CI = 1.25,71.5); low support from parents/siblings (RR = 3.5; CI = 0.9,12.9), and father-of-baby (RR = 3; CI = 1.1,9.9). These factors were also components of the...
Premature or preterm birth (PTB) is defined as a live birth that occurs before completion of 37 weeks of pregnancy. Over 380,000 PTBs occur per year in the United States with an estimate of 15 million per year worldwide.\(^1\)–\(^3\) In addition to the significant burden of infant morbidity and mortality associated with PTB, infants who survive are vulnerable to long-term health consequences and their economic costs. Numerous medical, demographic, and psychosocial factors have been associated with PTB, including genetics, prior spontaneous PTB, short cervix, African-American race, multiple gestations, urogenital tract infections, high blood pressure, diabetes/gestational diabetes, obesity, substance abuse, alcohol consumption, and smoking.\(^4\)–\(^13\) Maternal chronic stress\(^6\)–\(^8\),\(^14\)–\(^18\) and stress-related factors such as maladaptive coping skills, and depressive symptoms have been identified as risk factors for PTB.\(^19\),\(^20\) With respect to racial disparities in PTB: in 2019, African–American women showed an approximately 50% higher rate of PTB (14.4%) compared with White (9.3%) or Hispanic (10%) women. African–American women report racial discrimination as a significant source of chronic social stress.\(^21\) Therefore, it is possible that in addition to other factors, psychosocial and stress-related factors are significant contributors to racial disparities in PTB.

In the present study we aimed to: (1) Conduct an integrative investigation of a wide range of deleterious social and chronic stress-related factors and quantify their association with increased risk of PTB; (2) Investigate potential protective effects of “positive” factors such as happiness and social support; and (3) Investigate whether the newly designed Dhabhar Quick-Assessment Questionnaire for Stress and Psychosocial FactorsTM (DQAQ-SPF™) can be used to quantify a wide range of stress-related and psychosocial factors relatively rapidly and with minimal demands of time and resources on study participants and investigators.

Stress is defined as a constellation of events that begins with a stimulus (stressor), which precipitates a reaction in the brain (stress perception and evaluation) that results in the activation of fight-or-flight systems in the body (biological stress response).\(^22\) Stressors can be internal or external, and numerous factors (e.g., psychological, social, physical, physiological, or environmental) can act as stressors. Different stressors activate a common biological response with epinephrine, norepinephrine, and cortisol being the major hormones that are secreted in the circulation during stress (other hormones are also released, and the relative proportions of hormones can be affected by the type of stressor). Therefore, it is possible that the effects of PTB risk factors are at least partially mediated through stress-related biological pathways.

To understand the nature of the health-relevant effects of stress, it is useful to divide stress into two broad categories: short-term or acute stress has been defined as that which involves stress-related biological changes that last for minutes to hours in duration. In contrast, chronic stress is defined as that which induces changes that last for months to years.\(^23\) It has been shown that short-term stress experienced at the time of immune activation significantly enhances immune function.\(^22\) Indeed, it has been suggested that short-term stress (Mother Nature’s fight-or-flight stress response), experienced during dangerous situations (predator attack), medical procedures (vaccination or surgery), or challenging conditions (courtship, taking an examination, or a job interview), may be Mother Nature’s mechanism for enhancing protection and performance.\(^24\) In contrast,

### Key Points

- Newly designed questionnaire used for rapid quantification of stress and psychosocial factors early during pregnancy.
- Deleterious factors predict increased preterm birth (PTB) risk.
- Protective factors predict decreased PTB risk.
chronic stress can exert significant deleterious effects on brain and body.\textsuperscript{25}

In this study we used the newly designed DQAQ-SPF\textsuperscript{TM} to prospectively quantify chronic stress and psychosocial factors, as well as protective factors (that act as buffers against chronic stress) and investigated their predictive associations with risk of PTB. Expectant mothers took the DQAQ-SPF early in pregnancy, as part of a prospective cohort study conducted at the March of Dimes Prematurity Research Center at Stanford University. Our overarching hypotheses were that deleterious psychosocial factors and chronic stress would be associated with increased risk of PTB, and protective psychosocial factors would be associated with decreased risk of PTB.

**Materials and Methods**

The data analyzed here come from our Prematurity Research Center Cohort (PRCC). The PRCC was established in 2011 at Stanford University Hospitals and Clinics. Briefly, pregnant women were enrolled prior to 12 weeks gestation. These women were followed weekly until delivery to study PTB using both questionnaire-derived and biological data. We collected clinical and demographic data that included race/ethnicity, education, and age. Race/ethnicity was grouped as non-Hispanic White, non-Hispanic Black, Asian, Pacific Islander, Hispanic, American Indian/Alaskan Native, Other, and Missing. Education was categorized as some high school or less, high school diploma or equivalent, some college, college graduate or more, and missing. Age was binned into 5-year intervals with \( <20 \) and \( >40 \)-year tails, and missing. Pre-pregnancy body mass index (BMI) was obtained. Two questions regarding the mother’s perceived risk of birth complications, or birth defects, were personally communicated to Dr. Shaw by Dr. Michael S. Kramer (\textsuperscript{►}Supplementary Table S1, available in the online version only). Moreover, we collected data regarding stress- and health-related variables using the newly designed DQAQ-SPF\textsuperscript{TM} questionnaire (\textsuperscript{►}Supplementary Table S2, available in the online version only). The DQAQ-SPF\textsuperscript{TM} was designed to rapidly assess (with minimal burden on participants and investigators) deleterious and protective psychosocial factors, chronic stress, stressful life events, self-rated health, and factors related to emotions, personality, and sleep. Participants were requested to rate items on an eight-point scale, or to provide “yes/no” answers about whether they had experienced specific stressful life experiences (e.g., divorce). Participants were free to enter “not applicable” where relevant or to skip any question that they did not wish to answer. Based on their ratings, participants were divided into low (0, 1, 2), moderate (3, 4, 5), and high (6, 7) groups. The two questions related to self-rated health were modified from the general health question that is part of the Medical Outcomes Study 36-Item Short-Form Health Survey (SF-36).\textsuperscript{26} Questionnaires were administered during the enrollment visit – typically 11 (median) weeks of gestation (\textsuperscript{►}Fig. 1). All data collection associated with the PRCC was approved by the Stanford IRB.

**Risk Factors for Spontaneous Preterm Birth**

Becker et al.

**Fig. 1** Distribution of gestational age at questionnaire completion. In most instances (>95\%), the questionnaires were completed before 21 weeks of gestation. The median gestational age when completing the questionnaire was 11 weeks.

**Statistical Analyses**

**Relative Risk Analysis**

Poisson regression was used to estimate the crude relative risk (RR) and 95% confidence interval (CI) for PTB related to the assessed factors. Poisson regression was modeled using SAS 9.4 (Cary, NC).

**Interdependency Analysis**

The factors assessed by the DQAQ-SPF\textsuperscript{TM} can exhibit complex interdependencies among themselves, as well as associations with PTB or gestational age (GA) at delivery. To analyze these interdependencies, we calculated an interdependency network that enables visualization of the correlation structure between the obtained variables while at the same time enabling visualization of the association of each variable with GA. For this, we calculated the Spearman correlation between each variable pair across all collected samples (omitting undefined values). The resulting correlation vectors for each feature were then grouped into 15 clusters using the K-Means algorithm. The number of clusters was chosen based on Inertia (i.e., “within cluster sum-of-squares”) using the Kneadle algorithm\textsuperscript{27} in offline mode with sensitivity set to 1.0. To account for random fluctuations, we took the average number of determined clusters over 100 runs (number of clusters = 15.35 ± 2.67). The averaged elbow plot and the mean elbow value are shown in \textsuperscript{►}Fig. 2. The two-dimensional embedding used for placing the features was calculated using the t-distributed stochastic neighbor embedding (t-SNE)\textsuperscript{28} on the absolute values of the correlation vectors associated with each variable. Finally, we computed an association between the features and PTB using the Wilcoxon Rank-Sums coefficient for numeric and ordinal variables, the Chi-square test for categorical variables, and Fisher’s exact test for binary variables.

**Predictive Modeling of PTB**

For predictive modeling, we preprocessed data from the DQAQ-SPF questionnaire as follows. First, we removed all
The distribution of GA at delivery for these participants is shown in Fig. 2. Cessing procedure left 200 participants with 79 features. The imputed with medians for each feature. The overall preprocessing procedure left 200 participants with 79 features. The remaining missing values were imputed with medians for each feature. The overall preprocessing procedure left 200 participants with 79 features. The distribution of GA at delivery for these participants is shown in Fig. 2. We then trained, tested, and evaluated a nonlinear model (a support vector machine [SVM] with a radial basis function kernel with \( C = 1.0 \) and \( \gamma = \frac{1}{(\text{number of features} \cdot \text{global feature variance})} \)) using a repeated (100 times) 10-fold cross-validation setup. These parameters represent default settings. Further parameter optimization did not yield significant improvements. We reported mean area under the receiver operating characteristic curve (AUC) values as predictive power for the model. The mean was calculated by first computing the AUC across all folds of one repetition based on the predicted decision function values of the SVM and then taking the mean over all these repetitions.

**Model Reduction via Forward Selection**
To explore the possibility of using smaller sets of psychosocial and stress-related factors to predict PTB, we used an approach similar to forward selection in linear regression (see, e.g., Blanchet et al.\(^{29}\)). Specifically, we started with an empty set of features and iteratively added variables that improved the predictive power of the current model mostly according to the AUC evaluation metric. That is, in each step, we retrained our model adding one variable at a time. The added variable that results in the model with the highest AUC is considered the most valuable and added to the model. The new model is then based on one more variable. We continued this process until all variables were added to the model, yielding the overall model described above. Note that the AUC for selecting variables was calculated based on 100 times repeated 10-fold cross-validation equivalent to the previously described predictive experiments. The resulting variable trace is shown. For comparison, we also provide a similar visualization where variables are ordered by the statistical power of their association with PTB (see interdependency or IDP analysis for the applied statistical tests). Note that this forward selection process is “greedy” and that the variable selection itself is not embedded in a cross-validation procedure. This means that: (a) the presented results may not be the only valid or optimal combination of variables; and (b) the corresponding AUC values only hint at, but do not guarantee, increased predictive power. Nevertheless, this approach is well suited for exploring the predictive power of smaller variable subsets and can inform future investigations.

**Results**
Analyses are based on our cohort of 283 pregnant women, 248 of whom delivered at term and 35 of whom delivered before 37 weeks completed gestation. Table 1 shows the characteristics of the study population. Non-Hispanic White women formed the largest percentage in the Term group, while Hispanic women formed the largest percentage of the PTB group \((p < 0.01)\). A larger percentage of the Term group obtained a college degree or pursued a higher level of education, while a larger percentage of the PTB group had a high school or lower level of education \((p < 0.005)\). The Term and PTB groups did not differ significantly in maternal age. A larger percentage of the Term group had normal BMI, while a larger percentage of the PTB group were classified as having Obese I or higher BMI values \((p < 0.05)\). The Term and PTB groups differed in marital status: a larger percentage of the Term group was married \((p < 0.005)\).

**Relative Risk**
Table 2 shows the risk of PTB in association with psychosocial and chronic stress-related factors classified into three categories: (1) prenatal factors; (2) perinatal factors; and (3) “protective” factors that were hypothesized to reduce PTB...
### Table 1  Characteristics of study population

<table>
<thead>
<tr>
<th></th>
<th>Term</th>
<th>N</th>
<th>%</th>
<th>Preterm</th>
<th>N</th>
<th>%</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td>248</td>
<td>35</td>
<td></td>
<td>35</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Maternal race/ethnicity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.0038</td>
</tr>
<tr>
<td>Non-Hispanic White</td>
<td>103</td>
<td>11</td>
<td>31.43</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Hispanic Black</td>
<td>10</td>
<td>3</td>
<td>8.57</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>47</td>
<td>3</td>
<td>8.57</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pacific Islander</td>
<td>2</td>
<td>2</td>
<td>5.71</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>61</td>
<td>13</td>
<td>37.14</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>American Indian/Alaskan Native</td>
<td>1</td>
<td>2</td>
<td>5.71</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>18</td>
<td>1</td>
<td>2.86</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>6</td>
<td>2</td>
<td>2.42</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Maternal education</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.001</td>
</tr>
<tr>
<td>Some high school or less</td>
<td>23</td>
<td>6</td>
<td>17.14</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Highschool diploma/GED</td>
<td>15</td>
<td>5</td>
<td>14.29</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Some college</td>
<td>44</td>
<td>6</td>
<td>17.14</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>College graduate or more</td>
<td>162</td>
<td>14</td>
<td>40</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>4</td>
<td>4</td>
<td>11.43</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Maternal age</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.0889</td>
</tr>
<tr>
<td>&lt;20 y</td>
<td>6</td>
<td>2.42</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20–24 y</td>
<td>16</td>
<td>7</td>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25–29 y</td>
<td>62</td>
<td>5</td>
<td>14.29</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30–34 y</td>
<td>105</td>
<td>13</td>
<td>37.14</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>35–39 y</td>
<td>48</td>
<td>7</td>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40+ y</td>
<td>10</td>
<td>3</td>
<td>8.57</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>1</td>
<td>0.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Pre-pregnancy BMI</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.0171</td>
</tr>
<tr>
<td>Underweight</td>
<td>9</td>
<td>1</td>
<td>2.86</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>125</td>
<td>8</td>
<td>22.86</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overweight</td>
<td>52</td>
<td>11</td>
<td>31.43</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obese I</td>
<td>21</td>
<td>5</td>
<td>14.29</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obese II</td>
<td>14</td>
<td>6</td>
<td>17.14</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obese III</td>
<td>5</td>
<td>2</td>
<td>5.71</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>22</td>
<td>2</td>
<td>5.71</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Marital status</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.0141</td>
</tr>
<tr>
<td>Single</td>
<td>15</td>
<td>2</td>
<td>5.71</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In a relationship</td>
<td>52</td>
<td>10</td>
<td>28.57</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>176</td>
<td>19</td>
<td>54.29</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>5</td>
<td>4</td>
<td>11.43</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Gestational weeks at delivery</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;32 wk</td>
<td>6</td>
<td></td>
<td>17.14</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>32–36 wk</td>
<td>29</td>
<td></td>
<td>82.86</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>37+ wk</td>
<td>248</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: BMI, body mass index; GED, General Education Development.
risk directly, or through buffering/protecting the individual from the deleterious effects of chronic stress.

1. **Prenatal factors and risk of PTB**: 

   **Anxiety**: Compared with women in the low anxious group, women in the high anxious group showed increased risk for PTB (RR = 8.9, 95% CI: 2.0, 39.6). 
   
   **Sadness**: Compared with women in the low sadness group, women in the Moderate Sadness showed increased risk for PTB (RR = 2.8, 95% CI: 1.0, 7.5). 
   
   **Frequency of pain experience**: Compared with women in the low pain group, women in the high pain group showed increased risk for PTB (RR = 5.7, 95% CI: 1.7, 17.0). (Participants were also requested to list where pain was experienced: the most frequently reported anatomical regions as the source of pain were back and head.) 
   
   **Self-rating of health**: Compared with women in the high self-rated health group, those in the moderate self-rated health group showed increased risk for PTB (RR = 2.9, 95% CI: 1.1, 7.7). 
   
   **Divorce**: Compared with women who had not been divorced, those who had been divorced showed increased risk for PTB (RR = 2.9, 95% CI: 1.1, 7.8). Time from divorce to study enrollment ranged from 0 to 17 years and did not differ significantly between Term and PTB groups. 

2. **Perinatal stress-related factors and risk of PTB**: 

   **Perceived risk of birth complications**: Compared with women in the low perceived risk group, those in the high perceived risk group showed increased risk for PTB (RR = 4.0, 95% CI: 1.6, 10.1). 
   
   **Self-rating of current health**: Compared with women in the high self-rated health group, those in the moderate self-rated health group showed increased risk for PTB (RR = 2.6, 95% CI: 1.0, 7.0). 

3. **“Protective” factors that can buffer the individual against the effects of chronic stress**: 

   **Happiness**: Compared with women in the high happiness group, women in the low happiness group showed increased risk for PTB (RR = 9.1, 95% CI: 1.1, 71.5). Women in the Moderate Happiness group also showed increased risk for PTB (RR = 3.0, 95% CI: 1.1, 7.5). 
   
   **Support from parents and siblings**: Compared with women who reported receiving high levels of support from parents and siblings, women who reported receiving moderate levels of support showed increased risk for PTB (RR = 3.4, 95% CI: 1.2, 9.5). 
   
   **Support from father of baby**: Compared with women who reported receiving high levels of support from the father of the baby, women who reported receiving moderate levels of support showed increased risk for PTB (RR = 3.3, 95% CI: 1.1, 9.9).
**Interdependency Analysis**

The factors that were measured exhibit complex interdependency structures. - Fig. 4A shows how different factors cluster according to their correlation structure. Each factor is represented by a node. Round nodes indicate numeric or ordinal variables, squares indicate categorical variables, and crosses indicate binary variables. Proximity and color represent closely related variables on a global level, i.e., the closer two features are, the more similar they can be considered with regard to their correlation structure (see Interdependency Analysis Methods for details). Connecting edges between nodes represent the pairwise correlation between variables (thin edges represent a Spearman correlation ≥0.3, while thick lines represent a correlation >0.5, all edges correspond to correlations passing a Bonferroni corrected p-value of 0.05).

While the clusters of variables and their positioning are based on probabilistic algorithms and thus may fluctuate slightly between runs, it is nevertheless apparent that the variables assessed by the DQAQ-SPF™ questionnaire form logical clusters of semantically related factors. These clusters include, for example: emotion- and personality-related variables (e.g., happiness, sadness, anger, depression, and tiredness/fatigue), body image and health-related variables (e.g., satisfaction with appearance, weight and height, self-rated health—current and during past 3 years), a pain cluster (frequency and intensity), perceived risk of birth complications, and sleep-related variables. Among the most prominent clusters, the emotional variables are most strongly correlated with PTB and, at the same time, strongly related to one another.

- Fig. 4B shows factors within clusters that were significantly associated PTB. The size of the nodes represents the strength of association of each variable with PTB. Bold and underlined node labels indicate significant associations (p < 0.05; no multiple test correction for the visualization). One factor, perceived risk of birth complications, was significant (p < 0.05) after false discovery rate (FDR) correction (Benjamini and Hochberg) p < 0.009 (Wilcoxon Rank-Sums). Importantly, emotions, personality, and physical state-related variables such as “happy” and “anxious” as well as “tired” also showed prominent associations with PTB, as did variables associated with self-rated health such as “current health” and health in the previous 3 years (“health prev 3 y”). Interestingly, the variable regarding eye health ("glasses use") which asked whether the participant used glasses or contact lenses for reading or distance vision, or both, was also associated with PTB. Similarly, the frequency and amount of pain ("pain") were associated with PTB. Finally, there were several life stressors that were associated with PTB such as having experienced a divorce ("divorce self") or a major illness, and social support-related variables such as support from parents and siblings ("family support"), and the number of people you can count on ("support people").

- Fig. 4B also shows the direction of association among the factors themselves and between each factor and PTB. As explained previously, connecting edges between nodes/
factors represent the pairwise correlation between variables (thin edges represent a Spearman correlation $\geq 0.3$ while thick edges represent a correlation $>0.5$; edges also pass a Bonferroni-corrected significance test). Orange connections between nodes represent negative correlations while blue connections represent positive correlations. Similarly, orange nodes are negatively associated with PTB, while blue nodes are positively associated with PTB. We see a general trend for “negative”/health-aversive variables increasing the probability of having a PTB and for “positive”/health-promoting variables decreasing the probability of having a PTB. For example, “tired” and “anxious” from the emotions cluster are positively associated with PTB while “happiness” from the same cluster is negatively associated with PTB.

Feature Clusters and Their Associations with Preterm Birth
For a more detailed examination of specific clusters resulting from our IDP analysis and their individual association with PTB as well as their correlation structures, we visualize the clusters containing the features most associated with PTB (Fig. 5). The factors that showed the strongest association with PTB in each of the feature clusters, perceived risk of negative outcomes (5A), emotions and tiredness/fatigue (5B), and health-related factors (5C) were perceived risk of birth complications (“perceived birth complications risk," $p = 1.10e-04$, Wilcoxon Rank-Sums), anxiousness (“anxious,” $p = 8.32e-03$), self-rated current health (“current health,” $p = 2.62e-02$), respectively.

Predictive Modeling of Preterm Birth
Nonlinear predictive modeling shows a significant association between psychosocial and stress-related factors measured early during pregnancy, and PTB. In particular, our model reaches a mean AUC value of 0.73 $\pm$ 0.02 with a $p$-value of 2.27e $–$ 03 $\pm$ 1.84e $–$ 03 (Wilcoxon Rank-Sum). Fig. 6 shows the corresponding receiver operating characteristic (ROC) curve as well as the distribution of predicted model values (derived from the SVM’s decision function) for term and preterm pregnancies.

Variable Analysis by Model Reduction through Forward Variable Selection
Model reduction through forward selection (Figs. 7 and 8) reveals that even a small set of psychosocial and stress-related variables can potentially drive a model for PTB, and that predictive power may be improved in follow-up studies. In particular, Fig. 7 shows models with an increasing number of variables, while Fig. 8 shows the corresponding univariate associations sorted by $p$-value. Note that several features which are highly correlated with PTB are selected by the forward selection procedure (e.g., “perceived birth complication risk,” “tired,” “pain,” “sad”). The model also incorporates several other variables that are not significantly associated with PTB by themselves (e.g., “pessimistic,” “angry,” and “divorce parent”). Overall, through a greedy selection process we identify a set of 15 out of 79 variables that may, if validated, be particularly promising for building models with high predictive power for PTB. The selected variables consist of a mix of emotion, sleep, life event, social support, and perceived risk, features from 9 of the 15 established clusters. This indicates the importance of psychosocial and stress-related factors for predicting PTB. Also, a model using only the first three variables (from three different clusters), i.e., “pessimistic,” “angry,” and “divorce parent,” has the potential to perform similarly to our overall model (see the dashed line which represents the predictive power of the overall model). While these particular variables may be specific to our dataset.

Fig. 5 Three feature clusters and their relation to gestational age. Each subplot shows a specific cluster from Fig. 4: perceived pregnancy risks (A), emotions and tiredness/fatigue (B), and self-rated health (C). For each cluster we visualize the relation of feature values and PTB for the feature that shows the strongest association with PTB (marked by the black node border). PTB, preterm birth.
and may not be the only three that yield similarly predictive models, this analysis, including the previously mentioned 15 variable candidates, provides a general notion about variables that can be of interest for future studies.

**Discussion**

Further studies are needed to replicate and build upon the findings reported here. We present ideas below to provide a potential framework for future investigations and for designing interventions to ameliorate the harmful effects of chronic stress and deleterious psychosocial factors on risk of PTB.

**Rationale for Investigating Chronic Stress and Psychosocial Factors in the Context of PTB and for Creating the DQAQ-SPF**

Researchers have suggested that chronic stress-related factors such as perceived stress, pregnancy-related anxiety, adverse life experiences, depression, and lack of social support are risk factors for PTB.\(^7,11,12,14,30,31\) Pregnancy-related anxiety is one chronic stress-related factor that has consistently been associated with PTB in numerous studies.\(^8,32\) Here we set out to investigate the association between chronic stress, and deleterious or protective psychosocial factors, and risk of PTB using a newly created instrument, the DQAQ-SPF, that is designed to quantify stress and...
because of the above-mentioned logistical challenges. Tant information, but where they may not be measured epidemiological studies, for which they may provide impor-

and stress-related factors in the context of studies, including important for enabling the quanti-

and resources compared with traditional methods. This is social and stress-related factors using signi-

Therefore, the DQAQ-SPF used in this study was designed taking multiple psychometric questionnaires/surveys.

in a study because of the time and psychological burden of some instances, potential subjects may refuse to participate possible for studies in which stress is not the main focus. In large and lengthy batteries of questionnaires is generally not of time, resources, and personnel. Moreover, administering variables but are not able to do so given logistic constraints interested in quantifying psychosocial and stress-related factors using a minimal amount of time and resources. We examined factors that have been investigated previously (perceived risks related to pregnancy, stressful life experiences, general anxiety, and support from father-of-baby) and those that, to our knowledge, are reported here for the first time in the context of PTB (e.g., self-rated health, frequency of pain experience, happiness, familial social support, and tiredness/fatigue).

Batteries of psychometric tests designed to quantify an array of factors such as the ones measured by the DQAQ-SPF in this study, often require participants to answer multiple questionnaires in what can be a significantly time-consuming process. Even though the burden of taking tests (for the participant) and of administering and scoring tests (for the research team) can be significant, there can be benefits to using traditional psychometric questionnaires when feasible. However, there are instances when investigators are interested in quantifying psychosocial and stress-related variables but are not able to do so given logistic constraints of time, resources, and personnel. Moreover, administering large and lengthy batteries of questionnaires is generally not possible for studies in which stress is not the main focus. In some instances, potential subjects may refuse to participate in a study because of the time and psychological burden of taking multiple psychometric questionnaires/surveys. Therefore, the DQAQ-SPF used in this study was designed to enable investigators to relatively rapidly quantify psychosocial and stress-related factors using significantly less time and resources compared with traditional methods. This is important for enabling the quantification of psychosocial and stress-related factors in the context of studies, including epidemiological studies, for which they may provide important information, but where they may not be measured because of the above-mentioned logistical challenges.

Psychosocial and Chronic Stress-Related Factors and Risk of Spontaneous PTB

In this prospective study, we found that higher levels of deleterious factors were associated with a significantly increased risk of spontaneous PTB: general anxiety (8.9-fold increased risk), frequency of pain experience (5.7-fold); tiredness/fatigue (3.7-fold); perceived risk of birth complications (4-fold); self-rating of current health (2.6-fold) and of health during the previous 3 years (2.9-fold); and having experienced divorce (2.9-fold increased risk). In contrast to higher levels of deleterious factors, lower levels of protective factors were also associated with increased risk of PTB: low happiness (9-fold increased risk); low support from parents and siblings (3.5-fold), and low support from the father of the baby (3-fold). It is noteworthy that even though a large number of factors were included in the analysis, raising important concerns about Type 1 error or false positive findings, there was not a single instance where the associations discovered were contrary to our a priori hypothesis or to what one would logically expect given published stress-related findings in the context of other health outcomes. Nevertheless, we recognize the importance of replicating (and building on) these findings, which we will attempt to do in the context of additional subjects recruited as part of this study, and encourage other groups to attempt independently of this study.

Some of our findings confirmed previous reports, while others to our knowledge, are being reported here for the first time. Consistent with previous reports, we found strong associations between pregnancy-related anxiety and risk of PTB. Higher perceived risk of birth complications was associated with a 4-fold higher risk of PTB. Higher perceived risk of birth defects was associated with increased risk of PTB, but this association was not statistically significant.
Importantly, the mother’s self-report of high overall anxiety was associated with a 9-fold higher risk of PTB. Also, in agreement with previous reports,7,36 we found that low support from the father of the baby was associated with increased risk of PTB. While studies investigating associations between stressful life experiences and PTB have reported equivocal results, we found that life experiences (such as the mother having gone through a divorce or a major illness) were associated with increased risk of PTB.

We also investigated stress-related factors that, to our knowledge, had not been investigated previously: low self-rated health (current, and over the previous 3 years), high frequency of pain experience, low happiness and greater tiredness/fatigue over the previous 3 years, and low social support from parents and siblings were all associated with increased risk of PTB. Interestingly, while our finding is based on happiness over 3 years preceding the pregnancy, it is in agreement with findings showing that positive affect during pregnancy is associated with longer gestation and reduced risk of PTB.37 Moreover, factors such as self-rated health and social support have been associated with other health-related variables and outcomes such as sleep recovery and cardiovascular health,38–40 and cancer-related quality of life.41

Furthermore, we conducted IDP analysis to identify clusters of related factors, and to elucidate the strength, and direction of association among the various factors, and between each factor and risk of PTB. IDP analysis identified specific clusters of factors that were conceptually related and showed significant (p < 0.05) predictive associations with PTB: perceived risk of birth complications, general anxiety, happiness and tiredness/fatigue, self-rated health, social support, pain, and sleep. Importantly, negatively connoted or deleterious variables appear to promote, and positively connoted or protective variables appear to inhibit, PTB.

Supervised analysis of all factors, subject to cross-validation, produced a model highly predictive of PTB with an AUROC of 0.73 for PTB. This represents a model with significant predictive power (p < 0.0023) and thus illustrates the potential to effectively predict and support the treatment of PTB at an early stage in pregnancy based on psychosocial and stress factors that can easily be quantified by a rapid assessment questionnaire such as the DQAQ-SPF. For translational purposes, it is furthermore important to improve the predictive model and refine the DQAQ-SPF to efficiently assess psychosocial and stress-related variables affecting PTB. Our feature selection procedure illustrates that further improvement of the model may be possible through refinements of the DQAQ-SPF using a condensed set of questions. Doing this would leverage the synergistic effect of combinations of questions by employing multivariate nonlinear models as opposed to relying on univariate associations alone.

**Potential Biological Mechanisms and Targets for Future Studies and Interventions**

The biological systems/pathways through which psychosocial and chronic stress-related factors increase the risk of PTB could include: (1) Inhibition/disruption of the adaptive short-term fight-or-flight stress response23,24,42 that may be critical for maintaining full-term pregnancy and during parturition. (2) Dysregulation of immune function which involves22: (a) Suppression of protective immunity which could be important for maintaining full-term pregnancy and during parturition, and inhibition of which could also contribute to increased infection, including that of the urinary genital tract.22–24 (b) Enhancement of immuno-pathological inflammation driven by local and systemic increases in proinflammatory factors39,43,44 that could have pregnancy-specific and general adverse effects.45 (3) Chronic stress induced disruption of telomere—telomerase physiology and decreased telomere length.46–48 (4) Disruption of systemic and/or organ/cell-specific circadian rhythms.49 It has been suggested that such disruption may contribute to PTB.49

Specific biological factors that mediate stress-induced increases in risk of PTB include: components of the hypothalamic pituitary adrenal (HPA) axis and the sympathetic nervous system, and the downstream effects of chronic stress on the biological systems/pathways discussed above. Wadhwa et al showed that high corticotropin-releasing hormone (CRH) concentrations during gestation week 33 are associated with a 3-fold increase in risk of PTB. They also showed that women who delivered at term had higher circulating CRH levels at the beginning of the third trimester compared with women who delivered post-term.10 However, Kramer et al found no association between maternal stress levels, including pregnancy-related anxiety and CRH.16 These authors suggested that placental CRH, is the source of most of the maternal CRH, that is not significantly correlated with circulating adrenocorticotropic hormone or cortisol, and that this suggests that placental CRH is not significantly involved with HPA-axis stress responses during pregnancy.17

Cortisol could be another candidate mediator of the effects of chronic stress on PTB. Ruiz et al showed that Mexican–American mothers who reported having low family support, and higher acculturation (greater assimilation into dominant culture while losing aspects of native culture), showed increased risk of PTB that was mediated by higher circulating cortisol levels.50 However, Kramer et al reported the surprising (at first glance) finding that higher maternal hair cortisol levels were associated with longer gestation.16 However, given that hair cortisol quantification represents an integrated measure of cortisol over a period of several months, it could be that the “higher” cortisol levels observed by Kramer et al in mothers who delivered at term reflected salubrious physiological levels of cortisol that are normal/required to maintain a healthy pregnancy, and that mothers who delivered preterm were showing lower than “normal” cortisol levels due to dampening of their HPA axes due to chronic stress.23 Clearly, further research is needed to identify biological factors (including catecholamine and other stress-responsive factors) that mediate the effects of psychosocial chronic stress-related factors on increased risk of PTB.

Our findings raise other important questions that warrant further investigation: Are the associations observed between factors in different clusters and PTB risk mediated by
different biological pathways? Or are there a few (or one) common biological pathways (that can be activated by different clusters of psychosocial and stress-related factors) that mechanistically link stress, psychosocial factors, and PTB? We also observed considerable differences in the magnitude of PTB risk associated with different psychosocial and stress-related factors indicating that the risk-enhancing effects of some factors were stronger (e.g., general anxiety, lower levels of happiness) than others (e.g., having been through a divorce or major illness). Do these differences in magnitude of PTB risk indicate different biological mechanisms, or do they indicate differences in the magnitude of activation (of common biological pathways)? Another explanation for differences in magnitude of risk (and a topic for further investigation) is that we measured a limited number of life events fairly coarsely (i.e., with simple yes/no endorsement). In contrast, the more integrative factors (e.g., anxiety, happiness, etc.), that predicted greater PTB risk, may capture the overall effects of a longer period of time and much broader array of life events given that such factors can be shaped by life experiences.

It is also important for future studies to investigate biological mechanisms that mediate the effects of protective factors such as happiness, support from parents and siblings, and from the father of the baby (Ghosh et al. and data presented here). Protective factors could decrease risk of PTB either directly and/or by ameliorating the PTB risk-enhancing effects of deleterious factors. Protective factors could act through the parasympathetic nervous system, also known as the "rest-and-digest" system and the nerve endings carrying its principal drivers, acetylcholine and nitric oxide, which play a crucial role in resolving stress responses. It is also likely that protective factors may work through the "tend-and-befriend" system, acting through oxytocin and endogenous opioid mediators. Both these physiological systems, as well as others, could mediate direct protective effects (i.e., they may not act solely by countering the stress response or its effects) which are salubrious for the mother and/or the fetus. While considering potential biological mechanisms, it is important to determine effects on the mother, placenta, and fetus, and the proportional effect of each on risk of PTB.

While chronic stress can be difficult to ameliorate (especially depending on the driving stressor), it may be worth considering chronic stress reduction/management interventions for women who are trying to conceive and who fall into the higher/highest ranges of pre-pregnancy anxiety and/or chronic stress. In some cases, it is possible that an intervention (e.g., cognitive-behavioral therapy) started before or after conception, could have a clinically meaningful effect in lowering chronic stress levels and related risk of PTB.

Importantly, in our study, general anxiety was associated with a higher, 9-fold increased risk whereas perceived risk of birth complications, which could be related to pregnancy-specific anxiety, predicted a 4-fold increased risk of PTB. It is likely to be informative and useful if future studies could investigate the extent of the contribution of a general high-anxious phenotype versus pregnancy-specific factors that contribute to pregnancy-related anxiety, and the association between the two, because doing so could provide targets for psychosocial (e.g., cognitive-behavioral therapy) and/or pharmacologic intervention. For example, individuals who have low general anxiety, but high pregnancy-related anxiety, may benefit more from interventions designed to ameliorate the effects of pregnancy-specific anxiogenic factors.

Importantly, this study shows that factors such as anxiety and low self-rated health, quantified early during pregnancy, predict risk of PTB. Therefore, an important consideration for future investigations is the intriguing possibility that the mother’s brain senses biological changes that occur early during pregnancy which are associated with PTB, registers these changes, and expresses them as pregnancy-related anxiety and/or low self-rated health. Future studies should test the hypothesis that hormones, cytokines, and other factors released early during pregnancies which are likely to result in PTB, may stimulate abnormal, novel, or anxiogenic sensations which prompt the mother to perceive that her pregnancy has a higher risk of complications and/or to report low self-rated health. If this hypothesis is confirmed, identification of such biological factors could help elucidate potential mechanisms mediating spontaneous PTB and provide targets for early prediction of PTB risk and/or early intervention to prevent PTB.

It is also important to carefully investigate associations and interactions between harmful versus protective factors and risk of PTB in studies involving larger sample sizes. For example, if social support is validated to be an effective buffer against the deleterious effects of chronic stress, interventions could be designed to provide genuine social support for at-risk mothers, especially those who are not receiving support from parents, siblings, and the father of the baby. Such interventional support could be provided at regular meetings with nurses, physicians, and all members of the mother’s healthcare team, professionals such as counselors or social workers who are specifically trained to provide genuine and meaningful social support, and also through support groups.

This study also lays the groundwork for further investigation of the role of psychosocial and stress-related factors in contributing to significant racial disparities in PTB. In 2019, African–American women showed a 50% higher rate of PTB compared with White or Hispanic women. African–American women report racial discrimination as a significant source of chronic social stress. This suggests that in addition to other factors, the association between African–American race and PTB could also be mediated by psychosocial and chronic stress-related factors observed in this study, some of which could be accentuated and exacerbated by potential racial disparities, such as increased exposure to deleterious psychosocial factors and decreased availability of protective factors, which merits further investigation.

**Strengths and Limitations**

Strengths of this study include: (1) The prospective design, that enabled the measurement of psychosocial and chronic stress-related factors during the initial enrollment visit—typically 11 weeks gestation. (2) Introduction of a new instrument, the DQAQ-SPF, and the demonstration that
this questionnaire enables the rapid assessment of several deleterious and protective psychosocial and chronic stress-related factors in a time- and resource-efficient manner. (3) Discovery of associations between chronic stress-related factors and risk of PTB that are in agreement with previously reported findings, as well as novel associations that are in agreement with \textit{a priori} hypotheses. (4) Validation of the DQAQ-SPF, and confirmation of its efficiency, would strengthen its utility for quantifying psychosocial and stress-related factors in the context of studies where psychosocial factors and stress are not the primary focus of investigation, and/or studies that do not have the time, personnel, or resources to quantify these factors using traditional psychometric instruments, each of which generally involves many questions that need to be answered by the participant and scored by study personnel to assess one, or a few specific factors.

Important limitations of this study include: (1) Small sample size with respect to subjects who delivered preterm (the number is in keeping with the proportion observed in the general U.S. population). We aim to replicate and confirm these findings (and hope that other investigators will do the same). We will also replicate these analyses using follow-up data as we accrue more participants. (2) This is the first time that data and findings obtained from the DQAQ-SPF questionnaire are being reported. While this new instrument appears to be useful and effective for quantifying stress-related factors, it remains to be validated. (3) Limitations of self-reported measures such as recall bias, social desirability bias, and differences in the way in which participants understand/perceive the questions asked. (4) Increased chances of encountering Type I statistical errors that result in false positive findings due to multiple statistical comparisons. However, the consistency between our findings and those reported previously, and the fact that all the novel findings confirmed our \textit{a priori} hypotheses, suggest that there is a low probability that the findings reported here are the result of Type I error. (5) Absence of biological measures that could potentially establish cause-effect, mediator, or moderator relationships, and identify underlying mechanisms and targets for intervention. Importantly, biological factors are being quantified as part of the Stanford PRCC study and we aim to incorporate and analyze these factors in the context of the findings presented here. (6) Limitations of the forward selection analysis in which variable selection itself is not embedded in a cross-validation procedure, meaning that: (a) the presented results may not be the only valid or optimal combination of variables; and (b) the corresponding AUC values only hint at, but do not guarantee, increased predictive power. Nevertheless, this approach is well suited for exploring the predictive power of smaller variable subsets and can inform future investigations.

**Conclusion**

The findings presented here are an important step toward identifying psychosocial and chronic stress-related factors, or clusters of factors, that could be assessed quickly and efficiently before or after conception to serve as predictors of PTB risk and perhaps also other adverse pregnancy or health outcomes. Quantifying these factors, before or early in pregnancy, could identify women at risk of delivering preterm, pinpoint mechanisms/targets for intervention, and facilitate development of interventions to prevent PTB. Many of the findings described here are consistent with previous reports, and in the case of novel findings, are consistent with our \textit{a priori} hypotheses that were based on what is known in the literature about the harmful effects of deleterious psychosocial factors and chronic stress on other health outcomes. The logical consistency of our findings, their potential use for predicting adverse outcomes such as PTB, and their potential for impact in terms of identifying biological, psychological, and/or social targets for intervention, suggest that these findings merit further investigation.

**Reproducibility and Data Availability**

The data and source code for reproduction of the results are publicly available at https://nalab.stanford.edu/wp-content/uploads/stress-preterm.zip

**Author Contributions**

All authors have approved the manuscript and declared that no competing interests exist. J.A.M. conducted analyses involving demographic data and relative risk with guidance from G.M.S., wrote related methods sections, and reviewed multiple drafts of the manuscript. M.B. conducted the machine learning-based analyses, wrote related methods and results sections and limited portions of the discussion, all with the guidance and supervision from N.A. M.B. and N.A. reviewed and commented on multiple drafts of the manuscript. N.K.P. assisted with interpretation of initial analyses and reviewed the manuscript. C.C.Q. and A.L. facilitated numerous aspects of the overall project and worked with F.S.D. to integrate the Dhabhar Quick-Assessment Questionnaire for Stress and Psychosocial Factors\textsuperscript{TM} (DQAQ-SPF\textsuperscript{TM}) into the study. L.K. and I.H.G. reviewed and edited a draft of the manuscript. B.G. and M.S.A. commented on and approved the machine learning-based analyses conducted by M.B. and N.A. G.M. S. and D.K.S., obtained funding, and led, supported, and enabled all aspects of the March of Dimes Prematurity Research Center (PRC) at Stanford University. They reviewed multiple drafts of this manuscript. F.S.D. conceptualized and designed the DQAQ-SPF questionnaire, integrated it into the Stanford Prematurity Research Center project, interpreted results, and wrote the majority of this manuscript.

**Funding**

This work was supported by the March of Dimes Prematurity Research Center at Stanford University (D.K.S.) the Bill & Melinda Gates Foundation OPP1189911 (D.K.S.), the Robertson Family Foundation (D.K.S.), the National Institutes of Health R35GM138353 (N.A.), and the Burroughs Wellcome Fund (N.A.).
Conflict of Interest
F.S.D. reports grants from National Institutes of Health (CA107498) and The Office of Naval Research (N000141612096), outside the submitted work. The other authors declare no conflict of interest.

Acknowledgments
This project is part of a larger ongoing study sponsored by the March of Dimes Foundation that provided grant support for the March of Dimes Prematurity Research Center at Stanford University. F.S.D. conceptualized and designed the Dhabhar Quick-Assessment Questionnaire for Stress and Psychosocial Factors™ (DQAQ-SPF™), planned its integration into this project, and wrote the majority of this manuscript. His role in the study was supported by the Carl & Elizabeth Naumann Fund (Stanford University), and by the Sylvester Comprehensive Cancer Center, and the Department of Psychiatry and Behavioral Sciences (Miller School of Medicine, University of Miami). We thank Ronald J. Wong for assistance with this study.

References
19 Accortt EE, Schetter CD. Pregnant women screening positive for depressive symptoms at 24-28 weeks may have increased risk of preterm birth but more precise research is needed. Evid Based Nurs 2014;17(01):11 – 12


