

Predicting Perinatal Depression using Electronic Health Records

Varuni Sarwal

Advisor: Eleazar Eskin



Perinatal depression



Jilian Tamaki, for the NYTimes

Depression during pregnancy and following childbirth

- 10-20% of mothers
- often starts during pregnancy
- continues in postpartum period

One of the greatest causes of mortality and morbidity in mothers, high risk of suicide

Poor outcomes in children: learning delays, visits to the ER, cognitive and social delays, infanticide



But, perinatal depression



Berthe Morisot, "Le Berceau," 1872

Hypothesized to be more genetically homogeneous than non-PND depression

++ heritable than non-PND depression, $h^2 = 40-55\%$

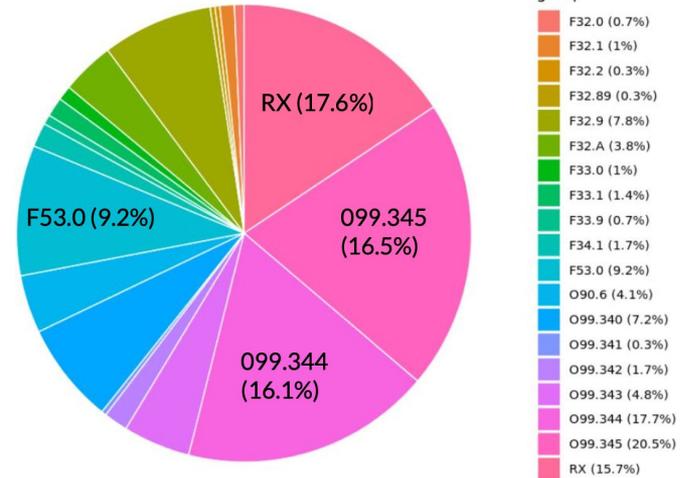
The period of interest is
well-defined
well-motivated
marked by high health care utilization

PND has been overwhelmingly understudied compared to other psychiatric disorders



EHR diagnosis \neq onset

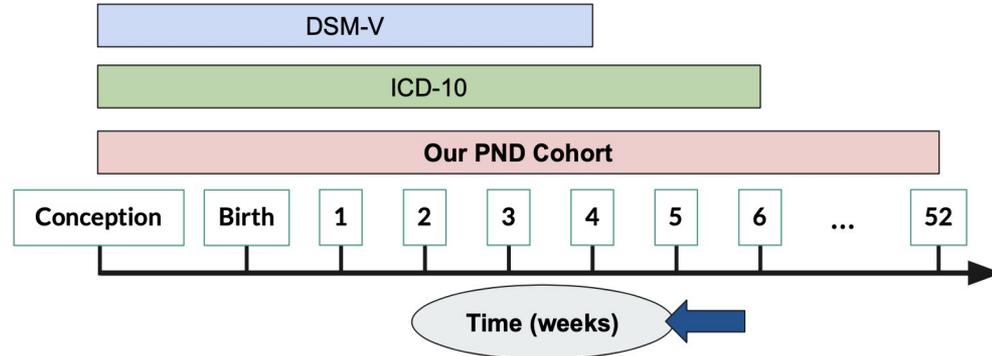
1. Stigma of mental health treatment
2. Uncertainty about the nature of their symptoms (postpartum “blues” vs. illness requiring treatment)
3. Lack of time for self care
4. Other socioeconomic factors



Perinatal period definition = from conception through 1 year of birth

Eligibility criteria:

- A new diagnosis of depression, postpartum depression, depression related to childbirth
- A new prescription for antidepressants



Methods

Goal

- Identify at-risk mothers
- Predict time to PND at the first prenatal visit
- Can subsequent visits help prediction?

Data

- UCLA Health System: diverse population of LA, ~4 million patients, 20,000 births with EHR.

Prediction

- Labs: first visit +/- 3 weeks
- Medications: first visit +/- 1 year
- Demographics (language, ethnicity, smoking)

Diagnosis

- A combination of ICD codes

Model

- Logistic Regression (baseline)
- Tree-based models
- Deep learning



Diagnoses



Demographics



Labs



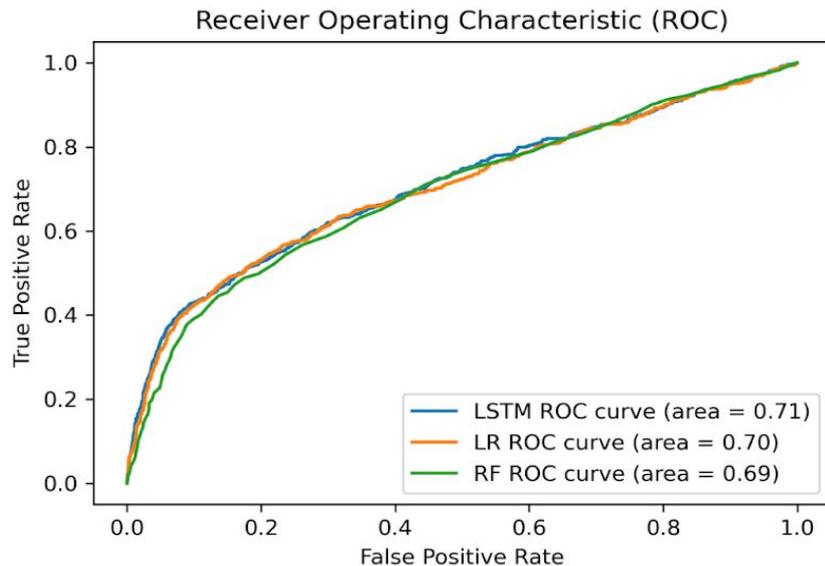
Medication



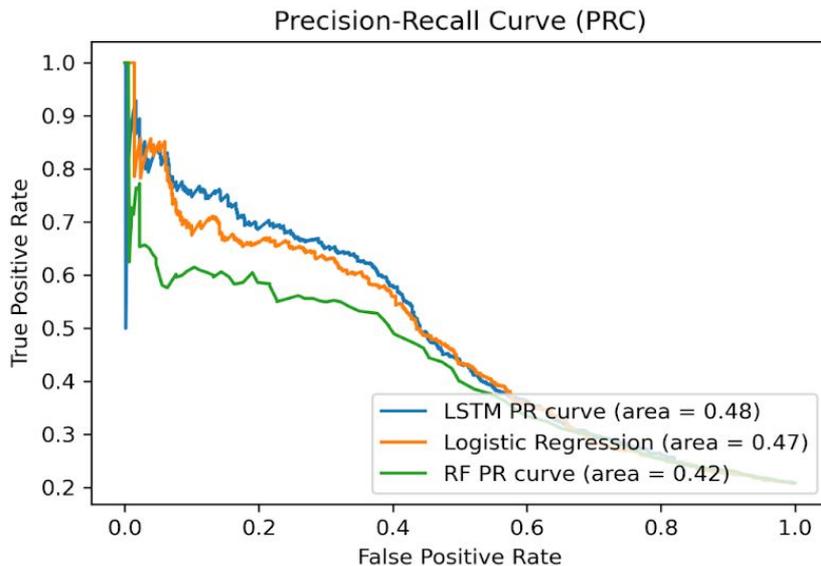
Other features



Results



Our prediction tasks is very challenging yet meaningful since we only use data from the first hospital visit

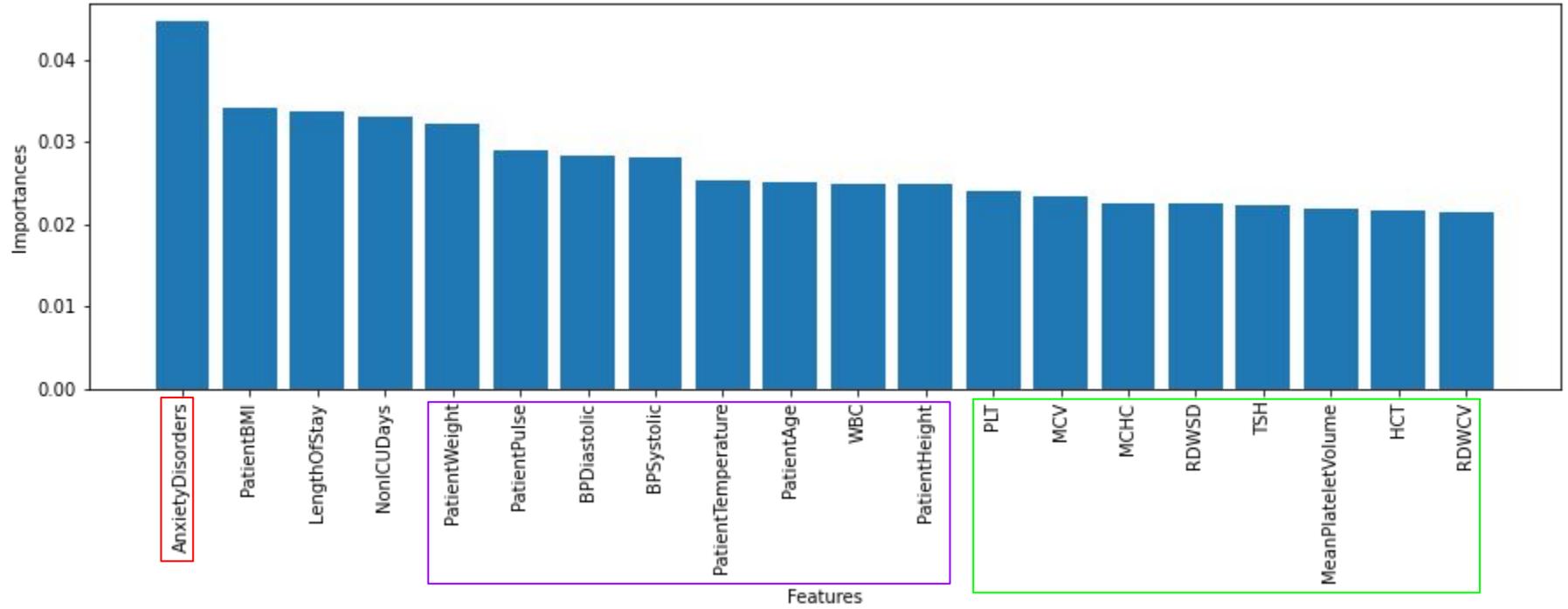


We were able to predict the onset of PND with an AUROC of 0.71 and AUPRC of 0.48

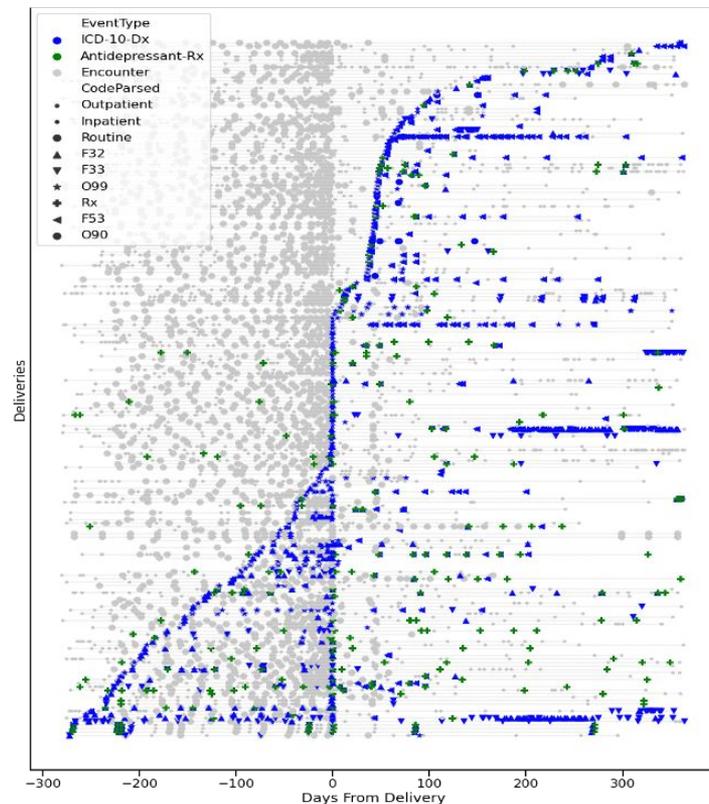
Results stratified by race

Race	# individuals	ROC	PRC	Accuracy
Latinx	2067	0.689	0.280	84.3%
Black	564	0.683	0.422	80.85%
White	4742	0.670	0.30	84.06%

Feature importances: Random Forest

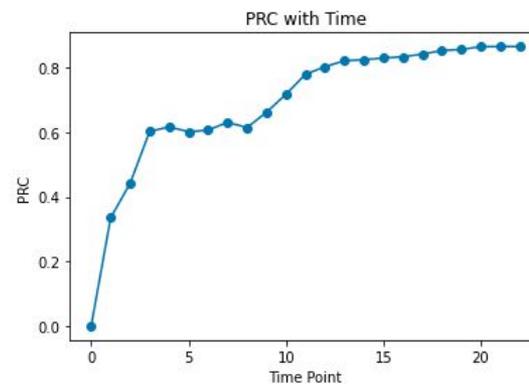
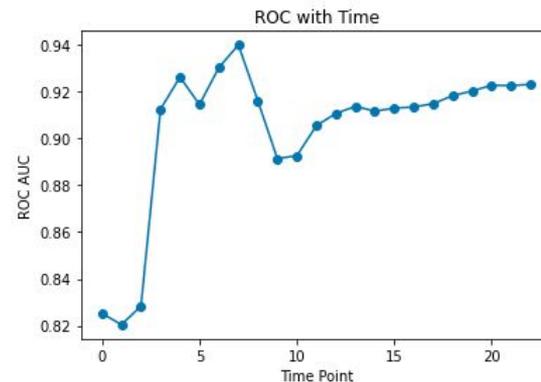
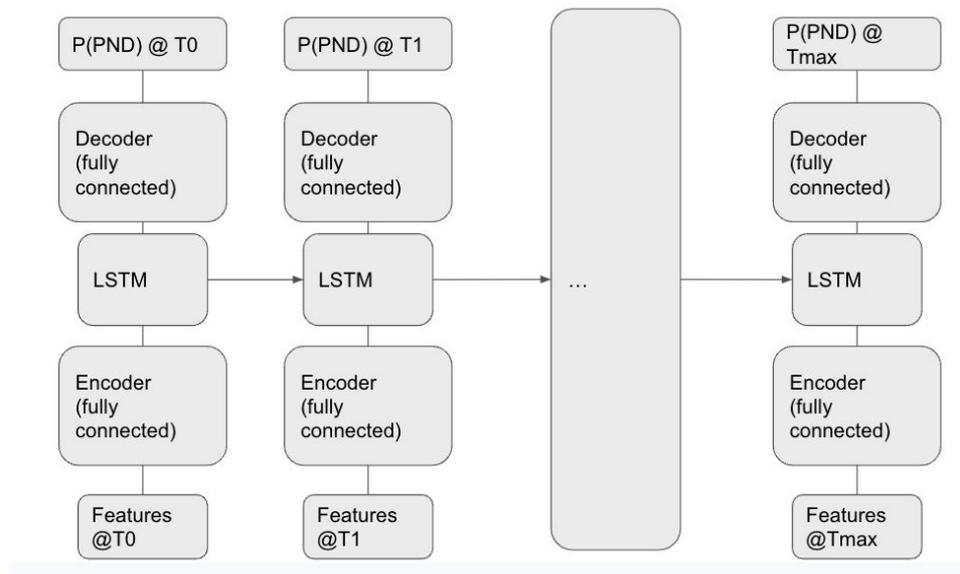


What next? Pregnant women have multiple encounters



Can we use this additional information to help prediction?

Proposed Method: LSTM network



Future Work

- How early can we predict?
- Why do we observe a low predictive performance at $t=0$?
- How can we explain the trends in AUROC and AUPRC
- What features become more important with time?