

A Risk Prediction Model for Postpartum Depression

Model Design

This model is a regularized logistic regression (LASSO) model which calculates the likelihood a patient will develop postpartum depression within 1 year after childbirth. For more information about the definitions and variables used in the model, refer to the this published [article](#) by Zhang et al.

Population Statistics

To train and test the model, we used data from 2015 to 2018 from Weill Cornell Medicine (WCM), drawn from over 15,000 patient encounters. Patients from the emergency department, inpatient, and outpatient contexts were used. See the Data Definition topic for specific population inclusion criteria. The model was trained on 5-fold cross validation of 80* of the data at WCM, while the other 20* was used to validate the performance of the model. The model was further validated using data from 2014 to 2018 from a consortium of health organizations (INSIGHT) of over 50,000 patients using the same inclusion criteria. The site information is listed in Table 1.

Table 1. Population statistics at 2 sites

Site	Patients	Prevalence
WCM	15197	6.7*
INSIGHT	53972	6.5*

Development Methodology

With the training population defined, we looked back in time up to the first record of prenatal visit and gathered all relevant data elements for each patient until 1 year after childbirth. We identified over 1000 data elements from the electronic health record (EHR), each representing a potentially predictive variable in the model. These include vital signs, medication orders, lab results, comorbidities, and demographic information. Values for a patient were grouped by trimester. Missing data were imputed using means. Variable selection was done through an iterative process of sequential feature selection (SFS) followed by clinical validation to remove clinically trivial variables. Model parameters were determined using a grid search that comprehensively searched for the best hyperparameters and parameters that resulted in the highest model performance as measured by area under the receiver operating characteristic curve (AUC). SFS was performed separately for patients with, and without, mental health history to ensure that the model can predict for both types of patients when in actual use. We combined features selected from both SFS into a single feature set such that a single model can be used for patients with and without a history of mental illness. In total, we ended up with 32 variables considered to be most predictive.

Model performance

Our testing set consisted of 20* of the patients seen at WCM and 100* of INSIGHT. Because we didn't use these patients in training the model, they provide more objective feedback on the model's performance. It is important to lay out a method for incorporating the temporal element of the pregnancy. That is, while it is important to understand that a patient may develop postpartum depression, it must also be timely to suggest a different clinical action. If a clinician is notified 24 weeks after gestation, it may be possible that

an intervention can be offered at that time as a primary prevention. On the contrary, if a clinician is notified after the symptom onset, an opportunity for prevention would have been lost.

Therefore, we created the following strategy for evaluating the model performance at multiple time periods during pregnancy. We tested the model performance using data that were available at up to 12 weeks, 18 weeks, 24 weeks, 30 weeks, 36 weeks, and at childbirth to evaluate how early the model can start performing satisfactorily as measured by the AUROC. Thresholds were defined using the population at WCM, but can be modified based on new populations and workflows. The performance at the 2 sites (AUROC, sensitivity, specificity, Brier score, PPV, NPV), is listed in Table 2 below.

Table 2. Model performance across 2 sites

Time	Classifier	AUROC	Sensitivity	Specificity	Brier Score	PPV	NPV
WCM (development site)							
12wk	Logistics regression	0.921(0.893,0.949)	0.79	0.97	0.074	0.61	0.99
18wk	Logistics regression	0.919(0.891,0.947)	0.79	0.97	0.074	0.61	0.99
24wk	Logistics regression	0.922(0.895,0.949)	0.79	0.97	0.074	0.61	0.99
30wk	Logistics regression	0.921(0.893,0.949)	0.79	0.97	0.074	0.61	0.99
Childbirth	Logistics regression	0.937(0.912,0.962)	0.83	0.96	0.082	0.59	0.99
CDRN (validation site)							
12wk	Logistics regression	0.810(0.801,0.819)	0.70	0.85	0.150	0.24	0.98
18wk	Logistics regression	0.817(0.808,0.826)	0.70	0.85	0.151	0.24	0.98
24wk	Logistics regression	0.821(0.812,0.830)	0.71	0.85	0.152	0.25	0.98
30wk	Logistics regression	0.824(0.815,0.833)	0.72	0.85	0.153	0.24	0.98
Childbirth	Logistics regression	0.886(0.879,0.893)	0.80	0.84	0.158	0.26	0.98

The distribution of predicted probabilities by case and controls in the test data are shown in Figure 1. As the figure shows, most of the cases are being predicted to be at risk (red), whereas controls are not (green).

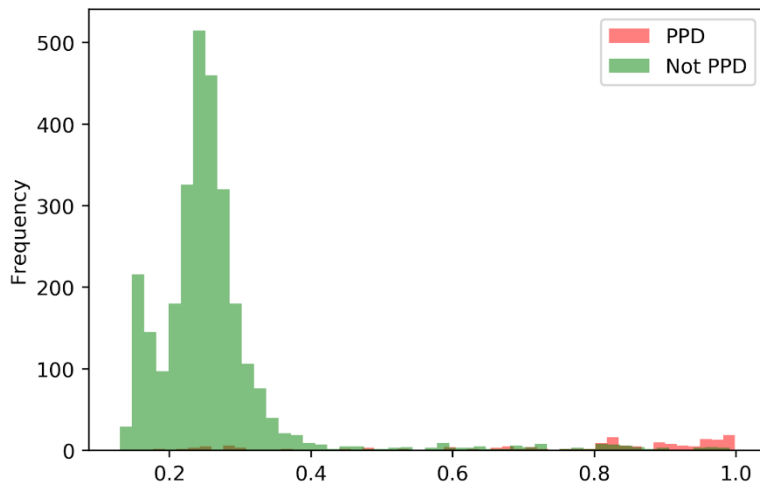


Figure 1. The distribution of predicted probability. Green: controls; Red: cases

Model bias and fairness were evaluated to ensure the model does not perform unequally across patient subgroups, as shown in Table 3.

Table 3. Bias and fairness evaluation

		N (test)	AUROC	Precision	Sensitivity	Specificity
All		3040	0.94	0.59	0.83	0.96
Race	White	1510	0.92	0.58	0.83	0.95
	Asian	569	0.95	0.57	0.67	0.99
	Black	156	0.85	0.42	0.83	0.95
	Other	349	0.93	0.53	0.89	0.95
	Unknown	456	0.96	0.79	0.85	0.99
Insurance	Public	360	0.99	0.63	1.00	0.98
	Private	2267	0.94	0.60	0.82	0.96
	Unknown	413	0.90	0.53	0.81	0.95

Data Definitions

Population

In order to be considered for the model, the following must have been true for a patient:

- Age 18 to 45 years old
- Receive prenatal care at the health organization

Outcome definition

The outcome is defined as having a diagnosis of postpartum depression within 1 year of childbirth. A postpartum depression diagnosis was defined using ICD-9, ICD-10, Systematized Nomenclature of Medicine (SNOMED) codes and the use of antidepressants within 1 year following childbirth (Dietz et al., 2007; Stewart and Vigod, 2016). The specific diagnostic codes for postpartum depression definition are listed in the published [article](#). The use of antidepressants was defined by Anatomical Therapeutic Chemical (ATC) codes under N06A (Petersen et al., 2018). To ensure that antidepressants were primarily used for treatment of mental health conditions, and not for other indications such as pain, we further excluded the following medications: Amitriptyline, Clomipramine, Duloxetine, Fluoxetine, and Nortriptyline (Schofield et al., 2016).

Variables

Demographics	ICD9/10/ATC Code (if applicable)
Race	
Single (vs. Married)	
Medical history	
Anxiety history	F40*, F41*, F42*, F93.0*, F93.1*, F93.2*, F43.0*, F43.1*, F43.8*, F43.9*, 300.0*, 300.2*, 300.3*, 309.8*, 308.3*
Other disorder history	F43.2*, F99*, F52*, F64*, F65*, F66*, F63*, F91*, F92*, F95*, F44*, F48*, F50*, F51*, F53*, F54*, F68*, F70*, F71*, F72*, F73*, F84*, F98*, F79*, F78*, F80*, F81*, F82*, F83*, F88*, F89*, F90*, F93.3*, F93.4*, F93.5*, F93.6*, F93.7*, F93.8*, F93.9*, F94, 309.0*, 309.1*, 309.2*, 309.3*, 309.4*, 309.8*, 309.9*, 302.0*, 302.1*, 302.2*, 302.3*, 302.4*, 302.5*, 302.6*, 302.7*, 302.8*, 302.9*, 297.0*, 297.1*, 297.2*, 297.3*, 297.8*, 297.9*, 312.0*, 312.1*, 312.2*, 312.3*, 312.4*, 312.8*, 312.9*, 298.0*, 298.1*, 298.2*, 298.3*, 298.4*, 299.0*, 299.1*, 299.8*, 299.9*, 300.1*, 300.5*, 300.6*, 300.7*, 300.8*, 300.9*, 307.0*, 307.1*, 307.2*, 307.3*, 307.4*, 307.5*, 307.6*, 307.7*, 307.9*, 308.0*, 308.1*, 308.2*, 308.9*, 310*, 313.0*, 313.1*, 313.2*, 313.3*, 313.8*, 313.9*, 314.0*, 314.1*, 314.2*, 314.8*, 314.9*, 315.0*, 315.1*, 315.2*, 315.3*, 315.4*, 315.5*, 315.9*, 316*, 317*, 318.0*, 318.1*, 318.2*, 319*
Mood disorder history	F30*, F31*, F34.0*, F32*, F33*, F34.1*, F38.1*, F34.8*, F34.9*, F38.0*, F38.8*, F39*, 291.0*, 291.1*, 291.2*, 291.3*, 291.4*, 291.5*, 291.6*, 291.7*, 291.8*, 291.9*, 292.0*, 292.1*, 292.2*, 292.3*, 292.4*, 292.5*, 292.6*, 292.7*, 292.8*, 292.9*, 303.0*, 303.1*, 303.2*, 303.3*, 303.4*, 303.5*, 303.6*, 303.7*, 303.8*, 303.9*, 304.0*, 304.1*, 304.2*, 304.3*, 304.4*, 304.5*, 304.6*,

	304.7* , 304.8* , 304.9* , 305.0* , 305.1* , 305.2* , 305.3* , 305.4* , 305.5* , 305.6* , 305.7* , 305.8* , 305.9*
Health issues	
Depression in pregnancy	F32, F33, F32.0, F32.1, F32.8, F32.9, F33.0, F33.1, F33.2, F33.9, F32.89, 296.2, 296.20, 296.25, 296.21, 296.3, 296.22, 296.31, 296.30, 296.32, 311, 296.33, F33.0, F33.2, F33.41, F32.8, F32.9, F33.1, F33.9, F32
Anxiety in pregnancy	F40* , F41* , F42* , F93.0* , F93.1* , F93.2* , F43.0* , F43.1* , F43.8* , F43.9* , 300.0* , 300.2* , 300.3* , 309.8* , 308.3*
Mental disorder in pregnancy	F99, 307, 300.1, 307.9, V40.2, 648.41, F69, F09, F68, F06, F99
Palpitations	R00.2, 785.1
Vomiting in pregnancy	O21.2, O21.8, O21.9, 643.9, 643.8, 643.80, 643.20, 643.83, 643.91, 643.93, 643.90, 643.81, 643.23, 643.2, 643.21, O21.8, O21.2, O21.9
Hypertensive disorder	402, 402.01, 402.9, 403.91, 404.01, 404.02, 404.93, 405.99, 642.12, 642.13, 642.14, H35.031, I10, I11, I11.9, I13, I97.411, O10.113, O10.219, 362.11, 401, 402, 402.1, 403.9, 403.9, 404, 404.11, 404.12, 437.2, 459.8, 642.11, 642.9, 642.9, 642.91, 642.92, 642.93, 997.91, H35.039, I11.0, I13.0, I15.8, I15.9, I16, I16.9, I51.6, I67.4, I97.41, I97.42, I97.620, I97.621, I97.63, N26.2, O10.211, O13, O13.2, O13.3, O13.4, P29.2, Q87.41, R03.0, 402.1, 402.11, 402.91, 403.1, 404, 404.03, 404.13, 404.9, 404.9, 404.92, 405.9, 429.2, 459, 642, 642.1, 642.1, 642.94, H35.032, I13.1, I13.10, I15, I97.418, I97.631, I97.638, I97.64, I97.640, I98, I98.8, I99.8, O10.11, O10.111, O10.13, O10.212, O10.22, O10.3, O13.1, V81.1, 401.9, 402, 402.9, 403, 403.1, 403.11, 404.1, 404.1, 404.91, 405, 459.89, 459.9, 796.2, H35.03, H35.033, I12, I12.0, I12.9, I13.11, I13.2, I13.9, I97.4, I97.410, I97.622, I97.630, I97.641, I97.648, I99, I99.9, O10.1, O10.112, O10.119, O10.12, O10.2, O10.21, O10.213, O10.23, O13.9, Q87.418,
Acute pharyngitis	J02, J02.8, J02.9, 462, J02.9, J02, J02.8
Hemorrhage in early pregnancy antepartum	O46, O46.8, O46.8X, O46.9, 641.80, 641.81, 641.91, 641.8, 641.93, 641.9, 641.90, 641.83, O46.90, O46.8X9, O46.92, O46.8, O46.8X3, O46.9, O46.91, O46.93, O46.8X1, O46.8X2, O46
Diarrhea	787.91, R19.7
Threatened miscarriage	640.03, 640.01, 640.00, 640.0, O20.0
Abdominal pain	R10.8, 789.09, 789.00, 789.0, 789.07, R10.84, R10.4, R10.9
Migraine	G43, G43.8, G43.80, G43.9, G43.90, 346.9, 346.90, 346, 346.80, 346.8, G43.A0, G43, G43.809, G43.8, G43.9, G43.909
Hypothyroidism	E03, E03.9, E03.9, E03.8
Placental infarct	O43.819, O43.811, O43.81, O43.813, O43.812

Primigravida	O09.51, 659.50, 659.5, O09.519, O09.512, O09.513, O09.511
Pre-eclampsia	O14, O14.0, O14.9, O14.04, O14.05, O14.94, O14.95, 642.40, 642.4, 642.41, 642.44, 642.43, 642.42, O14.02, O14.92, O14.03, O14.90, O14.0, O14.93, O14.9, O14, O14.00
Abnormality of organs and/or soft tissues of pelvis affecting pregnancy	O34.4, O34.5, O34.7, O34.8, O34.9, 654.43, 654.62, 654.63, 654, 654.42, 654.82, 654.4, 654.44, 654.83, 654.80, 654.64, 654.8, 654.41, 654.6, 654.60, 654.61, 654.40, O34.70, O34.81, O34.71, O34.40, O34.90, O34.92, O34.4, O34.42, O34.72, O34.41, O34.93, O34.7, O34.83, O34.82, O34.6, O34.5, O34.8, O34.80, O34.9, O34.73, O34.91, O34.43
False labor at or after 37 completed weeks of gestation	O47.1
Medications	
Antidepressants	N06A
Beta blocking agents	C07A
Direct acting antivirals	J05A
Other antibacterials	J01X
Antihistamines for systemic use	R06A
Vitals	
Diastolic blood pressure in third trimester	
Procedures	
Deliveries by cesarean	O82, 669.70, 669.7, O82.9, O82.8
Health Services	
Number of emergency department visits during pregnancy	