Title

DNA methylation biomarkers prospectively predict both antenatal and postpartum depression

Permalink

https://escholarship.org/uc/item/60v7c5dp

Authors

Payne, Jennifer L
Osborne, Lauren M
Cox, Olivia
et al.

Publication Date

2020-03-01

DOI

10.1016/j.psychres.2019.112711

Peer reviewed

DNA methylation biomarkers prospectively predict both antenatal and postpartum depression

Jennifer L. Payne[a,b], **Lauren M. Osborne**[a,b], **Olivia Cox**[a], **John Kelly**[a], **Samantha Meilman**[a], **Ilenna Jones**[a], **Winston Grenier**[a], **Karen Clark**[a], **Evelyn Ross**[a], **Rachel McGinn**[c], **Pathik D. Wadhwa**[d], **Sonja Entringer**[e,f], **Anne L. Dunlop**[g], **Anna K. Knight**[g], **Alicia K. Smith**[g], **Claudia Buss**[e,f], **Zachary A. Kaminsky**[a,c,h,i,*]

[a]Department of Psychiatry and Behavioral Sciences, Johns Hopkins University School of Medicine, Baltimore, MD, USA

[b]Department of Gynecology and Obstetrics, Johns Hopkins University School of Medicine, Baltimore, MD, USA

[c]The Royal's Institute of Mental Health Research, University of Ottawa, 1145 Carling Avenue, Rm, 6458 Ottawa, Ontario, Canada

[d]Department of Psychiatry & Human Behavior, UC Irvine Genetic Epidemiology Research Institute, University of California, Irvine, CA, USA

[e]Development, Health and Disease Research Program, University of California, Irvine, CA, USA

[f]Medical Psychology Department, Charité University Medicine Berlin, Berlin, Germany

[g]Department of Gynecology and Obstetrics, Emory University School of Medicine, Atlanta, GA

[h]Department of Mental Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA

[i]Department of Cellular and Molecular Medicine, University of Ottawa, Ottawa, Ontario, Canada

Abstract

[*]Corresponding author at: The Royal Ottawa Mental Health Center, Institute of Mental Health Research, University of Ottawa, 1145 Carling Avenue, Rm 6458, Ottawa Ontario Canada. zachary.kaminsky@theroyal.ca (Z.A. Kaminsky).

We sought to replicate and expand upon previous work demonstrating antenatal *TTC9B* and *HP1BP3* gene DNA methylation is prospectively predictive of postpartum depression (PPD) with ~80% accuracy. In a preterm birth study from Emory, Illumina MethylEPIC microarray derived 1st but not 3rd trimester biomarker models predicted 3rd trimester Edinburgh Postnatal Depression Scale (EPDS) scores   13 with an AUC=0.8 (95% CI: 0.63–0.8). Bisulfite pyrosequencing derived biomarker methylation was generated using bisulfite pyrosequencing across all trimesters in a pregnancy cohort at UC Irvine and in 3rd trimester from an independent Johns Hopkins pregnancy cohort. A support vector machine model incorporating 3rd trimester EPDS scores, *TTC9B*, and *HP1BP3* methylation status predicted 4 week to 6 week postpartum EPDS   13 from 3rd trimester blood in the UC Irvine cohort (AUC=0.78, 95% CI: 0.64–0.78) and from the Johns Hopkins cohort (AUC=0.84, 95% CI: 0.72–0.97), both independent of previous psychiatric diagnosis. Technical replicate predictions in a subset of the Johns Hopkins cohort exhibited strong cross experiment correlation. This study confirms the PPD prediction model has the potential to be developed into a clinical tool enabling the identification of pregnant women at future risk of PPD who may benefit from clinical intervention.

Keywords

1. Introduction

The field of psychiatry does not have many consistently validated biomarkers that enable the prediction of future risk for mental illness. This is due, in part, to the fact that it is difficult to prospectively obtain biological samples on subjects who may or may not develop mental illness and to follow these subjects for the months or years that it will take for the illness to develop. One exception to this conundrum is postpartum depression (PPD), which represents one of the few cases in psychiatry where we know both when to study women at risk as well as when they are likely to develop the illness. PPD develops after the gonadal hormone drop that accompanies childbirth in 10–20% of women without a psychiatric history (Josefsson et al., 2001; Miller, 2002; Pearlstein et al., 2009) and has significant adverse effects on both mother and child (Breese McCoy, 2011; Cuijpers et al., 2008; Field, 2011; Hirst and Moutier, 2010; O'Hara, 2009; Soufia et al., 2010). Further, PPD afflicts some populations at even higher rates, for example, 30% of women with a history of depression and 52% of women with bipolar disorder (Viguera et al., 2011).

By using prospective monitoring of mood during pregnancy and in the postpartum period we have previously identified two epigenetically modified biomarker genes with the ability to prognosticate future postpartum mood episodes. We originally generated estradiol associated epigenomic profiles in the mouse hippocampus and compared these data with peripheral blood derived DNA methylomes generated from pregnant women with pre-existing mood disorders who were at high-risk of developing PPD. Bioinformatic analysis subsequently identified a panel of epigenetic biomarkers in the *TTC9B* and *HP1BP3* genes(Guintivano et al., 2014).

Our initial work focused on the generation of a PPD predictive model that has the potential to generate a clinically efficacious tool. Using epigenetic variation in these genes, we generated a predictive model that was prospectively predictive of PPD with an area under the receiver operator characteristic curve (AUC) of 82% (Guintivano et al., 2014). Our initial findings were that the model generated PPD probability was higher for cases relative to controls if women who were antenatally euthymic while the reverse was true if women were antenatally depressed. Incorporating a proxy for cellular heterogeneity into our models corrected for the problem and generated a consistently higher model prediction score in women with PPD relative to controls, independent of antenatal depression status.

The predictive efficacy of this model was subsequently replicated in two additional cohorts including a prospective gene expression cohort of women with pre-existing mood disorders (Mehta et al., 2014) and a cross sectional study of women from the Franconian Maternal Health Evaluation Studies (FRAMES) study who had no mood disorder history (Mehta et al., 2012; Osborne et al., 2016). While the data provided promising evidence that the signal at *TTC9B* and *HP1BP3* was replicable in PPD, each replication had limitations. In the first cohort, gene expression values for *TTC9B* and *HP1BP3* were used and were limited to women with a prior mood disorder diagnosis, while in the second, blood was sampled long after pregnancy and as such, the results did not inform the efficacy of the development of a potential prospectively predictive blood test for PPD. Furthermore, more sophisticated machine learning modeling approaches like support vector machines (SVM) were not employed, and these represent an added potential to generate robust models that consistently predict PPD. In addition, questions still remain that limit our ability to translate this finding into a usable blood test including determining the ideal antenatal trimester of blood sampling to enable consistent predictions and what the effect of the model is, if any, on predicting antenatal depression status.

The goal of this study was therefore to address the limitations and unanswered questions of our previous work by evaluating the predictive efficacy of our model in alternative cohorts both with and without a prior psychiatric history at multiple time points during pregnancy. The first objective was to assess PPD model performance for predicting antenatal depression scores. The second objective was to evaluate our original model as well as newly generated machine learning models to predict PPD scores in women with and without a previous psychiatric history using blood taken at multiple time points during pregnancy. The third objective was to validate any newly generated models in a second independent cohort. The forth objective was to assess the ability of the biomarker model to generate consistent predictions across a range of varying experimental conditions.

2. Methods

2.1. Human samples

Subjects derived from four prospectively collected cohorts assessing mood symptoms during pregnancy and postpartum. The first was the Johns Hopkins Prospective PPD sample previously described by our group (Guintivano et al., 2014). The second was a prospective preterm birth study of pregnant women at Emory University for which only 1st and 3rd trimester DNA methylation and EPDS scores were available (Knight et al., 2018). The third

was a UC Irvine sample from which 1st, 2nd, and 3rd trimester DNA methylation. This cohort also contained DNA from 3 months postpartum in a subset of women, but as the 1 month postpartum EPDS scores were the focus of this analysis, this data was not evaluated in this study. The fourth was a novel cohort of samples collected at Johns Hopkins using the same protocol as our originally published Johns Hopkins Prospective PPD cohort. This cohort was named the Johns Hopkins Neuroimaging Cohort from which 3rd trimester blood, EPDS scores, and postpartum EPDS scores was available. Women from both the UC Irvine and Johns Hopkins Neuroimaging cohorts consisted of women with and without a previous psychiatric diagnosis. Detailed information on study subjects is available in Table 1 and Methods S1.

2.2. Analysis plan

Due to limitations in the consistency of data collected across cohorts, we analyzed data according to the following plan. The Edinburgh Postnatal Depression Scale (EPDS) was administered consistently across all cohorts and represented the primary outcome assessed in this study. The Johns Hopkins Prospective PPD cohort represents our originally published cohort (Guintivano et al., 2014) and was used as a training set for statistical model generation. As only antenatal mood metrics were available in the Emory Cohort, this dataset was used to assess for the predictive efficacy of biomarker models with antenatal mood status. The UC Irvine cohort was the most comprehensive dataset consisting of biological data across all antenatal trimesters and consisted of women with and without a previous psychiatric history. This cohort was used as the primary dataset to assess for the predictive efficacy of postpartum mood symptoms by various statistical modeling approaches including linear discriminate analysis and support vector machine (SVM) based prediction across trimesters and as a function of self reported previous psychiatric history. Antenatal mood information was not available in this cohort. The Johns Hopkins Neuroimaging cohort represents a prospective PPD cohort with ongoing subject collection. Biological data from 1st and 2nd trimesters was limited so analyses were performed on 3rd trimester data primarily as a replication of findings from the UC Irvine cohort. As we had access to the most biological material from this cohort, a subset of these samples were used to assess the technical replicability of the PPD prediction model to assess postpartum EPDS scores.

2.3. Sodium bisulfite pyrosequencing

Bisulfite conversion was carried out using EZ DNA Methylation Gold Kit (Zymo Research, Irvine, CA) according to the manufacturer's instructions. Nested PCR amplifications were performed with a standard PCR protocol in 25 ul volume reactions containing 3–4 μl of sodium-bisulfite-treated DNA, 0.2 uM primers, and master mix containing Taq DNA polymerase (Sigma Aldrich, St. Louis, MO). Primer sequences can be found in Table S1. PCR amplicons were processed for pyrosequencing analysis according to the manufacturer's standard protocol (QIAGEN, Germantown, MD) using a PyroMark MD system (QIAGEN) with Pyro Q-CpG 1.0.9 software (QIAGEN) for CpG methylation quantification. Sample processing order was randomized by sorting on a random number generated in Perl. Laboratory personnel were blind to mood status until the completion of data generation.

2.4. Statistical analysis

All statistical tests were performed in R (http://www.r-project.org/). Using an Anderson-Darling test from the nortest package, all distributions of data that rejected the null hypothesis of normality were subsequently evaluated with non-parametric tests. All statistical tests performed were two tailed and a $p < 0.05$ was considered significant. Unless otherwise specified ± denotes the standard error of the mean.

EPDS prediction for the antenatal time point in the Emory cohort was achieved using the previously published model (Guintivano et al., 2014):

$$E(T_z EPDS_i) = a + bHP1BP3 + cCelltype_j + d(HP1BP3xCelltype_j) + eTTC9B$$

Where EPDS for individual (i) at time point (T_z) was modeled as a function of an interaction of *HP1BP3* DNA methylation with cell type variable (j), where j is the ratio of monocyte to non-monocyte counts (See Method S2 for relevant rationale), controlling for additive covariate *TTC9B* DNA methylation. SVM models were built using the e1071 package in Bioconductor. All model training was performed in the Johns Hopkins Prospective cohort and models applied independently using the 'predict' method in R. Receiver operator characteristic curves were generated using the pROC package in Bioconductor. Algorithm predictive accuracy was assessed serially for each possible EPDS cut off, where by at a given cut off value, women greater than or equal to the cut off threshold are deemed a case, while those below are denoted a control. Previous studies suggest an EPDS value of 13 has the best psychometric properties to represent perinatal depression out of possible EPDS scores (Ji et al., 2011).

3. Results

3.1. PPD biomarkers taken in 1st trimester predict 3rd trimester antenatal depression

We evaluated the efficacy of our originally published linear model to predict antenatal EPDS scores in both the 1st trimester (T1) and 3rd trimester (T3) time points in the Emory cohort. The predictive model at T1 and T3 was not strongly predictive of subjects with concurrent EPDS scores 13 at their respective time points (T1 AUC= 0.67, 95% CI: 0.55–0.75, T3 AUC= 0.56, 95% CI: 0.37–0.56) and did not distinguish these subjects from those with EPDS < 13 (Fig. 1ab); however, the T1 biomarker model predicted the T3 EPDS scores 13 with a higher accuracy (AUC= 0.8, 95% CI: 0.63–0.8) that improved as a function of increasing depression severity ($R = 0.92$, $p = 1.4 \times 10^{-8}$) and clearly distinguished this group from those with EPDS < 13 (Fig. 1c). The maximum predictive accuracy was obtained for predicting women with T3 EPDS scores 20 (AUC= 0.89, 95% CI: 0.79–0.89) (Fig. 1c).

3.2. PPD biomarker model predicts postpartum EPDS scores in women with and without a previous psychiatric diagnosis

Our goal was not only to assess the efficacy of the original Linear Discriminate Analysis (LDA) model using antenatal biological measures in a general population sample without a previous psychiatric diagnosis but also to identify the trimester of optimal efficacy for biomarker model testing, which was only possible given the cohort design in the UCI cohort.

In this analysis, we assessed the ability of the LDA model developed from our original study (Guintivano et al., 2014) to predict elevated EPDS scores in women with no psychiatric history during the postpartum period in the UC Irvine cohort as this model did not require CBC information, which was not available for this cohort. The LDA model predicted cases with EPDS   13 with accuracies above 70% only at the T3 time point (AUC= 0.72, 95% CI: 0.6–0.72) (Fig. 2a). Furthermore, the T3 biomarker scores generated increasing predictive accuracies with increasing depression severity (T1: Rho= −0.63, $p = 0.096$, T2: Rho= −0.43, $p = 0.283$, T3: Rho= 0.86, $p = 3.6 \times 10^{-4}$).

In an attempt to improve the sensitivity of the prediction, we used the support vector machine (SVM) machine learning algorithm to retrain the data from the original Johns Hopkins Prospective PPD cohort. Using "leave one out cross validation," the new SVM model predicted PPD diagnostic status in the original cohort with an AUC of 0.83 (95% CI: 0.71–0.83). In the UC Irvine cohort using only data from women without a prior psychiatric history, model performance was similar to the LDA model across T1, T2, and T3 time points (Fig. 2b), with T3 biomarker scores predicting EPDS scores   13 with an AUC of 0.72 (95% CI: 0.56–0.72). Only T3 exhibited an increasing predictive accuracy with increasing depression severity (T1: Rho= 0.34, $p = 0.414$, T2: Rho= −0.55, $p = 0.154$, T3: Rho= 0.86, $p = 3.4 \times 10^{-4}$).

The LDA model performed poorly to distinguish cases with EPDS   13 (Fig. 2c) among UC Irvine cohort women with a prior PPD history and there was no correlation between prediction accuracy and increasing depression severity (T1: $N = 17$, Rho= −0.38, $p = 0.313$, T2: $N = 19$, Rho= −0.53, $p = 0.141$, T3: $N = 18$, Rho= −0.72, $p = 0.029$). The SVM model performed more favorably, exhibiting significant correlations of T1 and T3 prediction accuracies with increasing depression severity (T1: Rho= 0.68, $p = 0.042$, T2: Rho= 0.09, $p = 0.812$, T3: Rho= 0.7, $p = 0.035$), and generating a maximum accuracy for predicting EPDS   13 in the T1 samples (AUC= 0.78, 95% CI: 0.55–0.78) (Fig. 2d). Given the small sample size for this sub-cohort, the robustness of this finding is not high; however, the most important observation is that the direction of prediction for women with a previous PPD history was negative such that cases had a lower predicted PPD probability than controls (Fig. 2e). This is the opposite to that observed in the women without a psychiatric history, where predictive accuracies were similar (Fig. 2b) but cases had a higher predicted PPD probability than controls (Fig. 2f). When we attempted to predict the entire cohort together, no relationship between predictive accuracy and increasing depression severity was observed (T1: Rho= 0.62, $p = 0.074$, T2: Rho= 0.48, $p = 0.159$, T3: Rho= 0.52, $p = 0.081$). This observation is similar to that of our originally published study (Guintivano et al., 2014) where the model predicted PPD probability of women who were depressed antenatally was lower in women who remained depressed postpartum compared to those who became well (Figure S3). Conversely, in women euthymic during pregnancy, the model output was higher in the subgroup who became depressed in the postpartum period relative to those who remained well.

From a practical standpoint, in order to generate a model with the potential to be efficacious in a clinical environment, we trained a new SVM model on the Johns Hopkins Prospective PPD cohort data incorporating antenatal depression status as an interaction covariate in the

model. We applied this to the combined UC Irvine cohort, inputting previous history of PPD as the interacting covariate and observed significant correlations of predictive accuracy with increasing depression severity (T1: Rho= 0.94, *p* = 0, T2: Rho= 0.91, *p* = 0, T3: Rho= 0.57, *p* = 0.013) with T3 samples achieving an AUC of 0.77 (95% CI: 0.64–0.77) to detect an EPDS   13 (Fig. 2g). Importantly, the direction of prediction across all trimesters was such that women who became depressed postpartum had a higher model generated predicted PPD probability on average compared to those who did not become depressed (Fig. 2h). In the UC Irvine cohort, we did not have antenatal EPDS scores by which to directly assess antenatal mood; however, in light of our findings above that 1st trimester biomarker values predict antenatal depression status, we attempted to use our model for predicting antenatal mood status to generate such a measure (Method S2). To that end, we input 1st trimester DNA methylation levels for *HP1BP3, TTC9B,* and *SLC19A1* (used as a DNA methylation proxy of the ratio of monocytes to non-monocytes: see Result S1: Figure S1) into our originally published linear model to generate a best guess of 3rd trimester antenatal mood in the UC Irvine cohort. Incorporating this metric into the entire UC Irvine cohort as the interacting covariate in our SVM based PPD prediction model, we successfully identified women with an EPDS   13 in the postpartum period (AUC= 0.78, 95% CI: 0.64–0.78) (Fig. 3a). Together, these data suggest that the evaluation of T1 PPD biomarker loci in conjunction with T3 biomarker loci have the capability to predict both antenatal and postpartum depression.

3.3. Replication of novel SVM model for prospective PPD prediction

We sought to replicate the observation that our model predicts PPD in opposing directions as a function of antenatal mood status and that our newly trained SVM model correctly predicts postpartum mood status. To that end, we performed an analysis on available data at the time of this study in the Johns Hopkins Neuroimaging cohort and assessed the directionality and predictive accuracy of our model. In this cohort, EPDS scores were taken prospectively at the T3 time point and at 2, 6, and 12, 20, and 36 weeks postpartum. At the time of this analysis, there were *N* = 68 subjects with T3 biomarker data and 20 week postpartum data, from which *N* = 8 exhibited an average EPDS score from the 6 to 20 weeks   13 indicating a postpartum depressive state. Importantly, the T3 EPDS score in all 7 of the 8 subjects was 13 suggesting antenatal depression. Under the original two gene model without accounting for antenatal mood status, we would expect this to result in a negative prediction direction. As expected, the model was predictive of EPDS   13 in the postpartum period with cases exhibiting a lower model predicted PPD probability relative to controls (mean cases: 0.031 ± 0.16, mean controls: 0.37 ± 0.32 and generating an AUC of 0.87 (95% CI: 0.70–1). Application of the novel SVM model incorporating antenatal depression status generated in the above section resulted in an improved predictive accuracy and corrected the prediction direction (mean cases: 0.71 ± 0.22, mean controls: 0.33 ± 0.29, AUC= 0.84, 95% CI: 0.72–0.97) (Fig. 3b).

3.4. Replicability of PPD biomarkers

For a biomarker panel to be clinically useful, it must generate reproducible outputs under varying conditions to satisfy regulatory metrics of accuracy and precision. For this reason, we assessed *TTC9B* and *HP1BP3* DNA methylation and model prediction outputs in *N* = 20

samples from the Johns Hopkins Neuroimaging cohort, with three technical replicates per subject where all laboratory procedures were performed on different days by different operators in the laboratory. We observed significant agreement across *TTC9B* replicates (Rep1 vs. Rep2: Rho= 0.92, $p = 4.20 \times 10^{-6}$; Rep1 vs. Rep3: Rho= 0.91, $p = 2.81 \times 10^{-6}$; Rep2 vs. Rep3: Rho= 0.95, $p = 6.27 \times 10^{-6}$), while more variation was observed across *HP1BP3* technical replicates (Rep1 vs. Rep2: Rho= 0.28, $p = 0.24$; Rep1 vs. Rep3: Rho= 0.04, $p = 0.86$; Rep2 vs. Rep3: Rho= 0.1, $p = 0.66$)(Figure S4). Importantly, the models consistently identified individuals with high EPDS scores in the postpartum period and the pattern of model predictive accuracy as a function of increasing postpartum depression severity was highly correlated (Rep1 vs. Rep2: Rho= 0.9, $p = 3.81 \times 10^{-6}$; Rep1 vs. Rep3: Rho= 0.9, $p = 5.85 \times 10^{-6}$; Rep2 vs. Rep3: Rho= 1, $p = 5.83 \times 10^{-19}$) (Fig. 3c).The correlation of an SVM model built solely on *TTC9B* DNA methylation or *HP1BP3* DNA methylation alone did not generate a pattern of predictive accuracy as a function of increasing depression severity that was as consistent as the SVM model containing both *HP1BP3* and *TTC9B* (*HP1BP3*: Rep1 vs. Rep2: Rho= 0.33, $p = 0.22$; Rep1 vs. Rep3: Rho= 0.7, $p = 0.0040$; Rep2 vs. Rep3: Rho= 0.67, $p = 0.0061$; *TTC9B*: Rep1 vs. Rep2: Rho= 0.61, $p = 0.016$; Rep1 vs. Rep3: Rho= 0.72, $p = 0.0025$; Rep2 vs. Rep3: Rho= 0.89, $p = 9.28 \times 10^{-6}$). These results suggest that despite a higher degree of technical noise in the *HP1BP3* signal, the measured DNA methylation levels relative to that of *TTC9B* values derived from the same sample confer additional information of importance for consistent model prediction.

4.   Discussion

We sought to replicate our previously identified panel of PPD predictive biomarkers in several independent cohorts of pregnant women both with and without a prior psychiatric history using blood taken from either 1st, 2nd, or 3rd trimester. Similar to previous efforts to replicate our model in the FRAMES cohort(Osborne et al., 2016), the model relying solely on *TTC9B* and *HP1BP3* generated a reasonably strong predictive accuracy above 70% for detecting EPDS cases   13 in women without a psychiatric history but not in women with a prior PPD. Furthermore, the predictive accuracy was strongest using 3rd trimester blood derived DNA methylation as compared to 1st or 2nd trimester samples. Importantly, an SVM model trained on the training set data from our original publication (Guintivano et al., 2014) demonstrated a similar predictive efficacy in women both with and without a prior psychiatric history. Importantly and similar to our originally published studies, while subjects with high EPDS in the postpartum period are segregated from low EPDS individuals with ~80% accuracy, whether the average model output is higher or lower than that of the low postpartum EPDS subjects was contingent on antenatal depression status. In essence, the direction of prediction was opposite between antenatally depressed vs. antenatally euthymic subjects. In the past, we used a metric of cellular heterogeneity, modeling the ratio of monocytes to non-monocytes, to correct for this discrepancy (Guintivano et al., 2014; Osborne et al., 2016); however, our attempts to use a DNA methylation proxy of this ratio did not affect the direction of prediction per group (data not shown), either in the UC Irvine cohort or the Johns Hopkins Neuroimaging cohort. One possible explanation for this, as acknowledged in Guintivano et al., is that our original

metrics for cellular heterogeneity in our first publication may have been biased by batch effects that tracked with antenatal depression status (Guintivano et al., 2014) and that cellular heterogeneity metrics are not a good representative measure of antenatal depression status. Conversely, incorporating the estimated monocyte to non-monocyte ratio in the Emory cohort was necessary to accurately model third trimester EPDS scores. Furthermore, incorporation of the monocyte to non-monocyte proxy locus in the UC Irvine cohort using 1st trimester derived DNA methylation was used to generate an estimate of antenatal depression status for this cohort, which, when fed into the novel SVM model, accurately predicted postpartum EPDS scores in the UC Irvine cohort. Together, the data suggest that while the ratio of monocytes to non-monocytes may aid in the prediction of antenatal depression status, this relationship requires more study. Ultimately antenatal depression status is more important than cellular heterogeneity metrics for generating accurate PPD predictions. This is likely related to the timing of blood sampling relative to the outcome being modeled. Notably, in the Emory cohort, 1st trimester but not 3rd trimester PPD biomarker data was most accurately able to predict 3rd trimester EPDS scores   13, while prediction of postpartum EPDS scores   13 appears to be most consistently predicted by 3rd trimester PPD biomarkers. In this way, the PPD biomarkers appear to accurately predict the future but not the present depressive state. This may be related, in part, to the hypothesized role of *TTC9B* in regulating estrogen signaling, which in turn alters DNA methylation of these loci in a more extreme manner in women at risk(Guintivano et al., 2014; Kimmel et al., 2016; Osborne et al., 2016). Under a scenario of fluctuating hormone levels as occurs during pregnancy and then at birth, these effects would amplify any gonadal hormone induced differences in PPD biomarker DNA methylation, allowing them to distinguish PPD risk and non-risk groups, as well as to result in altered downstream changes to mood and anxiety associated with altered estrogen signaling. Subsequent ongoing studies evaluating changes in PPD biomarkers and serum hormone levels will allow us to address these hypotheses in the future.

Given the clinical nature of our current studies, a majority of women with a mood disorder diagnosis are being actively treated with antidepressant medications. As our model predicts who of these women will remain depressed or become well in the postpartum period, our data suggests the intriguing possibility that our PPD biomarkers may be an indication of responsiveness to treatment. In mice, an interaction of estrogen receptors and antidepressant treatment action has been indicated in some studies. Female mice given low doses of ketamine showed enhanced antidepressant responses while in proestrus or while being treated with ERα or ERβ agonists (Dossat et al., 2018). In another study, the antidepressant effect of duloxetine was diminished in female but not male estrogen deficient transgenic mice (Xu et al., 2017), resulting in regional changes in serotonin and dopamine and suggesting that estrogen is important for antidepressant therapeutic actions. *TTC9B* epigenetic variation has been suggested to mediate estrogen signaling, due in part to the demonstrated role of its close homologue, TTC9A, in mediating estrogen receptor alpha (ERα) signaling (Shrestha et al., 2015), as well as our prior observations that *TTC9B* DNA methylation is associated with changes in 17β-estradiol (E2) levels during pregnancy in a longitudinal design (Osborne et al., 2016). Furthermore, deletion of *TTC9A* in the context of estradiol administration results in measurable differences in rodent anxiety behavior as well

as serotonergic changes in the dorsal raphe nucleus (Lim et al., 2016). *TTC9A* and *B* are highly conserved and share all regulatory protein domains (Shrestha et al., 2015), suggesting they may have functional similarities such that TTC9B may contribute to the sensitivity of estrogen signaling. Thus, our model may not only be providing useful information for who is at risk for either developing or continuing depression in the postpartum period, but may indicate who is likely to respond to treatment. Future studies should investigate these effects in different classes of medication to assess the efficacy of our model in this domain.

Importantly, our results demonstrated that prediction of PPD across multiple technical replicates is highly consistent, despite being performed by different people in the laboratory and on different days. This is important, as technical variation affecting model output consistency would yield even the most promising statistical findings useless as a clinical tool. We observed that despite a strong cross experiment consistency for *TTC9B* and a relatively weak one for *HP1BP3*, the model predictive accuracy was more consistent when using the model incorporating two genes as opposed to one. This may be due to the fact that the true strength of model efficacy comes from the evaluation of each biomarker gene in the context of the other. Alone, each biomarker locus demonstrates some predictive efficacy to model PPD, but the accuracy of model prediction is lower (Fig. 4). Despite exhibiting a weaker predictive accuracy alone, genetic knock-out experiments of HP1BP3 generate a compelling model of altered maternal behavior mimicking elements of PPD and warrant its inclusion in the model. Recent evidence has suggested a role for HP1BP3 in mediating anxiety in female mice: the offspring of *HP1BP3* Knock-Out mice have significantly lower survival rates due to a deficit in maternal care that could be reversed upon cross fostering (Garfinkel et al., 2016). Cumulatively, these results suggest that together, *HP1BP3* and *TTC9B*, both of which exhibit interesting biological characteristics of potential etiological relevance to the disease, both more accurately and consistently model PPD together than apart.

This work addresses a number of the limitations of our previous studies including the optimal timing of the prospective blood draw, prediction of antenatal depression as well as evaluating women in the context of or lack of a prior psychiatric history. The results from this work demonstrate that this epigenetic PPD biomarker model is highly accurate in both women with and without a psychiatric history. Future studies are underway to find additional orthogonal sources of biological variation that will improve predictive accuracies beyond 80%. In its current state, the PPD biomarker model is technically replicable, displaying a high degree of analytical specificity and precision, suggesting that these findings could be developed into a diagnostic blood test to be administered during the standard blood screening that occurs at the beginning of the 3rd trimester as the standard of obstetric care. The importance of identifying patients at risk of PPD during pregnancy is reinforced by the January 26, 2016 guideline from the United States Preventive Services Task Force (USPSTF) recommending depression screenings for pregnant and post-partum women for the first time. These guidelines are expected to galvanize more health providers to provide screening and tangible patient care for identification of the more than 1 in 5 women who suffer from this disease and support longstanding recommendations by the American Psychiatric Association that encourage pediatric practices to create a system to identify "at risk" mothers.

Future studies should evaluate the added value of prospective identification of women at risk as compared to identification post-partum once the illness takes hold. Women identified with depression symptoms during the postpartum period can face a delay of weeks to months to be treated, which is problematic as it results in exposure of the offspring to the deleterious consequences of depression as well as higher health care usage by the depressed mothers(Dagher et al., 2012). As such, a PPD blood test given during pregnancy would allow for the possibility of intervention activities as well as the advanced preparation of psychological and psychiatric resources in the postpartum period, most likely resulting in a reduced latency to treatment and ultimately remission or even prevention, thus reducing the burden on both the healthcare system as well as the next generation. Future studies should carefully evaluate the potential clinical utility of these biomarkers and their potential for translation into a clinical tool.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

References

Breese McCoy SJ, 2011 Postpartum depression: an essential overview for the practitioner. South Med. J. 104 (2), 128–132. [PubMed: 21206417]

Cuijpers P, Brannmark JG, van Straten A, 2008 Psychological treatment of postpartum depression: a meta-analysis. J. Clin. Psychol. 64 (1), 103–118. [PubMed: 18161036]

Dagher RK, McGovern PM, Dowd BE, Gjerdingen DK, 2012 Postpartum depression and health services expenditures among employed women. J. Occup. Environ. Med. 54 (2), 210–215. [PubMed: 22267187]

Dossat AM, Wright KN, Strong CE, Kabbaj M, 2018 Behavioral and biochemical sensitivity to low doses of ketamine: influence of estrous cycle in C57BL/6 mice. Neuropharmacology 130, 30–41. [PubMed: 29175352]

Field T, 2011 Prenatal depression effects on early development: a review. Infant Behav. Dev. 34 (1), 1–14. [PubMed: 20970195]

Garfinkel BP, Arad S, Neuner SM, Netser S, Wagner S, Kaczorowski CC, Rosen CJ, Gal M, Soreq H, Orly J, 2016 HP1BP3 expression determines maternal behavior and offspring survival. Genes Brain Behav. 15 (7), 678–688. [PubMed: 27470444]

Guintivano J, Arad M, Gould TD, Payne JL, Kaminsky ZA, 2014 Antenatal prediction of postpartum depression with blood DNA methylation biomarkers. Mol. Psychiatry 19 (5), 560–567. [PubMed: 23689534]

Hirst KP, Moutier CY, 2010 Postpartum major depression. Am. Fam. Phys. 82 (8), 926–933.

Ji S, Long Q, Newport DJ, Na H, Knight B, Zach EB, Morris NJ, Kutner M, Stowe ZN, 2011 Validity of depression rating scales during pregnancy and the postpartum period: impact of trimester and parity. J. Psychiatr. Res. 45 (2), 213–219. [PubMed: 20542520]

Josefsson A, Berg G, Nordin C, Sydsjo G, 2001 Prevalence of depressive symptoms in late pregnancy and postpartum. Acta Obstet. Gynecol. Scand. 80 (3), 251–255. [PubMed: 11207491]

Kimmel M, Clive M, Gispen F, Guintivano J, Brown T, Cox O, Beckmann MW, Kornhuber J, Fasching PA, Osborne LM, Binder E, Payne JL, Kaminsky Z, 2016 Oxytocin receptor DNA methylation in postpartum depression. Psychoneuroendocrinology 69, 150–160. [PubMed: 27108164]

Knight AK, Conneely KN, Kilaru V, Cobb D, Payne JL, Meilman S, Corwin EJ, Kaminsky ZA, Dunlop AL, Smith AK, 2018 SLC9B1 methylation predicts fetal intolerance of labor. Epigenetics 13 (1), 33–39. [PubMed: 29235940]

Lim LW, Shrestha S, Or YZ, Tan SZ, Chung HH, Sun Y, Lim CL, Khairuddin S, Lufkin T, Lin VC, 2016 Tetratricopeptide repeat domain 9A modulates anxietylike behavior in female mice. Sci. Rep. 6, 37568. [PubMed: 27869229]

Mehta D, Newport DJ, Frishman G, Kraus L, Rex-Haffner M, Ritchie JC, Lori A, Knight BT, Stagnaro E, Ruepp A, Stowe ZN, Binder EB, 2014 Early predictive biomarkers for postpartum depression point to a role for estrogen receptor signaling. Psychol. Med. 44 (11), 2309–2322 8. [PubMed: 24495551]

Mehta D, Quast C, Fasching PA, Seifert A, Voigt F, Beckmann MW, Faschingbauer F, Burger P, Ekici AB, Kornhuber J, Binder EB, Goecke TW, 2012 The 5-HTTLPR polymorphism modulates the influence on environmental stressors on peripartum depression symptoms. J. Affect. Disord. 136 (3), 1192–1197. [PubMed: 22209125]

Miller LJ, 2002 Postpartum depression. JAMA 287 (6), 762–765. [PubMed: 11851544]

O'Hara MW, 2009 Postpartum depression: what we know. J. Clin. Psychol. 65 (12), 1258–1269. [PubMed: 19827112]

Osborne L, Clive M, Kimmel M, Gispen F, Guintivano J, Brown T, Cox O, Judy J, Meilman S, Braier A, Beckmann MW, Kornhuber J, Fasching PA, Goes F, Payne JL, Binder EB, Kaminsky Z, 2016 Replication of epigenetic postpartum depression biomarkers and variation with hormone levels. Neuropsychopharmacology 41 (6), 1648–1658. [PubMed: 26503311]

Pearlstein T, Howard M, Salisbury A, Zlotnick C, 2009 Postpartum depression. Am. J. Obstet. Gynecol. 200 (4), 357–364. [PubMed: 19318144]

Shrestha S, Sun Y, Lufkin T, Kraus P, Or Y, Garcia YA, Guy N, Ramos P, Cox MB, Tay F, Lin VC, 2015 Tetratricopeptide repeat domain 9A negatively regulates estrogen receptor alpha activity. Int. J. Biol. Sci. 11 (4), 434–447.

Soufia M, Aoun J, Gorsane MA, Krebs MO, 2010 [SSRIs and pregnancy: a review of the literature]. Encephale 36 (6), 513–516. [PubMed: 21130237]

Viguera AC, Tondo L, Koukopoulos AE, Reginaldi D, Lepri B, Baldessarini RJ, 2011 Episodes of mood disorders in 2,252 pregnancies and postpartum periods. Am. J. Psychiatry 168 (11), 1179–1185. [PubMed: 21799064]

Xu Y, Ma L, Jiang W, Li Y, Wang G, Li R, 2017 Study of sex differences in duloxetine efficacy for depression in transgenic mouse models. Front. Cell Neurosci. 11, 344. [PubMed: 29163055]
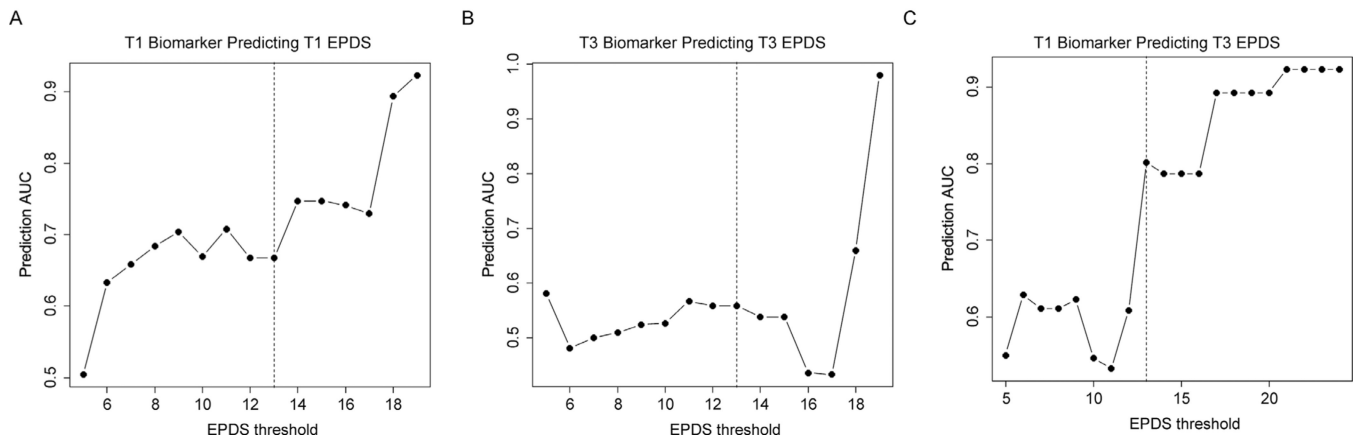
Fig. 1. Biomarker model based prediction of antenatal EPDS scores.
A plot of the AUC of model prediction (y axis) as a function of EPDS cut off score (x axis) for 1st trimester predicting 1st trimester EPDS scores (A), 3rd trimester biomarker scores predicting 3rd trimester EPDS scores (B), and 1st trimester biomarker scores predicting 3rd trimester EPDS scores (C). Vertical black lines represent an EPDS of 13, representing the cut off for depression.
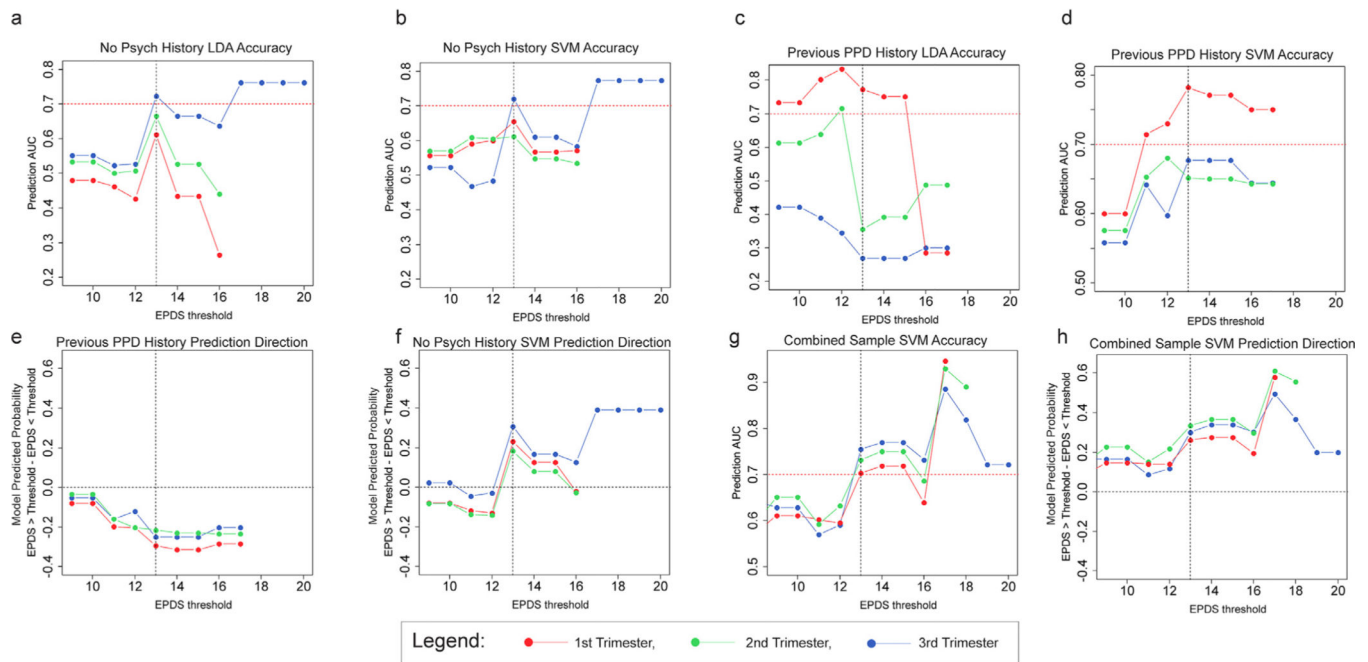
Fig. 2. Biomarker model performance as a function of antenatal time point and psychiatric diagnosis.

a.)Plot of EPDS threshold values (x axis) as a function of the AUC of prediction for the number of women above that threshold (y axis) for the originally published LDA model in women without a psychiatric history. b.)Plot of EPDS threshold values (x axis) as a function of the AUC of prediction for the number of women above that threshold (y axis) for the SVM model in women without a psychiatric history. c.) Plot of EPDS threshold values (x axis) as a function of the mean model output (predicted probability) for women above the EPDS threshold minus that for women below the threshold (y axis) for the originally published LDA model in women without a psychiatric history. d.)Plot of EPDS threshold values (x axis) as a function of the AUC of prediction for the number of women above that threshold (y axis) for the SVM model in women with a previous history of PPD. e.) Plot of EPDS threshold values (x axis) as a function of the mean model output for women above the EPDS threshold minus that for women below the threshold (y axis) for the SVM model in women with a previous history of PPD. f.)Plot of EPDS threshold values (x axis) as a function of the mean model output for women above the EPDS threshold minus that for women below the threshold (y axis) for the SVM model in women without a psychiatric history. g.) Plot of EPDS threshold values (x axis) as a function of the AUC of prediction for the number of women above that threshold (y axis) for the novelly SVM model accounting for antenatal depression status in a combined sample of women with and without a previous history of PPD. h.)Plot of EPDS threshold values (x axis) as a function of the mean model output for women above the EPDS threshold minus that for women below the threshold (y axis) for the novelly SVM model accounting for antenatal depression status in a combined sample of women with and without a previous history of PPD. Horizontal dashed red lines denote an AUC of 70% while a dashed vertical black line denotes an EPDS of 13, signifying likely PPD.
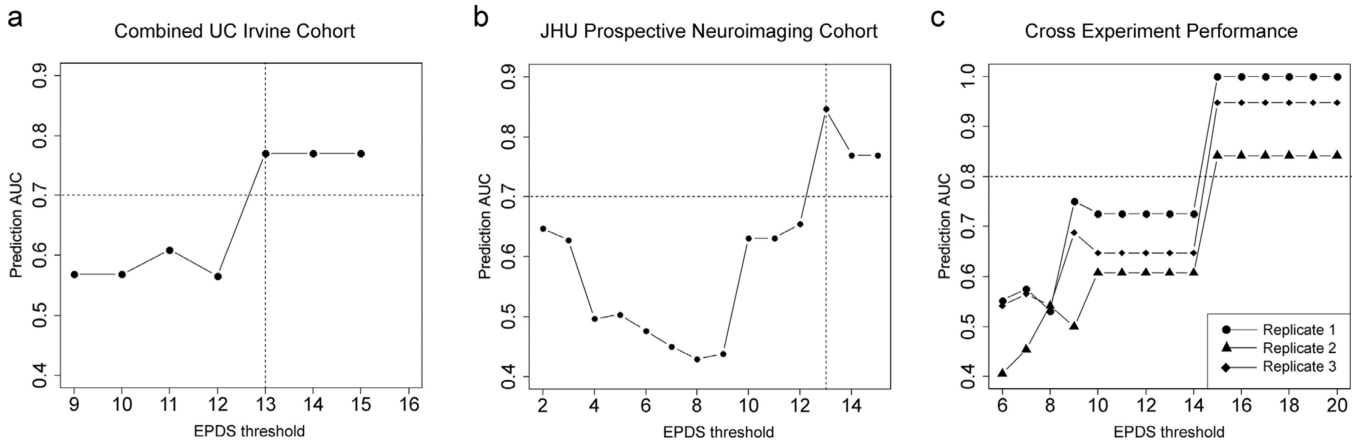
Fig. 3. Biomarker model performance incorporating antental depression.
a.)Plot of EPDS threshold values (x axis) as a function of the AUC of prediction for the
number of women above that threshold (y axis) for the SVM model accounting for antenatal
depression status in women from the UC Irvine cohort where antenatal depression status is
determined with T1 time point biomarker output. b.)Plot of EPDS threshold values (x axis)
as a function of the AUC of prediction for the number of women above that threshold (y
axis) for the SVM model accounting for antenatal depression status in the JHU Prospective
Neuroimaging cohort. Horizontal dashed red lines denote an AUC of 70% while a dashed
vertical black line denotes an EPDS of   13, signifying likely PPD. c.) A plot of EPDS
threshold values (x axis) as a function of the AUC of prediction for the number of women
above that threshold (y axis) for an SVM model to detect PPD status trained on a model
incorporating variation in both *TTC9B* and *HP1BP3* across three technical replicates in a
subset of *N* = 20 women from the JHU Prospective Neuroimaging cohort. Horizontal dashed
red lines denote an AUC of 80%.

Table 1

Sample demographics.

Cohort	Johns Hopkins Prospective PPD Cohort	Johns Hopkins Neuroimaging Cohort	UC Irvine Cohort	Emory Cohort
Total	51	68	113	53
EPDS ≥ 13	29 *	8	13	4
EPDS < 13	22 *	60	82	49
PPD Assessment	Prospective Clinical	Prospective EPDS	Prospective EPDS	Prospective EPDS
1st trimester	9	0	107	53
2nd trimester	22	0	111	0
3rd trimester	20	68	113	53

*
Numbers reflect women clinically assessed as having postpartum depression.