Application and Comparison of Machine Learning Models for the Prediction of Postpartum Depression: Research Based on a Cohort Study

Weina Zhang, Han Liu, Vincent Michael Bernard Silenzio, Peiyuan Qiu, Wenjie Gong

Submitted to: JMIR Medical Informatics
on: July 18, 2019

Disclaimer: © The authors. All rights reserved. This is a privileged document currently under peer-review/community review. Authors have provided JMIR Publications with an exclusive license to publish this preprint on its website for review purposes only. While the final peer-reviewed paper may be licensed under a CC BY license on publication, at this stage authors and publisher expressively prohibit redistribution of this draft paper other than for review purposes.
# Table of Contents

**Original Manuscript** ........................................................................................................................................................................... 4

**Supplementary Files** ........................................................................................................................................................................... 23

**Figures** .................................................................................................................................................................................................... 24
  - Figure 6 .............................................................................................................................................................................................. 25
  - Figure 2 .............................................................................................................................................................................................. 26
  - Figure 1 .............................................................................................................................................................................................. 27
  - Figure 3 .............................................................................................................................................................................................. 28
  - Figure 4 .............................................................................................................................................................................................. 29
  - Figure 5 .............................................................................................................................................................................................. 30

**Other materials for editor/reviewers only** ......................................................................................................................................... 31
  - Other materials for editor/reviewers only 0 .................................................................................................................................... 32

**Multimedia Appendixes** ................................................................................................................................................................... 33
  - Multimedia Appendix 0 ................................................................................................................................................................. 33

**Other materials for editor/reviewers only** ......................................................................................................................................... 34
  - Other materials for editor/reviewers only 0 .................................................................................................................................... 34
Application and Comparison of Machine Learning Models for the Prediction of Postpartum Depression: Research Based on a Cohort Study

Weina Zhang BSc, ; Han Liu MSc, ; Vincent Michael Bernard Silenzio MD, MPH, ; Peiyuan Qiu PhD, ; Wenjie Gong PhD,

Corresponding Author: Wenjie Gong PhD, Phone: +8613607445252 Email: gongwenjie@csu.edu.cn

Abstract

Background: Postpartum depression (PPD) is a serious public health problem. Building a predictive model of PPD using data in pregnancy can facilitate earlier identification and intervention.

Objective: To compare the effects of four different machine learning models using data in pregnancy to predict PPD and explore which factors in the model are the most important for PPD prediction.

Methods: Information on the pregnancy period from a cohort of 508 women, including demography, social environmental factors and mental health, was used as predictors in the models. Edinburgh Postnatal Depression Scale (EPDS) score within 42 days after delivery was used as the outcome indicator. Using two feature selection methods (expert consultation and filter feature selection based on random forest (FFS-RF)) and two algorithms (support vector machine (SVM) and random forest (RF)), we developed four different machine learning PPD prediction models and compared their prediction effects.

Results: There was no significant difference in the effectiveness of the two feature selection methods in terms of model prediction performance, but 10 fewer factors were selected with the FFS-RF than with the expert consultation method. The model based on SVM and FFS-RF had the best prediction effects (sensitivity = 0.69, AUC = 0.78). In the feature importance ranking output by the RF algorithm, psychological elasticity, depression during the third trimester and income level were the most important predictors.

Conclusions: In contrast to the expert consultation method, FFS-RF was important in dimension reduction. When the sample size is small, the SVM algorithm is suitable for predicting PPD. In the prevention of PPD, more attention should be paid to the psychological resilience of mothers.

(JMIR Preprints 18/07/2019:15516)
DOI: https://doi.org/10.2196/preprints.15516

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?
✔ Please make my preprint PDF available to anyone at any time (recommended).
   Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
   Only make the preprint title and abstract visible.
   No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?
✔ Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).
   Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain visible to all users.
   Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in the PubMed Now! service, my accepted manuscript PDF will automatically be made openly available.

https://preprints.jmir.org/preprint/15516 [unpublished, non-peer-reviewed preprint]
Introduction
Postpartum depression (PPD) is a serious public health problem that affects 10−20% of pregnant women[1-3]. PPD not only adversely affects the physical and mental health of mothers, it is detrimental to the growth and development of infants. In extreme cases, suicide and infanticide may even occur[4]. Establishing an effective PPD prediction model that can be used in pregnancy may enable earlier identification, thus helping health care providers to provide more effective management of at-risk patients[5]. Previous studies have explored this possibility and demonstrated feasibility[6-7].

As machine learning (ML) may be useful in making accurate predictions based on data from many sources, these approaches have been applied in prediction studies in recent years[8]. There are many predictive factors for postpartum depression, involving demographic, psychological, and environmental aspects[5,9,10]. Using risk factors during pregnancy can allow enough time for subsequent interventions. The expert consultation method has often been used to generate guidelines for PPD detection, based on expert opinion and clinical experience. In contrast, ML approaches rely on the use of empirical data to generate prediction models. The key to building good ML models is in the rigorous selection of appropriate features and algorithms. There are two approaches to address the important challenge in ML of feature selection: Filter and Wrapper[11]. A random forest-based filter feature selection algorithm (FFS-RF) can use the importance score of a so-called random forest of variables as the evaluation criterion for feature selection, in order to identify the subsets of data feature that may be most relevant to accurately predict the targeted outcome variable(s) of interest. Such strategies to identify the most relevant data features have proven to be effective ways to explore the risk factors for some diseases[12]. There are two main algorithms used in depression prediction studies, namely, the support vector machine (SVM) and random forest (RF) algorithms[8]. Depression prediction studies using these two methods have achieved relatively good results[13-15]. SVM is an example of supervised learning. It focuses on minimizing structural risks within the set of available data[16]. It has great advantages in solving high-dimensional modeling problems and performs well in situations where relatively smaller sample data may be available[17]. In contrast, RF models are built using a decision tree as the basic classifier. RF approaches have high classification accuracy, strong inductive capacity, a simple parameter adjustment process, fast calculation speed, relatively low sensitivity to missing data values, and the ability to output feature importance[12,18].

Comparison between those ML methods concerning PPD has not been studied. This study is based on data drawn from a large, ongoing cohort study of pregnant women in the Hunan province of south central China. In the work described here, we combined the two feature selection methods and the two ML algorithms described above to assess four PPD prediction models using data during pregnancy, to compare the effect of PPD prediction models, pick the optimal predictive model, and provide a reference for the development of ML in PPD.

Methods
This study was part of a larger cohort study. All the data included here is original and previously unpublished. Researchers in the study collected the following measures at a series of seven visits conducted in the first trimester through six weeks postpartum: depression (using the Edinburgh Postnatal Depression Scale), social environment, and psychological and biological factors associated with depression. The study was approved by the institutional review board of the institute of clinical pharmacology of Central South University (ChiCTR-ROC-16009255).

Sampling
Participants were recruited from two Maternity and Child Care Centers of Hunan Province and Yiyang City. The former is a major provincial teaching hospital located in Changsha, a city with approximately 8.15 million residents. The latter (Yiyang City) is a less economically developed area of Hunan province, with approximately 4.39 million residents. Researchers sought for recruit women
in the obstetric clinics of the two hospitals from September 2016 to February 2017. Following is the inclusion criteria for participants: 1) woman, 2) age \( \geq 18 \) years old, 3) gestational weeks \( \leq 13 \) weeks (pregnancy weeks are estimated by the first day of the last menstrual period). All participants signed informed consent. Eventually, 1,126 women were recruited.

Measures

Following tools were used to collect data:

1) Study Questionnaire: A purpose-built questionnaire, designed for this study and optimized through a pilot survey, was used to collect information including the following: age, education degree, income level, occupation, marital satisfaction, first pregnancy, folic acid intake, premenstrual syndrome, history of mental health concerns, family history of mental illness, mother's menopausal symptoms, childhood experiences and life events.

2) The EPDS was used to self-report maternal symptoms of depression[19]. The EPDS is a 10-item self-rated questionnaire, with each item scored from 0 to 3, with a total score ranging from 0 to 30. The Chinese language EPDS used in this study was translated by Wang Yuqiong[20]. The EPDS is the most common PPD screening tool[21,22]. The critical value was 9.5.

3) The Brief Resilience Scale (BRS) was used to determine the level of psychological resilience. The BRS is a six-item questionnaire that reflects the respondent's ability to bounce back or recover from stress. The score is the average score of each item, and the higher the score, the stronger the strain and adaptability[23].

4) The Pittsburgh Sleep Scale (PSQI) is a comprehensive scale that reflects the sleep quality of subjects. It is composed of 7 dimensions: "Sleep Quality", "Sleep Latency", "Sleep Duration", "Sleep Efficiency", "Sleep Disorders", "Use of Sleep Medications", and "Daytime Dysfunction". The scores of each dimension are summed to obtain the total PSQI score. The higher the score is, the worse the sleep quality. According to the total score, sleep quality can be divided into different grades: 6-10 indicates "good sleep quality", 11-15 indicates "average sleep quality", and 16-21 indicates "poor sleep quality"[24]. The scale has good reliability and validity[25].

5) The Social Support Rating Scale (SSRS), which was designed by Shuiyuan Xiao[26], was used to measure social support. The SSRS is a ten-item questionnaire with three dimensions, namely, objective support, subjective support, and utilization of social support. The higher the total score and the scores for each dimension, the better an individual's level of social support.

6) The Generalized Anxiety Scale (GAD-7) was developed by Spitzer[27]. The score is obtained by summing the scores of 7 items. Most current studies consider a total score of \( \geq 10 \) as indicative of anxiety[27,28].

Procedure

Seven time points were selected for depression screening, corresponding to the women’s routine obstetric examinations. We divided these into first trimester (gestational week 13 or earlier), second trimester (weeks 17-20 and 21-24), third trimester (weeks 31-32 and 35-40) and postpartum (7 days and 6 weeks postpartum). Except for the first period, screening for perinatal depression by EPDS was performed twice in the remaining periods respectively. If one or more of the EPDS \( \geq 9.5 \) in each of these grouped sets of visits were positive, the participant was regarded as at risk for depression in this period. The Study Questionnaire, BRS and GAD-7 were assessed during the first trimester, whereas the PSQI was used during the second trimester and SSRS during the third trimester. At the study’s conclusion, 508(508/1126, 45.12%) of 1126 subjects completed all screening (See Figure 1).

Figure 1 Participant recruitment and response condition

https://preprints.jmir.org/preprint/15516 [unpublished, non-peer-reviewed preprint]
Feature Selection
Two simple and easy-to-implement methods were used for feature selection, namely, the expert consultation method and the FFS-RF methods. The expert consultation method was utilized to select clinically relevant factors as appropriate predictors of pre-existing or potential PPD. This was accomplished by consulting experts in the area of obstetrics and gynecology as well as mental health practitioners. The FFS-RF was utilized to identify proper predictors for PPD. Under this approach, features within a certain bound value range (p>.05) were selected as potential predictors and incorporated into the final prediction model.

Model Development
Of the 508 participants, 75% (381/508) were randomly selected for model training. Data from the remaining participants (n=127) was held back for use in model testing and verification. Table 1 shows the model selection scheme. Based on the expert consultation method and FFS-RF, four PPD prediction models were generated by SVM and RF algorithms. The parameters of the models were optimized, and the specific parameters are shown in Table 5.

Table 1. Names of the PPD prediction models

| Machine learning modeling algorithm | Feature selection method | FFS-RF<sup>a</sup> |
|------------------------------------|-------------------------|--------------------|
| Random forest (RF)                 | E-RF<sup>b</sup>        | F-RF<sup>c</sup>   |
| Support vector machine (SVM)       | E-SVM<sup>d</sup>       | F-SVM<sup>e</sup>  |

<sup>a</sup>FFS-RF: Filter feature selection based on random forest
<sup>b</sup>E-RF: Model built using the RF algorithm and expert consultation method
Evaluation of Model Effects
For the test set, we used the trained models to test and compare their prediction of PPD with real data and create a confusion matrix (see Table 2). A series of indicators were obtained of each model. The index formula are as follows:

1) Accuracy = \frac{a + d}{a + b + c + d}
2) Mis-classification rate = \frac{b + c}{a + b + c + d}
3) Positive predictive value = \frac{a}{a + b}
4) Negative predictive value = \frac{c + d}{d}
5) Sensitivity (Sen) = \frac{a}{a + c}
6) Specificity (Spe) = \frac{d}{b + d}
7) G-Mean = \sqrt{Sen * Spe}

The sensitivity and Area under ROC Curve (AUC) were used to evaluate the effects of each model and choose the best prediction model. To select the optimal model, we first selected the model with an AUC > .75 to confirm that it had a good comprehensive prediction effect. On this basis, we then selected the model with the highest sensitivity as the best prediction model to ensure that as many mothers as possible with a high risk of PPD would be detected.

Table 2. Confusion matrix

| Predicted Results | Real Results |
|-------------------|-------------|
| Positive          | a           | c           |
| Negative          | b           | d           |

Statistical Analysis
This study used the Redcap system to build a database and SPSS version 18.0 to clean the data. The training and test sets were analyzed by the ‘sklearn.model_selection.train_test_split’ package. The RF data were analyzed by the ‘sklearn.ensemble.randomforestclassifiers’ package. The SVM data were analyzed by the ‘sklearn.svm.SVC’ package. Cross-validation was performed using the ‘sklearn.cross_validation’ package. All these packages were available in the Python 3.6 software.

Results
Candidate predictors
Table 3 shows the 25 candidate predictors of the subjects with and without PPD. Among the 508 subjects, 173 (173/508, 34.06%) were regarded as having PPD. The average age of the pregnant women was 28.64 years old (SD 4.344). The average BRS score was 3.10 (SD 0.371). The average personal monthly income of the women and their spouses was between 2,000 and 5,000 yuan ($393–785). Most of the subjects had a bachelor’s degree. Of the 173 women with PPD, 116 (116/173, 67.05%) had positive EPDS screening results in the third trimester. Table shows the results of the single factor analysis (p[].05).
Table 3. Comparison of candidate predictors, including data sets of women without PPD (n=335) and women with PPD (n=173) among the total sample (n=508) of pregnant women.

| Item                        | Women without PPD (n=335) | Women with PPD (n=173) | P value |
|-----------------------------|---------------------------|------------------------|---------|
| Age (years), Mean (SD)      | 28.7(4.5)                 | 28.5(4.0)              | 0.7     | .49     |
| Education degree            |                           |                        |         |
| Junior high school and below| 29(8.7)                   | 19(11.0)               | -0.6    | .54     |
| High school                 | 77(23.1)                  | 40(23.1)               |         |         |
| Bachelor’s                  | 198(59.3)                 | 99(57.2)               |         |         |
| Master’s degree and above   | 30(9.0)                   | 15(8.7)                |         |         |
| Education degree (husband)  |                           |                        |         |
| Junior high school and below| 42(12.6)                  | 20(11.9)               |         |         |
| High school                 | 79(23.7)                  | 36(21.4)               | -0.2    | .87     |
| Bachelor’s                  | 177(53.2)                 | 99(58.9)               |         |         |
| Master’s degree and above   | 35(10.5)                  | 13(7.7)                |         |         |
| Income level                |                           |                        |         |
| ¥0                          | 80(24.3)                  | 56(33.3)               |         |         |
| ¥0 and ¥2000                | 20(6.1)                   | 9(5.4)                 | -1.7    | .09     |
| ≥¥2000 and ≤¥5000           | 178(54.1)                 | 78(46.4)               |         |         |
| Item                                      | Women without PPD(n=335) | Women with PPD(n=173) | P value |
|-------------------------------------------|--------------------------|-----------------------|---------|
| Income level (husband)                    |                          |                       |         |
| ≥¥5000 and ¥10000                         | 41(12.5)                 | 22(13.1)              |         |
| ≥¥10000                                   | 10(3.0)                  | 3(1.8)                |         |
| Occupation                                |                          |                       |         |
| Public officials                          | 80(24.2)                 | 28(16.4)              |         |
| Corporation management                    | 59(17.8)                 | 33(19.3)              |         |
| In business (self-employed)               | 52(15.7)                 | 20(11.7)              | 10.8    | .03 |
| Unemployed                                | 73(22.1)                 | 58(33.9)              |         |
| Others                                    | 67(20.2)                 | 32(18.7)              |         |
| Marital satisfaction                      |                          |                       |         |
| Satisfied                                 | 294(88.6)                | 133(77.3)             |         |
| Basically satisfied                       | 38(11.4)                 | 38(22.1)              | -3.3    | <.001 |
| Dissatisfied                              | 0(0.0)                   | 1(0.6)                |         |

Table 3. Comparison of candidate predictors, including data sets of women without PPD (n=335) and women with PPD (n=173) among the total sample (n=508) of pregnant women.
**First pregnancy**

|                | Women without PPD (n=335) | Women with PPD (n=173) | P value |
|----------------|---------------------------|------------------------|---------|
| No             | 231 (70.4)                | 119 (68.8)             | 0.1     |
| Yes            | 97 (29.6)                 | 54 (31.2)              | .70     |

**Folic acid intake before this pregnancy**

|                | Women without PPD (n=335) | Women with PPD (n=173) | P value |
|----------------|---------------------------|------------------------|---------|
| No             | 127 (38.0)                | 64 (37.4)              | 0.0     |
| Yes            | 207 (62.0)                | 107 (62.6)             | .90     |

**Premenstrual syndrome - mood instability**

|                | Women without PPD (n=335) | Women with PPD (n=173) | P value |
|----------------|---------------------------|------------------------|---------|
| No             | 256 (76.4)                | 99 (57.2)              | 20.0    |
| Yes            | 79 (23.6)                 | 74 (42.8)              | <.001   |

**Premenstrual syndrome - sleep changes**

|                | Women without PPD (n=335) | Women with PPD (n=173) | P value |
|----------------|---------------------------|------------------------|---------|
| No             | 316 (94.3)                | 160 (92.5)             | 0.7     |
| Yes            | 19 (5.7)                  | 13 (7.5)               | .42     |

**Depression history**

|                | Women without PPD (n=335) | Women with PPD (n=173) | P value |
|----------------|---------------------------|------------------------|---------|
| No             | 224 (96.1)                | 107 (91.5)             | 3.3     |
| Yes            | 9 (3.9)                   | 10 (8.5)               | .07     |

**Other mental illness history**

|                | Women without PPD (n=335) | Women with PPD (n=173) | P value |
|----------------|---------------------------|------------------------|---------|
| No             | 323 (99.1)                | 167 (98.2)             | 0.1     |
| Yes            | 9 (3.9)                   | 8 (8.5)                | .70     |

Table 3. Comparison of candidate predictors, including data sets of women without PPD (n=335) and women with PPD (n=173) among the total sample (n=508) of pregnant women.
Table 3. Comparison of candidate predictors, including data sets of women without PPD (n=335) and women with PPD (n=173) among the total sample (n=508) of pregnant women.

| Item                                                                 | Women without PPD(n=335) | Women with PPD(n=173) | P value |
|----------------------------------------------------------------------|--------------------------|-----------------------|---------|
| Yes                                                                  | 3(0.9)                   | 3(1.8)                |         |
| **Depression history of other family members**                       |                          |                       |         |
| No                                                                   | 317(94.9)                | 162(94.2)             | 0.9     |
| Yes                                                                  | 8(2.4)                   | 3(1.7)                | .64     |
| Not clear                                                            | 9(2.7)                   | 7(4.1)                |         |
| **Other mental illness history of other family members**             |                          |                       |         |
| No                                                                   | 325(97.3)                | 164(95.3)             | —       |
| Yes                                                                  | 4(1.2)                   | 1(0.6)                | .18     |
| Not clear                                                            | 5(1.5)                   | 7(4.1)                |         |
| **Mother's menopausal symptoms**                                     |                          |                       |         |
| No                                                                   | 198(59.3)                | 76(44.4)              | 10.0    |
| Yes                                                                  | 58(17.4)                 | 41(24.0)              | .007    |
| Others                                                                | 78(23.1)                 | 54 (31.6)             |         |
| **Suffered sexual/psychological/physical violence in early age**     |                          |                       |         |
| No                                                                   | 324(96.7)                | 155(89.6)             | 10.7    |
| Yes                                                                  |                          | 45(26.0)              | <.001   |
| Others                                                                |                          | 79(45.7)              |         |

https://preprints.jmir.org/preprint/15516 [unpublished, non-peer-reviewed preprint]
### Results of EPDS\(^a\) (first trimester)

|                  | Women without PPD (n=335) | Women with PPD (n=173) | P value |
|------------------|---------------------------|------------------------|---------|
| Negative         | 268 (80.0)                | 76 (43.9)              |         |
| Positive         | 67 (20.0)                 | 97 (56.1)              | 67.9    | <.001 |

### Results of EPDS (second trimester)

|                  | Women without PPD (n=335) | Women with PPD (n=173) | P value |
|------------------|---------------------------|------------------------|---------|
| Negative         | 254 (75.8)                | 68 (39.3)              | 65.5    | <.001 |
| Positive         | 81 (24.2)                 | 105 (60.7)             |         |

### Results of EPDS (third trimester)

|                  | Women without PPD (n=335) | Women with PPD (n=173) | P value |
|------------------|---------------------------|------------------------|---------|
| Negative         | 259 (77.3)                | 57 (32.9)              | 95.5    | <.001 |
| Positive         | 76 (22.7)                 | 116 (67.1)             |         |

### Scores of BRS\(^b\), Mean (SD)

|                  | Women without PPD (n=335) | Women with PPD (n=173) | P value |
|------------------|---------------------------|------------------------|---------|
|                  | 3.1 (0.3)                 | 3.1 (0.4)              | -1.2    | .25    |

### Scores of PSQI\(^c\), Mean (SD)

**Table 3.** Comparison of candidate predictors, including data sets of women without PPD (n=335) and women with PPD (n=173) among the total sample (n=508) of pregnant women.

---

[unpublished, non-peer-reviewed preprint]
Scores of SSRS\textsuperscript{d}, Mean (SD)

|                           | Expert consultation method | FFS-RF\textsuperscript{e} |
|---------------------------|----------------------------|---------------------------|
| Objective support         | 10.3(2.5)                  | 9.2(2.4)                  |
| Subjective support        | 21.8(3.7)                  | 19.8(4.3)                 |
| The utilization of support| 8.2(1.9)                   | 7.9(2.0)                  |
| Total scores              | 40.3(6.1)                  | 36.9(7.0)                 |

Results of GAD-7\textsuperscript{e}

|                |               |
|----------------|---------------|
| Negative       | 324(97.9)     |
| Positive       | 7(2.1)        |

Feature Selection

The predictive features obtained by the expert consultation method FFS-RF approaches are shown in Table 4. This study included a total of 25 features: 17 were selected as predictive characteristics by expert consultation method, and 7 were selected by FFS-RF. (Table 4 is feature description on expert consultation method and FFS-RF).

| Feature Description | Expert consultation method | FFS-RF\textsuperscript{e} |
|---------------------|----------------------------|---------------------------|
| 1. Age              | 1. Level of psychological resilience |
| 2. Education degree | 2. Depressive symptoms in first trimester |
| 3. Income level     | 3. Income level |

\textsuperscript{a}EPDS: Edinburgh Postnatal Depression Scale  
\textsuperscript{b}BRS: Brief Resilience Scale  
\textsuperscript{c}PSQI: Pittsburgh Sleep Scale  
\textsuperscript{d}SSRS: Social Support Rating Scale  
\textsuperscript{e}GAD-7: Generalized Anxiety Scale
4. Husband’s education degree
5. Husband’s income level
6. Marital satisfaction
7. Sexual/psychological/physical spousal abuse
8. Childhood abuse history
9. Premenstrual syndrome–mood instability
10. Premenstrual syndrome–sleep changes
11. Depression history of woman
12. Depression history of family members
13. Other mental illness history of woman
14. Other mental illness history of family members
15. Mother’s menopausal symptoms
16. Level of psychological resilience
17. Depressive symptoms in the third trimester

*FFS-RF: Filter feature selection based on random forest*

Model Effects

PPD prediction models were established using the RF and SVM modeling applied to the training data set, using the feature sets constructed through our two feature selection methods. The optimal parameters of each model are shown in Table 5. After five-fold cross-validation, we found that when `n_estimator = 200, max_features = sqrt` and `criterion = entropy`, the E-RF model had the best sensitivity. When `n_estimator = 200, criterion = gini`, and `max_features = auto`, the F-RF model had the best sensitivity. Therefore, the software default setting was `max_features = auto`. With the SVM algorithm, regardless of the feature selection strategy, the kernel function with the highest model sensitivity was a linear kernel function.

| Table 5. Optimal parameters for each model |
|-------------------------------------------|
| PPD prediction                            |
| Parameter settings                        |

https://preprints.jmir.org/preprint/15516  [unpublished, non-peer-reviewed preprint]
The model evaluation index is shown in Table 6, and the ROC curves for the four PPD models are shown in Figures 3-6. The SVM models had a slightly lower classification error rate than the RF models, while their sensitivity was significantly higher than latter. No significant difference in the specificity of each prediction model was observed. Both the positive predictive and negative predictive values of the SVM models were significantly higher than those of the RF models. With regard to feature selection, the G-mean value for the expert consultation method was slightly higher than that for the FFS-RF. The AUC value under the SVM was slightly higher than that under the RF. In summary, among the four models tested, F-SVM was the optimal model.

Table 6. Test data sets for each model evaluation index

| Model name | E-RF<sup>a</sup> | E-SVM<sup>b</sup> | F-RF<sup>c</sup> | F-SVM<sup>d</sup> |
|------------|------------------|------------------|------------------|------------------|
| Mis-classification rate | 0.28 | 0.20 | 0.27 | 0.22 |
| Sensitivity | 0.48 | 0.68 | 0.48 | 0.69 |
| Specificity | 0.86 | 0.87 | 0.86 | 0.83 |
| Positive predictive | 0.63 | 0.72 | 0.63 | 0.68 |
| Negative predictive | 0.76 | 0.84 | 0.76 | 0.84 |
| G_mean | 0.84 | 0.76 | 0.64 | 0.76 |
| AUC<sup>e</sup> | 0.75 | 0.81 | 0.70 | 0.78 |

<sup>a</sup>E-RF: Model built using the RF algorithm and expert consultation method
<sup>b</sup>E-SVM: Model built using the SVM algorithm and expert consultation method
<sup>c</sup>F-RF: Model built using the RF algorithm and FFS-RF method
<sup>d</sup>F-SVM: Model built using the SVM algorithm and FFS-RF method

Figure 3-6. ROC curves for the four PPD prediction models
The features selected by the expert consultation method and FFS-RF method were put into the E-RF and F-RF models, respectively. The importance of the features was ranked as shown in Figure 3. The importance of mental elasticity in the model is significantly higher than other factors. Symptoms of depression in late pregnancy was the second most important predictor. Income levels were also important predictors of PPD. There was no significant difference in the importance of each factor to PPD. The top ten most important features in these two models are shown in Table 7.

Figure 2. Ordering of features by importance for the E-RF and F-RF models

(a) Feature factor importance ordering for th (b) Feature factor importance ordering for the F-RF
Model built using the RF algorithm and expert consultation method

\(^b\)F-RF: Model built using the RF algorithm and FFS-RF method

Table 4. Top ten features according to the E-RF and F-RF (Descending order)

| E-RF\(^a\)                          | F-RF\(^b\)                        |
|-------------------------------------|-----------------------------------|
| 1. Level of psychological resilience | 1. Level of psychological resilience |
| 2. Depressive symptoms in the third trimester | 2. Depressive symptoms in early pregnancy |
| 3. Income level                     | 3. Income level                   |
| 4. Husband’s education degree       | 4. Husband’s income level         |
| 5. Education degree                 | 5. Husband’s education degree      |
| 6. Husband’s income level           | 6. Education degree               |
| 7. Mother’s menopausal symptoms     | 7. Mother’s menopausal symptoms   |
| 8. Premenstrual syndrome–mood instability |                                  |
| 9. Marital satisfaction             |                                   |
| 10. Age                             |                                   |

\(^a\)E-RF: Model built using the RF algorithm and expert consultation method

\(^b\)F-RF: Model built using the RF algorithm and FFS-RF method

Discussion

We compared four PPD prediction models and provide a reference for the application of ML in PPD. Compared with the expert consultation method approach, the FFS-RF method identified fewer predictive factors. We found that the F-SVM model was the best model. The strongest predictive factor was the psychological resilience of pregnant women. Between the expert consultation method and FFS-RF methods, the latter selected far fewer predictive factors. Furthermore, there was no significant difference between the two methods in terms of their effects on model performance, indicating that the FFS-RS method could reduce dimensions and improve the efficiency of the algorithmic function without changing model predictive performance. The reduction of the number of predictive factors means that the burden of collecting information is reduced, making the model easier to implement and popularize, especially in busy obstetrics clinics. The SVM was chosen as the better algorithm because it showed higher sensitivity than the RF algorithm (D-SVM = 0.67, F-SVM = 0.69, D-RF = 0.48, F-RF = 0.48). SVM had a clear advantage over RF in processing our research data, and the smaller sample size may be the main reason for this finding. Previous research on depression suggested that sample size is a key factor affecting the
performance of ML models. When the sample size is small, SVM can avoid overfitting while providing efficient computing time and produces better prediction results in depression[29,30]. Our results also support this view. Therefore, we believe that when the dataset is small, SVM is more practical than RF in prediction research for PPD. Several previous studies used the SVM algorithm to make PPD predictions. Jiménez collected data on postpartum women from seven Spanish hospitals and used the EPDS score as the outcome indicator to train a SVM prediction model based on SVM[13]. Sriraam used social media as a data source and, based on the mental health data of 173 mothers, an SVM-based PPD prediction model was established[15]. De Choudhury developed an SVM model to identify high-risk emotions and behaviors predictive of PPD using the content of Twitter posts[31]. As these studies either target different populations or use different methods to detect the occurrence of PPD, the model prediction effects cannot be easily compared. However, the results of the optimal F-SVM model in our study are within range (sensitivity = 0.69, AUC = 0.78) are consistent with the findings of previous studies (sensitivity = 0.56-0.78, AUC = 0.63-0.81)[13,15,31]. Due to the negative effects of PPD on mothers and infants[32,33] such as the negative effects on the physical and mental health of mothers, the closeness of the mother-infant bond, and infant development, it is important to have a model with high sensitivity, while maintaining a high AUC value. The selection of indicators in evaluating depression prediction models varies across studies. For example, Sriraam[15] and De Choudhury[31] emphasized the accuracy of the model's prediction of PPD. Jiménez emphasized model sensitivity and specificity. The balance between them is the G-mean[13]. The AUC is also widely used to evaluate the comprehensive performance of a model[14,15]. Our evaluation criteria provide a reference for prediction research for screening purposes, but the approach may be different in research studies.

We found that the top three most important predictors in models were psychological resilience, depression during the third trimester and income level. First, psychological resilience is the most important factor in the prediction of PPD, which can be attributed to the protective effect of psychological elasticity. Pregnancy and childbirth are challenging time for women emotionally and physiologically and the mother's body and mind are under greater stress[15]. Previous research has shown that psychological resilience as an important regulatory process can enable people to recover from and adapt to stress and life events, reducing the occurrence of adverse outcomes[34-36]. Our finding also supports the findings of Lu, who found that the level of psychological elasticity was negatively correlated with the occurrence of PPD in a study[37]. Second, the results regarding depression in the third trimester are consistent with most previous studies. Depression in the third trimester is associated with PPD[9,38,39]. A review by Robertson mentioned that "depression and anxiety during pregnancy are the strongest predictors of PPD" [5]. Mora's research suggests that depression in the third trimester may continue to develop into the postpartum period[40]. Third, the income levels remains important factors affecting PPD, which supports Rhonda's findings that mothers with low income levels faced obstacles in using mental health resources and were more likely to be frustrated[41]. Epidemiological studies of PPD worldwide have also found that the incidence in developing countries is higher than that in developed countries[42].

The identification of these predictors also reveals the different aspects of PPD risk factors. A pregnant woman's psychological elasticity may reflect her personality traits. Depression in the third trimester may be a special symptom accompanying pregnancy. The income of a pregnant woman and her partner reflects the stability and coping resources available to them. It indicates that PPD risk should be assessed based on a combination of individual long-term, short-term and environmental characteristics.

This study has several limitations. First, there was potential selection bias. Women who were not lost to follow-up might have had a greater awareness of mental health services. Second, the 50% loss to follow-up and small sample size may have negatively affected the applicability of the PPD model, indicating that more extensive validation is required. Third, a larger number of potential predictive factors would have been useful. Further studies should develop different PPD models using other ML...
algorithms and data from different sources as well as incorporating additional cultural factors to expand the application of the PPD models.

Acknowledgments
We acknowledge the people who have contributed to the field of this study, including professor K.K. Cheng from the University of Birmingham, Liu Lu from Central South University. This project is funded by the National Natural Science Foundation of China (Grant No. 81402690; 81773446); the National Natural Science Foundation of Hunan Province (Grant No. 2019JJ40351); and the Graduate Research and Innovation Project of Central South University (Grant No. 1053320183626).

Authors' Contributions
WZ, as the first author, developed the initial manuscript. She helped with recruitment of the participants and collected the data. WZ and HL performed the statistical analysis. HL and VMBS contributed substantially to the revision and refinement of the final manuscript study. WG and PQ guided the overall design of the study, supervised the model development and manuscript. WG and PQ contributed equally to this paper.

References:
1. VanderKruik, R, Barreix M, Chou D, Allen T, Say L, Cohen LS. The global prevalence of postpartum psychosis: a systematic review. BMC Psychiatry 2017.17(1):272. [PMID:28754094] [doi:10.1186/s12888-017-1427-7]
2. Gavin NI, Gaynes, BN, Lohr KN, Meltzer-Brody S, Garthlehner G, Swinson T. Perinatal depression: a systematic review of prevalence and incidence. Obstet Gynecol 2015.106(S Pt 1):1071-83. [PMID:16260528] [doi:10.1097/AOG.0000183597.31630.db]
3. Fisher J, Cabral de Mello M, Patel V, Rahman A, Tran T, Holton S, Holmes W. Prevalence and determinants of common perinatal mental disorders in women in low- and lower-middle-income countries: a systematic review. Bull World Health Organ 2012.90(2):139g-149g. [PMID:22423165] [doi:10.2471/blt.11.091850]
4. Muzik M, Borovska S. Perinatal depression: implications for child mental health. Ment Health Fam Med 2010.7(4):239-47. [PMID:22477948]
5. Robertson E, Grace S, Wallington T, Stewart DE. Antenatal risk factors for postpartum depression: a synthesis of recent literature. Gen Hosp Psychiatry 2004.26(4):289-95. [PMID:15234824] [doi:10.1016/j.genhospsych.2004.02.006]
6. Righetti-Veltema M, Conne-Perreard E, Bousquet A, Manzano J. Risk factors and predictive signs of postpartum depression. J Affect Disord 1998.49(3):167-80. [PMID:9629946] [doi:10.1016/s0165-0327(97)00110-9]
7. Fergusson DM, Horwood LJ, Thorpe K. Changes in depression during and following pregnancy. ALSPAC Study Team. Study of Pregnancy and Children. Paediatric and perinatal epidemiology 1996.10(3):279-93. [PMID:MEDLINE:8822771] [doi:10.1111/j.1365-3016.1996.tb00051.x]
8. Shatte, Adrian BR, Hutchinson DM, Teague SJ. Machine learning in mental health: a scoping review of methods and applications. Psychol Med 2019.49(9):1426-1448. [PMID:30744717] [doi:10.1017/s0033291719000151]
9. Beck CT. Predictors of postpartum depression: an update. Nurs Res 2001.50(5):275-85. [PMID:11570712] [doi:10.1097/00006199-200109000-00004]
10. Johnstone SJ, Boyce PM, Hickey AR, Morris-Yates AD, Harris MG. Obstetric risk factors for postnatal depression in urban and rural community samples. Aust N Z J Psychiatry 2001.35(1):69-74. [PMID:11270460] [doi:10.1046/j.1440-1614.2001.00862.x]
11. Peters J, Verhoest NEC, Samson R, Meirvenne, VM, Cockx L, Baets BD. Uncertainty propagation in vegetation distribution models based on ensemble classifiers. Ecological Modelling 2009.220(6):791-804. [PMID:WOS:000264645000005] [doi:10.1016/j.ecolmodel.2008.12.022]
12. Yao DJ. 2016. Research on feature selection and classification method based on random forest for medical datasets. 2016. [http://cmdm.cnki.com.cn/Article/CMFD-10217-1017245762.htm]
13. Jimenez-Serrano S, Tortajada S, Garcia-Gomez JM. A Mobile Health Application to Predict Postpartum Depression Based on Machine Learning. Teledmed J E Health 212015.(7):567-74. [PMID:25734829] [doi:10.1089/tmj.2014.0113]
14. Jin H, Wu S, Di Capua P. Development of a Clinical Forecasting Model to Predict Comorbid Depression Among Diabetes Patients and an Application in Depression Screening Policy Making. Prev Chronic Dis 2015.12:E142. [PMID:26334714] [doi:10.5888/pcd12.150047]

15. Natarajan S, Prabhakar A, Ramanan N, Baglione A, Connelly K, Siek K. Boosting for Postpartum Depression Prediction. 2nd IEEE/ACM International Conference on Connected Health - Applications, Systems and Engineering Technologies (CHASE): 232-240 2017 JUL. [https://dl.acm.org/citation.cfm?id=3204126]

16. Andrew AM. An Introduction to Support Vector Machines and other kernel-based learning methods. 2000. [doi:10.1108/k.2001.30.1.103.6]

17. Boser BE, Guyon IM, Vapnik VN. ACM Press the fifth annual workshop - Pittsburgh, Pennsylvania, United States 1992 Proceedings of the fifth annual workshop on Computational learning theory A training algorithm for optimal margin classifiers.144-152.

18. Hapfelmeier A, Hothorn T, Ulm K, Strobl C. A new variable importance measure for random forests with missing data. Statistics and Computing 2014.24(1):21-34. [PMID:WOS:000329246300002] [doi:10.1007/s11222-012-9349-1]

19. Costafreda SG, Chu C, Ashburner J, Fu CHY. Prognostic and diagnostic potential of the structural neuroanatomy of depression. PLoS One 2009.4(7):e6353. [PMID:19633718] [doi:10.1371/journal.pone.0006353]

20. Wang Y, Guo X, Lau Y, Chan KS, Yin L, Chen J. Psychometric evaluation of the Mainland Chinese version of the Edinburgh Postnatal Depression Scale. Int J Nurs Stud 2009.46(6):813-23. [PMID:19217107] [doi:10.1016/j.ijnurstu.2009.01.010]

21. Hewitt CE, Gilbody SM, Mann R, Brealey S. Instruments to identify post-natal depression: Which methods have been the most extensively validated, in what setting and in which language? Int J Psychiatry Clin Pract 2010.14(1):72-6. [PMID:24917236] [doi:10.3109/13651500903198020]

22. Cox JL, Holden JM, Sagovsky R. Detection of postnatal depression. Development of the 10-item Edinburgh Postnatal Depression Scale. Br J Psychiatry 1987.150:782-6. [PMID:3651732] [doi:10.1192/bjp.150.6.782]

23. Smith BW, Dalen J, Wiggins K, Tooley E, Christopher P, Bernard J. The brief resilience scale: Assessing the ability to bounce back.15(3):194-200. [doi:10.1080/10705500802222972]

24. Buysse DJ, ReynoldsIII CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Res 1989.28(2):193-213. [PMID:2748771] [doi:10.1016/0165-1781(89)90047-4]

25. Lu TY, Yan L, Ping X, Zhang GQ, Wu DR. Analysis on reliability and validity of the Pittsburgh sleep quality index. Chongqing Medicine 2014(3):260-263. [doi:10.3969/j.issn.1671-8348.2014.03.002]

26. Xiao SY. The theoretical basis and research application of social support rating scale. JOURNAL OF CLINICAL PSYCHOLOGICAL MEDICINE. 1994.(02):98-100. [http://www.cnki.com.cn/Article/CJFDTotal-LCJS402.019.htm]

27. Spitzer RL, Kroenke K, Williams JB, Lowe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006.166(10):1092-7. [PMID:16717171] [doi:10.1001/archinte.166.10.1092]

28. Wang L, Lu K, Ding RJ, Hu DY, Li S. GW25-e4488 The value of Chinese version GAD-7 and PHQ-9 to screen anxiety and depression in cardiovascular outpatients. Journal of the American College of Cardiology. 2014.64(16):C222. [doi:10.1016/j.jacc.2014.06.1038]

29. Patel MJ, Khalaf A, Aizenstein HJ. Studying depression using imaging and machine learning methods. NeuroImage-Clinical 2016.10:115-123. [PMID:WOS:000373189000013] [doi:10.1016/j.nicl.2015.11.003]

30. Malki K, Koritskaya E, Harris F, Bryson K, Herbster M, Tosto MG. Epigenetic differences in monozygotic twins discordant for major depressive disorder. Translational Psychiatry 6 2016. [PMID:WOS:000383417400004] [doi:10.1038.tp.2016.101]

31. Choudhury M De, Count S, Horvitz E. [ACM Press the SIGCHI Conference - Paris, France (2013.04.27-2013.05.02)] Proceedings of the SIGCHI Conference on Human Factors in Computing Systems- Predicting postpartum changes in emotion and behavior via social media.3267. [doi:10.1145/2470654.2466447]

32. Glasheen C, Richardson GA, Fabio A. A systematic review of the effects of postnatal maternal anxiety on children. Arch Womens Ment Health 2010.13(1):61-74. [PMID:19789953] [doi:10.1007/s00737-009-0109-y]

33. Martini J, Petzoldt J, Einsle F, Beesdo-Baum K, Hofler M, Wittchen HU. Risk factors and course patterns of anxiety and depressive disorders during pregnancy and after delivery: a prospective-longitudinal study. J Affect Disord 2015.175:385-95. [PMID:25678171] [doi:10.1016/j.jad.2015.01.012]

34. Lutha, SS, Cicchetti D. The construct of resilience: implications for interventions and social policies. Dev Psychopathol 2000.12(4):857-85. [PMID:11202047] [doi:10.1017/s0954579400004156]

35. Zautra, AJ, Hall JS, Kate EM. Resilience: A new definition of health for people and communities. 2010. [https://
36. Fletcher D, Sarkar M. Psychological resilience: A review and critique of definitions, concepts, and theory. European psychologist 2013.18(1):12-23. [doi:10.1027/1016-9040/a000124]

37. Lu QR, Ding Q, Wang YL, Wang Y. Moderating effect of psychological resilience between prenatal perceived stress and postnatal depression among perinatal women. 2019.33(11):1906-1910. [doi:10.12102/j.issn.1009-6493.2019.11.019]

38. O'hara MW, Swain AM. Rates and risk of postpartum depression—a meta-analysis. International review of psychiatry. 1996.8(1):37-54. [doi:10.3109/09540269609037816]

39. Josefsson A, Angelsiöö L, Berg G, Ekström CM, Gunnervik C, Nordin C, Sydsjö G. Obstetric, somatic, and demographic risk factors for postpartum depressive symptoms. Obstetrics gynecology.2002.99(2):223-228. [doi:10.1016/s0029-7844(01)01722-7]

40. Sutter-Dallay AL, Cosnefroy O, Glatigny-Dallay E, Verdoux H, Rascle N. Evolution of perinatal depressive symptoms from pregnancy to two years postpartum in a low-risk sample: the MATQUID cohort. J Affect Disord 2011.139(1):23-9. [PMID:22410506] [doi:10.1016/j.jad.2011.08.018]

41. Boyd RC, Mogul M, Newman D, Coyne JC. Screening and Referral for Postpartum Depression among Low-Income Women: A Qualitative Perspective from Community Health Workers. Depress Res Treat 2011:320605. [PMID:21603131] [doi:10.1155/2011/320605]

42. Gulamani SS, Premji SS, Kanji Z, Azam SI. A review of postpartum depression, preterm birth, and culture. J Perinat Neonatal Nurs 2013.27(1):52-9; quiz 60-1. [PMID:23360942] [doi:10.1097/JPN.0b013e318277cf24]

Abbreviations:
PPD:Postpartum depression
SVM:support vector machines
RF:Random Forest
FFS-RF:Filter feature selection algorithm based on random forest
E-RF:Model built using the RF algorithm and expert consultation method
F-RF:Model built using the RF algorithm and FFS-RF method
E-SVM:Model built using the SVM algorithm and expert consultation method
F-SVM: Model built using the SVM algorithm and FFS-RF method
EPDS: Edinburgh Postnatal Depression Scale
BRS: Brief Resilience Scale
PSQI: Pittsburgh Sleep Scale
SSRS: Social Support Rating Scale
GAD-7:Generalized Anxiety Scale
Supplementary Files
Figures
The relative feature importance rankings of the E-RF and the F-RF based on the two feature selection methods.
The receiver operating characteristic curve of E-RF. AUC: area under the curve; ROC: receiver operating characteristic.
Participant recruitment and response condition.
The receiver operating characteristic curve of E-SVM. AUC: area under the curve; ROC: receiver operating characteristic.
The receiver operating characteristic curve of F-RF. AUC: area under the curve; ROC: receiver operating characteristic.
The receiver operating characteristic curve of F-SVM. AUC: area under the curve; ROC: receiver operating characteristic curve.
Other materials for editor/reviewers onlies
Cover letter to editor.
URL: https://asset.jmir.pub/assets/50f2ba4dabb0f908716882b3bef00d70.pdf

Original comments from Erin Bankes.
URL: https://asset.jmir.pub/assets/88bb5daf1159550cffe1615fca2515f6.pdf
Multimedia Appendixes

Comparison of demographic characteristics, including data sets of 618 pregnant women lost in the cohort and 508 mothers who left the cohort study after childbirth.
URL: https://asset.jmir.pub/assets/a7902bd184f026d49a993c4bbe6dee48.docx

Definitions and coding of analyzed variables.
URL: https://asset.jmir.pub/assets/9083e3319c9d9d9129f73d60c034f51f.docx
Other materials for editor/reviewers onlies

Editing Certificate.
URL: https://asset.jmir.pub/assets/89aa7bfa6e55296115ac54a5579ff36.pdf

Reply to editors and reviewers.
URL: https://asset.jmir.pub/assets/7be634fe6d79e86c270d4b1be136722.docx