Cureus Journal of Computer Science
https://doi.org/10.7759/s44389-026-00057-3

Cureus
Part of SPRINGER NATURE

Original Article

Assessing the Association Between Maternal Age and Postpartum Depression Using Statistical Analysis and Explainable Artificial Intelligence

Trisha Roy Chowdhury [1], Dhrubasish Sarkar [2], Moumita Chatterjee [3]

1. *Department of Information Technology (Data Science), Maulana Abul Kalam Azad University of Technology, Kolkata, IND*
2. *Department of Computer Science and Engineering, NSHM College of Management & Technology, Kolkata, IND*
3. *Department of Computer Science and Engineering, Aliah University, Kolkata, IND*

Received: March 05, 2026 | Review began: March 16, 2026 | Review ended: April 03, 2026 | Published: April 13, 2026

Abstract

Postpartum depression (PPD) can feel like a curse for a new mother in an otherwise happy family after childbirth. PPD can occur due to a wide range of factors, including physical and psychosocial issues. This research focuses on middle-aged mothers (40-45 years) to examine the prevalence of PPD. It further investigates whether age plays a crucial role in the occurrence of PPD. Although the initial dataset was unclassified, a popular unsupervised algorithm, K-means, was used to cluster the data. The result of the cluster for mothers aged 40-45 was further cross-checked using several statistical tests, including the chi-square test, the Mann-Whitney U-test, and logistic regression with odds ratios. While the Mann-Whitney U-test showed no significant difference in the severity of PPD between younger and older mothers, the chi-square test and logistic regression indicated that the age group 40-45 is a moderately crucial period for motherhood due to the risk of PPD. Finally, explainable AI was applied using a supervised algorithm (random forest) to assess the influence of age and other symptoms. This study clearly shows which symptoms are mainly responsible for PPD and which have a moderate effect. An accuracy of approximately 94% was achieved, with key symptoms including feelings of guilt, appetite problems, and irritability. The model achieved precision, recall, and F1-scores of 91.77%, 95.30%, and 93.50%, respectively. The receiver operating characteristic score was 98.40%.

Categories: AI/ML-based decision support systems, Explainable AI, Machine Learning (ML)

Keywords: postpartum depression, machine learning, unsupervised learning, clustering, statistical test, middle age, supervised learning, explainable ai, symptoms

Introduction

Every mother has the equal right to live a healthy life. When the matter is about taking care of a newborn with the mother herself, it sometimes becomes turmoil. Postpartum depression (PPD) plays a catalyst role in deteriorating the situation. PPD is a treatable mental health disorder among new mothers. This study found that in low- and middle-income countries, the prevalence of PPD ranges from 6.5% to 12.9% or higher [1]. It includes sadness, loss of hope, anxiety, sleep disturbances, appetite changes, trouble concentrating, loss of interest, suicidal tendencies, and sometimes harming the baby. Though hormonal changes may be an initial reason for PPD, there are several psychosocial factors involved here. Previously untreated depression or poor mental health are also important factors in PPD. PPD is a major depressive disorder. According to researchers, the rising tendency of PPD begins before delivery [2]. Researchers stated that taking care of a newborn with older children throughout the day, often in the middle of the night to take care of the newborn, keeping up household responsibilities, and returning to work from maternity leave are burdens that affect the physical and mental health of the new mother [2]. In addition, factors such as male infant preference, cultural expectations, conflict with traditional postpartum rituals, family dynamics, poverty, loss of autonomy, and low social support can act as catalysts to trigger PPD.

In our current study, we have focused on the 17 Sustainable Development Goals given by the United Nations (2015). Our research area plays an important role in mental health, mainly maternal mental health. From Goal 3 of the Sustainable Development Goals, i.e., “Ensure healthy lives and promote well-being for all at all ages”, we have chosen target 3.4, i.e., “promote mental health and well-being” [3]. The first four weeks after delivery are a very crucial and important period for new mothers [2]. Due to hormonal fluctuations, psychosocial reasons, and economic factors, PPD may develop. A new mother needs social support, economic support, and most importantly, freedom from health issues to lead a happy life with her newborn and family. First, it should be known that there is a difference between “Baby Blues” and “Postpartum Depression”. “Baby Blues” or “Postpartum Blues” refer to temporary sadness, tearfulness, or mood swings in a new mother. It may last up to two weeks after onset. However, when it comes to “Postpartum Depression”, it is very important to seek care. Medical treatment and psychological counselling are required for recovery. Sometimes, the outcome of PPD can be very tragic, leading to suicidal tendencies in the mother or incidents such as infanticide [2]. According to WHO [4], globally 13% of women suffer from PPD after childbirth, whereas 19.8% of mothers in developing countries suffer from PPD. Mothers who experience domestic violence, including physical, mental, or sexual abuse, are in a more vulnerable state than other new mothers. Sometimes, career-focused women become mothers after the age of 30, which may be considered middle age. Middle age is a crucial time for women to become mothers because the end of the fertile period approaches, and on the other hand, the urge or wish to become a mother (sometimes for the first time) is also significant.

Our research focus is to check whether age is really an important parameter in triggering PPD or not. Nowadays, machine learning and statistics combinedly have made significant breakthroughs in the medical field. Both physical health problems and mental health issues can be easily diagnosed using subfields of artificial intelligence (AI), and through statistical methods, hidden patterns of disorders (both physical and mental) can be easily identified.

Several conventional methods exist to recognize PPD. The most common method is the Edinburgh Postnatal Depression Scale. It consists of 10 questions to answer and excludes common symptoms of depression during pregnancy and the postpartum period. Although it is available in 50 different languages, this scale is considered reliable and efficient for diagnosing PPD. Other popular methods include the Patient Health Questionnaire-9, the Postpartum Depression Screening Scale, and the Beck Depression Inventory I and II [5,6]. In this technology era, we depend on mobile apps and websites. Due to this, machine learning-based and deep learning-based diagnosis is very important. Several researchers have used various machine learning techniques for PPD screening purposes. Many of them work on different datasets across populations, while others use the same dataset but different techniques. Gopalakrishnan et al. [7] used extremely randomized trees to predict PPD. Qi et al. [8] employed SHAP (SHapley Additive exPlanations) values to make their predictions on PPD interpretable. Prenatal risk factors were detected by Sibbald et al. [9] using various machine learning models, where Lasso-regularized linear regression was the best performer. Amit et al. [10] used various machine learning techniques and evaluated model performance using different validation techniques. They cross-verified the results using the Edinburgh Postnatal Depression Scale. Extreme Gradient Boosting (XGBoost) was the best-performing model among logistic regression, support vector machine, random forest (RF), light gradient boosting machine, and multilayer perceptron (MLP) in this research [11]. Zhang et al. [12] used RF, decision tree, XGBoost, regularized logistic regression, and MLP as classification models for two kinds of datasets, and regularized logistic regression showed the best performance among the algorithms. Research on interpretable AI was done by Zhang et al. [13]. They studied 2,055 participants, i.e., mothers, to predict PPD using SHAP on an XGBoost model to explain the predictions. Andersson et al. [14] analyzed data collected from 2009 to 2018. First, they divided the data into previously mental health issues and non-previously mental health issues. After that, several machine learning models were applied to the datasets. Extremely randomized trees showed the best performance among the datasets. Women with a previous mental health history dataset achieved 64% accuracy, whereas the non-previous mental health history dataset achieved 73% accuracy.

How to cite this article:
Roy Chowdhury T, Sarkar D, Chatterjee M (April 13, 2026) Assessing the Association Between Maternal Age and Postpartum Depression Using Statistical Analysis and Explainable Artificial Intelligence. Cureus J Comput Sci 3 : es44389-026-00057-3.
DOI https://doi.org/10.7759/s44389-026-00057-3

A nationwide study was conducted using data from 2008 to 2015. The gradient-boosted decision tree algorithm was applied to data that had sociodemographic, clinical, and obstetric features [15]. Wang et al. [16] worked on a dataset of Chinese women to examine the prevalence of PPD. A total of 3,174 women participated in the study to assess PPD at various stages after childbirth from October 2021 to January 2023. RF was the most effective model in this study. Matsuo et al. [17] conducted research on Japanese postpartum mothers. Although their machine learning models, such as ridge regression and elastic net, did not perform well like conventional logistic regression, they contributed to identifying high-risk PPD cases. Raisa et al. [18] collected data from 150 postpartum women in Bangladesh. They applied several machine learning models, where RF performed best with 89% accuracy. According to their study, the prevalence of PPD in Bangladesh was 66.7%. Kim [19] applied multiple machine learning algorithms, including decision tree classifier, RF classifier, AdaBoost classifier, and logistic regression, to classify the prevalence of PPD among Korean women. The researcher found that conflict with a partner and stress level were the strongest predictors of PPD. Zhang et al. [20] studied an AI-based clinical decision support system for PPD prevention, diagnosis, and management. L2-regularized logistic regression was trained on electronic health records, and Microsoft Azure was used to provide a scalable, secure, and efficient operational framework. Prabhashwaree and Wagarachchi [21] worked on Sri Lankan postpartum mothers. They used feed-forward neural network (FFANN), adaptive neuro-fuzzy inference system with genetic algorithm (ANFIS-GA), RF, and support vector machine for classification into four categories of risk. Although the best-performing models were the FFANN model (97.08% accuracy) and the ANFIS-GA model (testing error: 0.0496), the highest performance for multi-class classification was achieved by the FFANN model [21]. Shivaprasad et al. [22] also worked using explainable AI (XAI) to identify risk factors of PPD. In their dataset, they targeted the "suicide attempt" feature as the target variable. They used four XAI methods, SHAP, LIME (Local Interpretable Model-agnostic Explanations), ELI5, and Anchor, on k-nearest neighbor and a stacking model. For both cases, the accuracy achieved was 97% [22]. Anand et al. [23] also worked on the same dataset, but their target variable was "feeling anxious" to determine the prevalence of PPD. They used eight machine learning models individually and then applied an ensemble model combining all eight models. The ensemble model achieved the highest accuracy of 94.31%. Suganthi and Geetha [24] introduced a new technique called Osprey Parameter Optimized MLP. They collected data from social networking sites like Twitter (tweets) and Instagram (Comments). Natural language processing was used to pre-process the extracted data. Linguistic inquiry word count and latent semantic analysis were used to extract additional features like mental health and behavioral changes among women. The accuracy of Osprey Parameter Optimized MLP model was 96.12% on the collected dataset. The novel Meta DT-KN-RF model was introduced by Nasim et al. [25]. The researchers named this model as it is a combined version of decision tree, K-nearest neighbor, and RF. The model achieved 99% accuracy in predicting anxiety, where "Feeling Anxious" was the target variable. A 10-fold cross-validation approach was applied in this study.

An accuracy of 87.95% was reported by Fischbein et al. [26]. The ensemble method, named "Super Learner," was a hybrid model combining conditional RF and stochastic gradient boosting. The researchers used data from the Pregnancy Risk Assessment Monitoring System for this purpose. Lin et al. [27] developed a calculator to predict the prevalence of PPD. The calculator was based on machine learning algorithms such as RF and XAI. SHAP was used to explain the model's outputs. The area under the receiver operating characteristic curve was 91%. Finally, risk stratification was performed using K-means clustering analysis.

We have used a publicly available dataset from the Kaggle repository [28] to find the hidden patterns of mothers diagnosed with PPD. By analyzing clinical and psychosocial parameters, PPD-affected mothers and healthy mothers were first clustered. After that, based on the results, several statistical tests were conducted to check the reliability of the identified patterns. Finally, we classified the clusters using XAI techniques to ensure trust in the decisions and outputs [29-31].

Study gap

The existing research studies separately identify age, family and social support, delivery-related risks, and clinical symptoms as risk factors of PPD. However, these studies are limited to classical regression or score-based cut-offs, where the pattern behind the symptoms is not considered, and clustering is also not addressed. Although some studies focus on PPD in middle-aged (40-45 years) women, other age groups are often categorized into multiple age groups. Moreover, XAI has not been specifically used to analyze the relationship between age, symptoms, and their severity in detecting PPD. On the other hand, although some studies have attempted to identify multiple subtypes of PPD using cluster analysis, they did not validate the findings using statistical tests such as the chi-square test, Mann-Whitney U-test, or logistic regression with odds ratio (OR).

In this research, unsupervised learning (K-means clustering), inferential statistics (chi-square test, Mann-Whitney U-test, and logistic regression with OR), and XAI (RF, SHAP, LIME) are combined to examine whether age, particularly middle age (40-45 years), plays a crucial role in detecting PPD or not. Additionally, if age is not the main factor, the study analyzes and discusses which factors play a more significant role in predicting whether a mother is experiencing depression.

Materials And Methods

Data collection and preprocessing

The proposed methodology (Figure 1) shows the complete flow diagram of our implementation of this study. The dataset is collected from the Kaggle public repository [28]. Initially, there were 1,503 instances in the dataset. A Google Form for survey purposes was distributed by the data collector to 1,503 new mothers. The respondents answered questionnaires that included clinical and sociopsychological symptoms. Although there were initially 15 features and 1 target variable, i.e., "feeling anxious", finally nine features were considered as final features for clustering, and "feeling anxious" was considered separately. This is because anxiety and depression (PPD) are different conditions, and a mother may or may not have an anxiety problem while suffering from PPD. Among the 1,503 data records, some instances contained one or more null values, as there was a non-mandatory response format. These records were discarded, and only complete records were considered. Finally, 1,490 records were used for clustering and other analyses. The dataset contained categorical responses such as Yes, No, Sometimes, Two or more days a week, Often, Not at all, Maybe, Not interested to say, 25-30, 30-35, 35-40, 40-45, and 45-50. Extreme responses such as "Yes" were encoded as 3, while 2 was assigned to "Sometimes," "Maybe," "Not interested to say," "Often," and "Two or more days a week." Encoding of "No" and "Not at all" was handled carefully; in "overeating and loss of appetite," "No" was encoded as 2 and "Not at all" as 1, whereas in other features, "No" was encoded as 1. The age groups were encoded from 1 to 5, where 1 represents 25-30 years, 2 represents 30-35 years, 3 represents 35-40 years, 4 represents 40-45 years, and 5 represents 45-50 years. The variable "feeling anxious" was binary encoded, where 1 represents "yes" and 0 represents "no." The dataset was stored in tabular format as a CSV file, and the "Timestamp" column was removed as it was unnecessary during preprocessing. The processed dataset was stored in the local machine for final processing purposes.

How to cite this article:
Roy Chowdhury T, Sarkar D, Chatterjee M (April 13, 2026) Assessing the Association Between Maternal Age and Postpartum Depression Using Statistical Analysis and Explainable Artificial Intelligence. Cureus J Comput Sci 3 : es44389-026-00057-3.
DOI https://doi.org/10.7759/s44389-026-00057-3

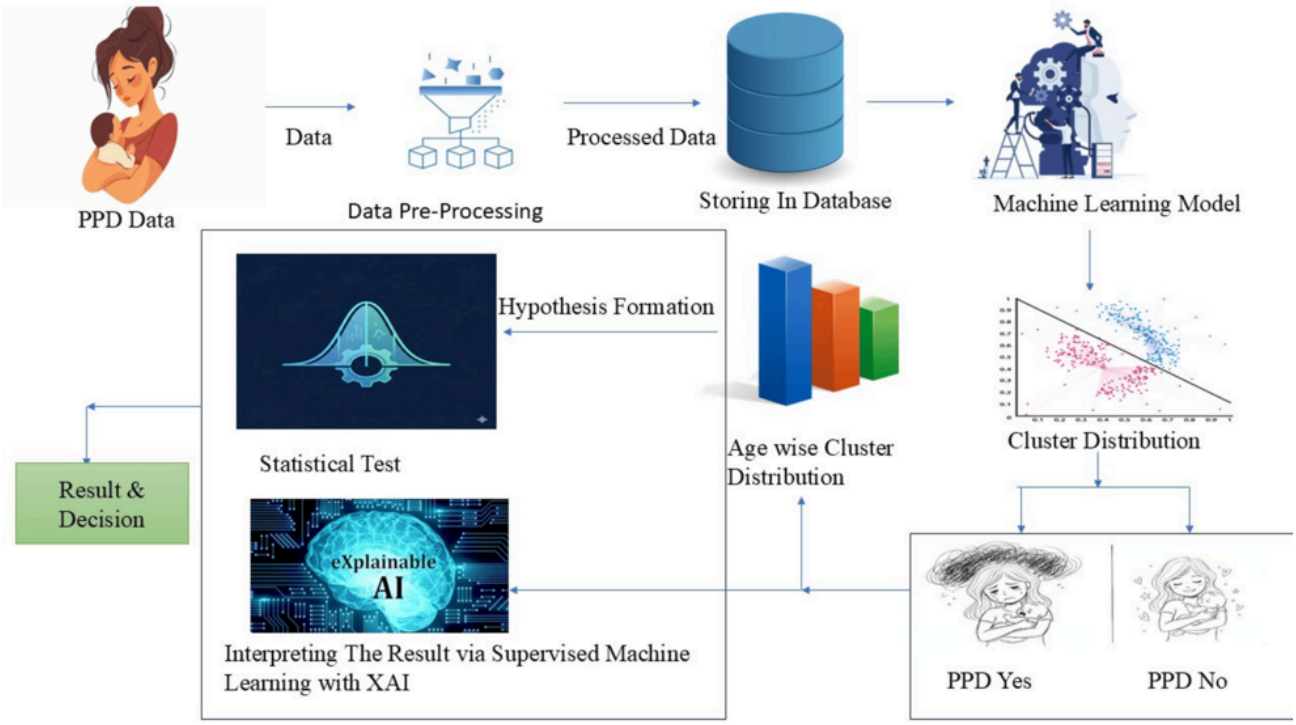

FIGURE 1: Proposed methodology

PPD, Postpartum Depression; XAI, Explainable Artificial Intelligence

Unsupervised machine learning

K-means is an unsupervised algorithm that divides the dataset into *K* number of clusters, where each cluster has an individual centroid. The same cluster points are similar to each other and dissimilar to another cluster of points. Let us have *n* number of data points:

$$X = \{x_1, x_2, \ldots, x_n\}, \quad x_i \in \mathbb{R}^d \tag{1}$$

We want to divide the points into *K* numbers of clusters:

$$C_1, C_2, \ldots, C_K \tag{2}$$

In this research, we determine the value of $K = 2$ using the elbow method (a graphical method). The elbow point decides the number of clusters, as it is the optimal value of *K*. The main concept is that as *K* increases, the within-cluster error decreases rapidly, andnd after that, the decrease slows down; that point is the elbow point (Figure 2).

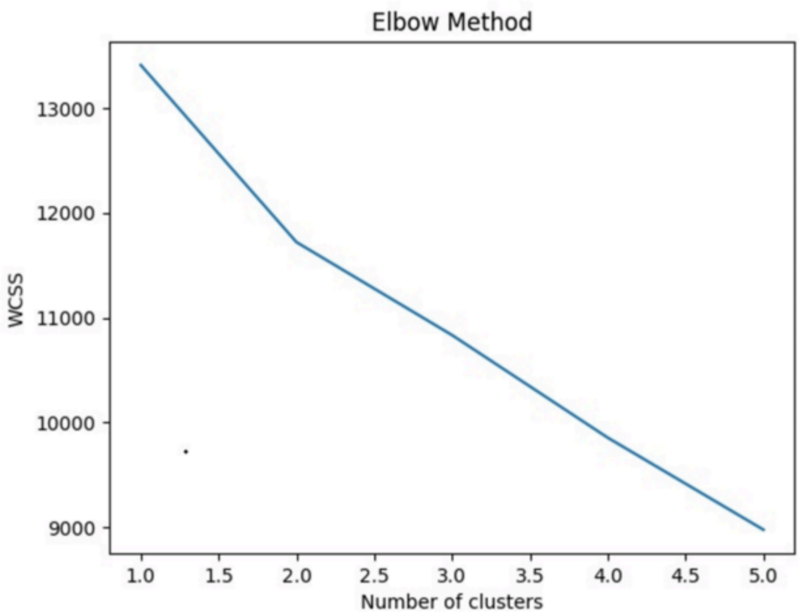

FIGURE 2: Elbow point

The principal component analysis algorithm is used to plot the three-dimensional data by reducing the dimension into two dimensions. It is used for dimensionality reduction purposes. After the dimension of data is reduced, it is visualized by the scatter plot. This algorithm is very popular for visualizing the cluster of the K-means algorithm (Figure 3).

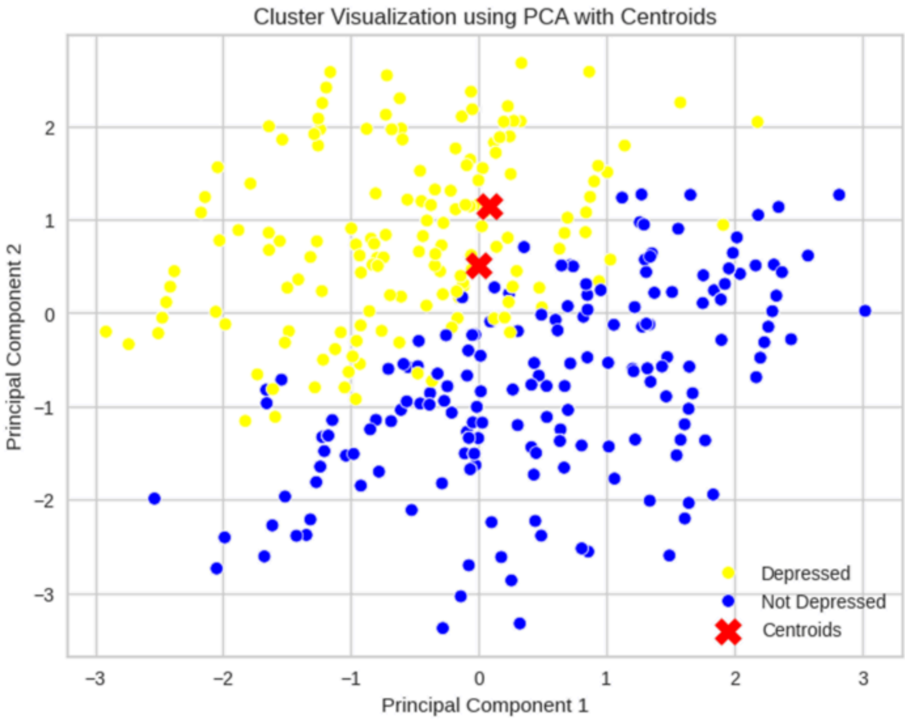

FIGURE 3: The cluster visualization of K-means algorithm

PCA, Principal Component Analysis

Figure 4 shows the age-wise cluster distribution of "PPD yes" and "PPD no". It is clearly visible that in the 40-45 age group, the number of "PPD yes" cases is higher than the number of "PPD no" cases.

How to cite this article:
Roy Chowdhury T, Sarkar D, Chatterjee M (April 13, 2026) Assessing the Association Between Maternal Age and Postpartum Depression Using Statistical Analysis and Explainable Artificial Intelligence. Cureus J Comput Sci 3 : es44389-026-00057-3.
DOI https://doi.org/10.7759/s44389-026-00057-3

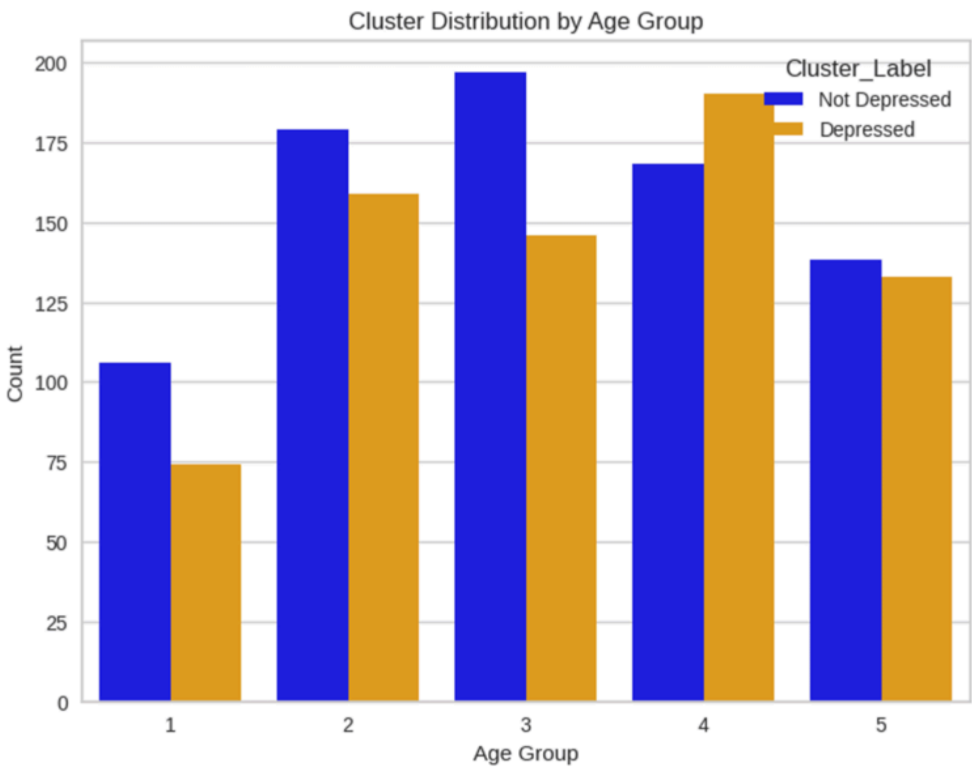

FIGURE 4: Age-wise cluster distribution (yellow = PPD yes, blue = PPD no)

PPD, Postpartum Depression

The binary cluster variables “PPD yes” and “PPD no” are encoded as 1 and 0, respectively, and the results are stored in a separate CSV file for further processing in this study. From this point onward, the research work is carried out using this new file. We focus on mothers aged 40-45 years (previously encoded as 4), as they belong to the second-to-last phase of pre-menopause and also to the middle-age group from a sociopsychological perspective, where the 45-50-year age group is considered the first group. However, our target group is 40-45 years of age, which is encoded as 1, while other age groups are encoded as 0. A composite PPD score is calculated by summing all nine symptom items. These operationalizations reflect the pathway to an ordinal symptom scale in postpartum mental health research.

Hypothesis

Symptom Score Hypothesis

H_0: There is no difference between the median PPD score of middle-aged, i.e., 40-45-year-old, mothers and other-aged (both younger and older) mothers.

$$P(PPDyes|Age = 1) = P(PPDyes|Age \neq 1) \tag{3}$$

H1: The 40-45-year-old mothers are more likely to be affected by PPD, i.e., the PPD score of middle-aged mothers is higher than that of other age groups (both younger and older).

DOI https://doi.org/10.7759/s44389-026-00057-3

$$P(PPDyes|Age = 1) > P(PPDyes|Age \neq 1) \tag{4}$$

Statistical tests

Chi-Square Test

It is a very popular statistical test used to validate a hypothesis. This test examines the relationship between two variables. In this research, the first variable is the age group, where two categories are considered: “40-45-year-old mothers vs. other-aged mothers,” and the second variable is “PPD yes vs. PPD no.” First, a 2×2 contingency matrix is constructed, where each cell represents the observed frequency. According to the theory, the number of cases expected if the two variables were completely independent of each other is called the expected frequency. The chi-square test measures the difference between observed and expected frequencies. The formula is as follows:

$$\chi^2 = \sum \frac{(O - E)^2}{O} \tag{5}$$

where O = observed and E = expected. This χ^2 value is evaluated against a table or p-value according to the specific degrees of freedom and significance level (e.g., 0.05). If the p-value is small (below 0.05), it is said that there is a statistically significant relationship between the two variables; if the p-value is large, it is assumed that the observed difference in the data is due to chance, and no real association is proven. Here, $\alpha = 0.05$ is the significant threshold.

Mann-Whitney U-Test

The Mann-Whitney U-test is a non-parametric test used to determine whether the scores or values of two independent groups are significantly different from each other. This test is typically used when the data cannot be assumed to be normally distributed or when a t-test is inappropriate due to unequal variances. All observations are combined and ranked from smallest to largest, and it is then observed which group accumulates more of the higher ranks. If the p-value is small (i.e., < 0.05), it is concluded that the distributions or medians of the two groups are different; a large p-value indicates that there is no significant difference.

Logistic Regression With OR

Logistic regression is a statistical model used to analyze how a binary outcome (such as depressed = 1, not depressed = 0) depends on a predictor variable, and its results are typically interpreted in terms of OR. The model first estimates the log-odds of depression using the logit function, and then the OR is obtained by exponentiating the coefficient (eβ). An OR = 1 indicates no difference between the two groups; an OR > 1 means the odds of the outcome are higher in the specified group, and an OR < 1 means they are lower.

Machine Learning and XAI

RF is trained to predict “PPD yes” and “PPD no” from age along with nine other symptom variables using a stratified train-test split (70:30) and class-balanced weighting. SHAP is used to globally examine the important features for predicting PPD. It also identifies the significance of age as a predictor. LIME is applied to a record of women aged 40-45 diagnosed with “PPD yes.” LIME presents the prediction in a human-readable format for each variable, allowing evaluation of how strongly age, relative to other features (i.e., symptoms), drives the predictions toward the “PPD yes” category.

Results And Discussion

Here is the sample constructed by 1,490 mothers. Table *1* and Table *2* show the age-wise distribution of participants and cluster-wise distribution of participants, respectively. We can see that the 25-30 age group (the youngest group in this study) and the 45-50 age group (the oldest group in this study) have the lowest and second lowest number of participants, respectively. The sociological reasons behind this pattern may be that younger women are busy with their careers and studies and do not want to take on the responsibility of a child early. For the older age group, the reasons may include physiological and sociological factors. Many women in the 45-50 age group may try to have a child but are unable to do so due to physiological issues such as early menopause, hormonal problems, and PCOS. On the other hand, the 45-50 age range is often a peak career stage, and at that time a woman may wish to become a mother, but career demands may not allow her to pursue that wish. Interestingly, our research-focused group, i.e., 40-45 years, has the highest number of participants, i.e., 24.0%. This is a relatively balanced age group, neither too early nor too late for motherhood, which may explain why it has the highest participation.

	PPD no	PPD yes
Other Ages	620	512
Age 40–45	168	190

TABLE 1: 2×2 contingency matrix

PPD, Postpartum Depression

Age Code	Age Group	No. of Participants	Percentage (%)
1	Age = 1 (25-30 years old)	180	12.1 %
2	Age = 2 (30-35 years old)	338	22.7 %
3	Age = 3 (35-40 years old)	343	23.0 %
4	Age = 4 (40–45 years old)	358	24.0 %
5	Age = 5 (45-50 years old)	271	18.2 %
Total No. of Participants		1,490	100.0 %

TABLE 2: Number of participants in each age group

Association between age 40-45 years and PPD cluster

From Table *3*, which shows a 2×2 contingency matrix, it is found that 190 mothers out of 358, i.e., 53.1%, are clustered in the PPD-diagnosed group. However, 512 of 1,132 mothers, i.e., 45.2%, from other age groups are affected by PPD. The chi-square test was significant, with $\chi^2 = 6.403$, degrees of freedom (df) = 1, and a p-value = 0.0114, which is less than 0.05. That is, the p-value is below the significance threshold. It indicates that we can strongly reject the null hypothesis (H_0) and accept the alternative hypothesis (H_1). This means that 40-45-year-old mothers have a higher tendency to be diagnosed with PPD. We can say that this middle-age group (40-45 years) is a risk group for postnatal mental health.

	PPD no	PPD yes
Other Ages	620	512
Age 40–45	168	190

TABLE 3: 2×2 contingency matrix

PPD, Postpartum Depression

From Table 4, it is clearly visible that the OR > 1, i.e., 1.370, for the age group 40-45, with a p-value = 0.0097 (approximately 0.01), which is less than 0.05. The 95% confidence interval for the OR is above 1 (1.079-1.738). This indicates that the data does not support the idea that the OR could be 1. Therefore, the effect of Age_40_45 is indeed positive. It implies that mothers aged 40-45 years have a 37% higher chance of being in the "PPD Yes" group compared to the older and younger groups combined. This effect is statistically significant but moderate in magnitude.

Final Result - Age 40-45 Effect (Simple Logistic Regression)
Odds Ratio = 1.370
95% CI = [1.079, 1.738]
p-value = 0.0097

TABLE 4: The output of simple logistic regression with odds ratio

Symptom severity across age groups

The results of the Mann-Whitney U-test given in Table 5 clearly show that the total depression symptom score of mothers aged 40-45 is not different from that of mothers in other age groups. In the 40-45 age group, the median value of the "Postpartum Depression Score" among 358 mothers is 16.00, and the median is exactly the same, 16.00, for the remaining 1,132 mothers (all other ages). This indicates that both groups are at approximately the same level in terms of median symptom severity. The Mann-Whitney U statistic was 197995.0, and the one-sided p-value is 0.7456. Since this p-value is much greater than 0.05, there is no statistical evidence that the scores of the 40-45 age group are higher than those of other age groups. In simple terms, no significant difference was found between the total depression scores of the two groups; therefore, although being 40-45 years old slightly increases the risk of being classified in the PPD cluster, the overall symptom burden is not different from that of mothers of other ages.

Median Age 40–45: 16.00 (N = 358)
Median Others: 16.00 (N = 1,132)
Mann–Whitney U-test = 197995.0, one-sided p = 0.7456

TABLE 5: Output of Mann–Whitney U-test

SHAP global importance

The RF classifier shows reasonable discrimination between "PPD yes" and "PPD no" clusters. From Figure *5*, it is clearly visible that "feeling of guilt," "irritable towards baby & partner," "overeating and loss of appetite," "feeling sad or tearful," "suicide attempt," "feeling anxious," "problems concentrating or making decisions," and "age" are ranked from highest to lowest value. "Age" exhibits the smallest SHAP values and is placed at the bottom of the importance ranking. This indicates that there is minimal global contribution to the model output compared with other symptom features.

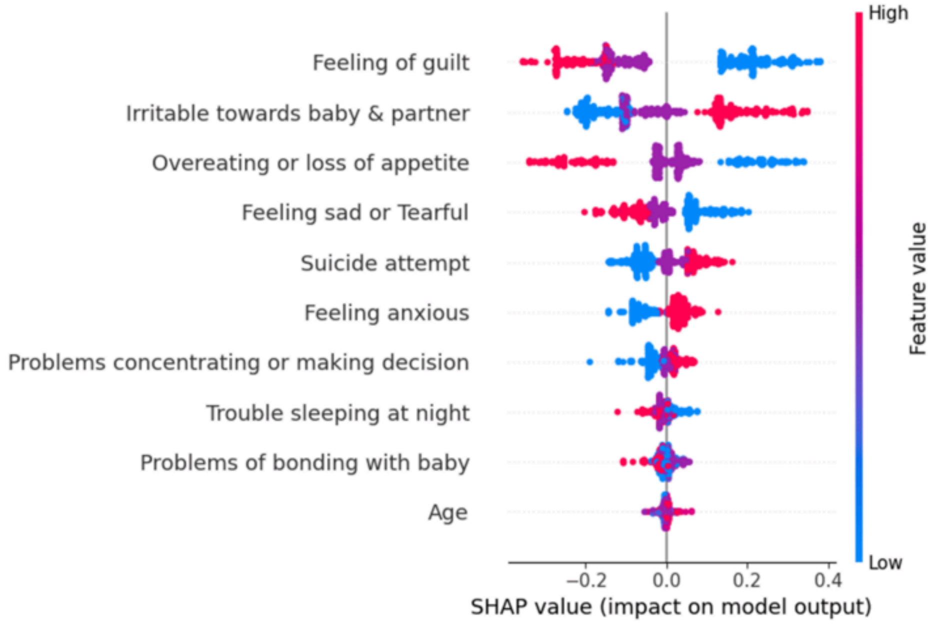

FIGURE 5: SHAP value (the impact on the model output)

SHAP, SHapley Additive exPlanations

LIME local explanation for 40-45-year-old PPD-diagnosed mother

For a PPD-diagnosed 40-45-year-old mother, LIME weights are shown in Figure 6. LIME attributed positive contributions to high levels of guilt, appetite disturbance, irritability, sadness, suicide attempts, and problems with concentration. "Age" appeared as a small positive weight, i.e., 3.00 < Age ≤ 4.00. This result confirms that age pushes the prediction toward the "PPD yes" class, but its effect is substantially outweighed by other sociopsychological and clinical symptom contributions.

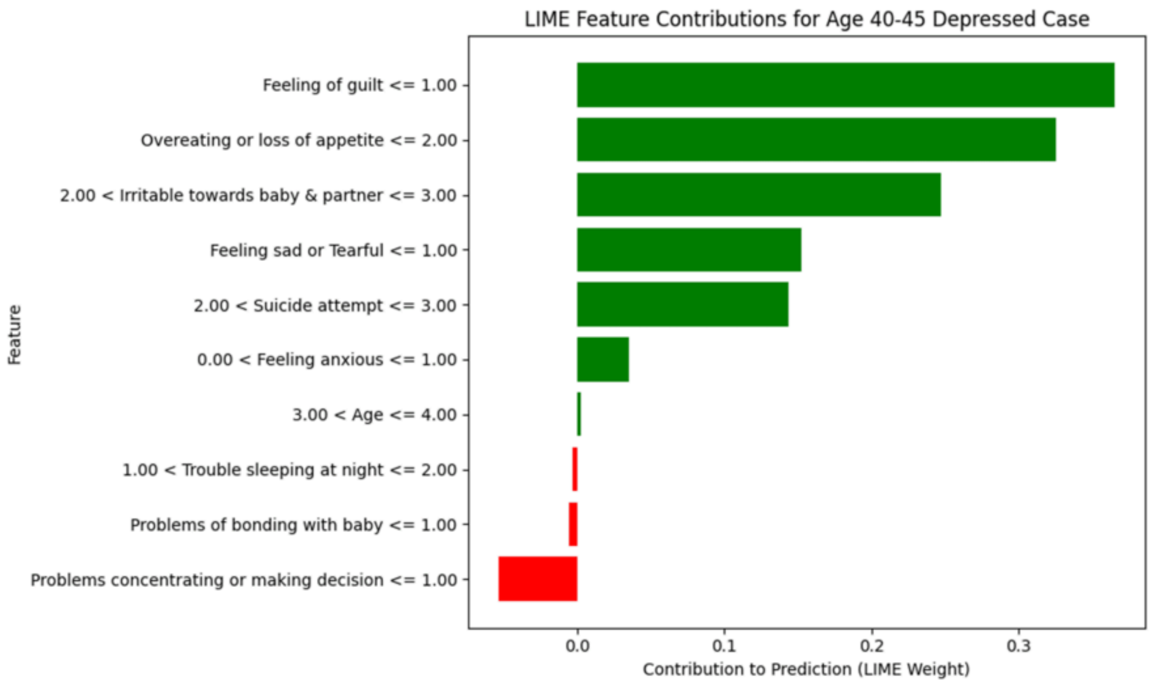

FIGURE 6: LIME (Local Interpretable Model-agnostic Explanations) weight values

LIME, Local Interpretable Model-agnostic Explanations

Random forest classification with top 3 features

Finally, we classify "PPD yes" and "PPD no" classes only by the top 3 predictor symptoms, which are "Feeling guilty," "Overeating and loss of appetite," and "Irritable towards baby & partner," and the result achieves a very good score. The accuracy of that prediction is 0.9376 ± 0.0130 or 94%. The precision, recall, and F1-score are 0.9177 (91.77%), 0.9530 (95.30%), and 0.9350 (93.50%), respectively.

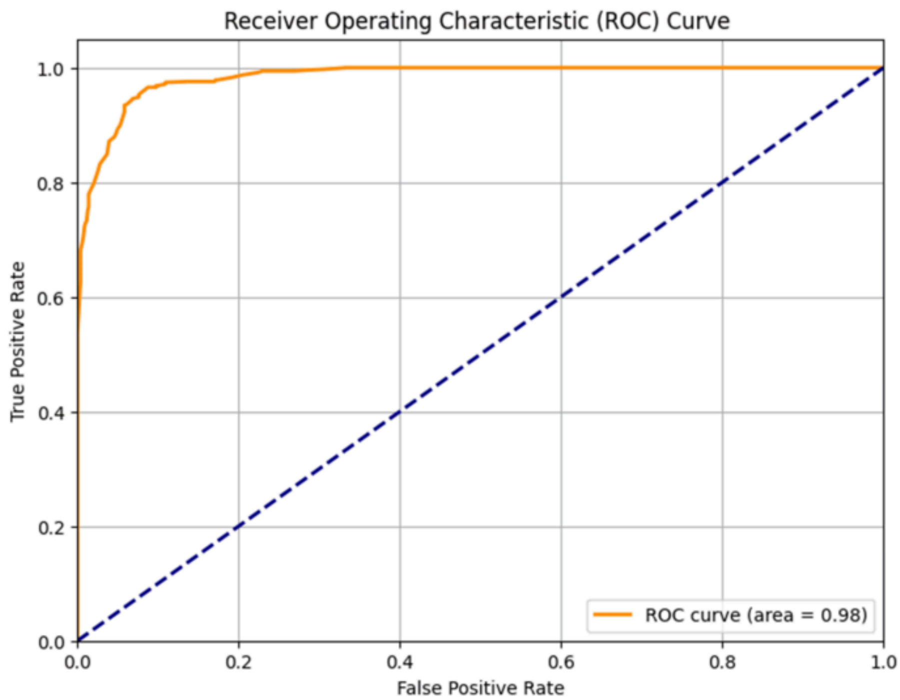

FIGURE 7: Receiver operating characteristic (ROC) curve

It can be seen from Figure 7 that the area under the ROC curve is 0.9840, which is very high and indicates that the final model excellently discriminates between "PPD yes" and "PPD no" mothers using only the top 3 symptoms.

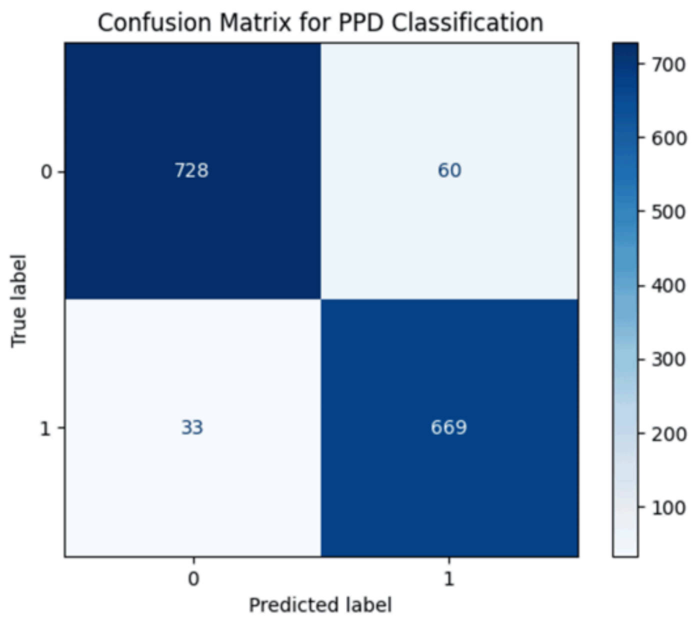

FIGURE 8: Confusion matrix of PPD classification by top 3 predictors

PPD, Postpartum Depression

From the confusion matrix of the final classification with the top 3 symptoms shown in Figure 8, the following results were obtained: True negatives: 728 correctly predicted as not depressed (Class 0, i.e., “PPD no”). False positives: 60 were incorrectly predicted as depressed (Class 1, i.e., “PPD yes”) when they were not depressed. False negatives: 33 were incorrectly predicted as not depressed (Class 0, i.e., “PPD no”) when they were depressed. True positives: 669 correctly predicted as depressed (Class 1, i.e., “PPD yes”). Here, the confusion matrix shows a relatively low number of misclassifications, particularly for false negatives, which aligns with the high recall score. The model is effective at identifying depressed cases.

Conclusions

In this study, we investigated whether middle age can be considered a potential risk factor in PPD using a combination of statistical analysis, machine learning techniques, and XAI. The findings indicate that age shows a measurable association with several psychological and behavioral indicators linked to PPD. In particular, middle-aged mothers show distinct patterns in emotional responses, stress levels, and related factors, suggesting that age may play a contributory role in the manifestation of PPD symptoms. The integration of machine learning models enabled the identification of hidden patterns within the dataset, while the application of XAI techniques provided transparency in model predictions. The combined approach demonstrates the effectiveness of data-driven methods in finding complex relationships in maternal mental health studies. Overall, this research highlights the importance of considering age as a meaningful factor in PPD assessment and demonstrates the value of integrating statistical methods with explainable machine learning for improved mental health analysis and decision support.

However, the study has certain limitations. The dataset used does not explicitly specify the geographical or demographic origin of participants, which may restrict the generalizability of the findings. Future work will focus on incorporating clinically validated datasets, expanding demographic diversity, and applying advanced approaches to improve predictive accuracy.

Additional Information

Author Contributions

All authors have reviewed the final version to be published and agreed to be accountable for all aspects of the work.

Concept and design: Moumita Chatterjee, Trisha Roy Chowdhury

Drafting of the manuscript: Moumita Chatterjee, Trisha Roy Chowdhury

Supervision: Moumita Chatterjee

Acquisition, analysis, or interpretation of data: Trisha Roy Chowdhury, Dhrubasish Sarkar

Critical review of the manuscript for important intellectual content: Dhrubasish Sarkar

Disclosures

Human subjects: All authors have confirmed that this study did not involve human participants or tissue. **Animal subjects:** All authors have confirmed that this study did not involve animal subjects or tissue. **Conflicts of interest:** In compliance with the ICMJE uniform disclosure form, all authors declare the following: **Payment/services info:** All authors have declared that no financial support was received from any organization for the submitted work. **Financial relationships:** All authors have declared that they have no financial relationships at present or within the previous three years with any organizations that might have an interest in the submitted work. **Other relationships:** All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

Data Availability Statements

The datasets (and/or code) supporting this study are available from the corresponding author upon reasonable request.

References

1. Miller LJ: Postpartum depression. JAMA. 2002, 287:762-765. 10.1001/jama.287.6.762
2. Wisner KL, Chambers C, Sit DK: Postpartum depression: a major public health problem. JAMA. 2006, 296:2616-2618. 10.1001/jama.296.21.2616
3. United Nations: The 2030 Agenda for Sustainable Development's 17 Sustainable Development Goals (SDGs). United Nations, New York, NY; 2015.
4. Howard LM, Khalifeh H: Perinatal mental health: a review of progress and challenges. World Psychiatry. 2020, 19:313-327. 10.1002/wps.20769
5. Wells T: Postpartum depression: screening and collaborative management. Primary Care: Clinics in Office Practice. 2023, 50:127-142. 10.1016/j.pop.2022.10.011
6. Gupta A, Pajai S, Gupta A, Singh Thakur A, Muneeba S, Batra N, Patel DJ: In the shadows of motherhood: a comprehensive review of postpartum depression screening and intervention practices. Cureus. 2024, 16:e54245. 10.7759/cureus.54245

7. Gopalakrishnan A, Venkataraman R, Gururajan R, Zhou X, Zhu G: Predicting women with postpartum depression symptoms using machine learning techniques. Mathematics. 2022, 10:4570. 10.3390/math10234570

8. Qi W, Wang Y, Wang Y, et al.: Prediction of postpartum depression in women: development and validation of multiple machine learning models. Journal of Translational Medicine. 2025, 23:291. 10.1186/s12967-025-06289-6

9. Sibbald L, van den Heuvel MI, Haas MR, et al.: Identifying prenatal risk factors of postpartum depression with machine learning. Scientific Reports. 2025, 15:34610. 10.1038/s41598-025-18204-6

10. Amit G, Girshovitz I, Marcus K, Zhang Y, Pathak J, Bar V, Akiva P: Estimation of postpartum depression risk from electronic health records using machine learning. BMC Pregnancy and Childbirth. 2021, 21:630. 10.1186/s12884-021-04087-8

11. Liu H, Dai A, Zhou Z, et al.: An optimization for postpartum depression risk assessment and preventive intervention strategy based machine learning approaches. Journal of Affective Disorders. 2023, 328:163-174. 10.1016/j.jad.2023.02.028

12. Zhang Y, Wang S, Hermann A, Joly R, Pathak J: Development and validation of a machine learning algorithm for predicting the risk of postpartum depression among pregnant women. Journal of Affective Disorders. 2021, 279:1-8. 10.1016/j.jad.2020.09.113

13. Zhang R, Liu Y, Zhang Z, Luo R, Lv B: Interpretable machine learning model for predicting postpartum depression: retrospective study. JMIR Medical Informatics. 2025, 13:e58649. 10.2196/58649

14. Andersson S, Bathula DR, Iliadis SI, Walter M, Skalkidou A: Predicting women with depressive symptoms postpartum with machine learning methods. Scientific Reports. 2021, 11:7877. 10.1038/s41598-021-86368-y

15. Hochman E, Feldman B, Weizman A, et al.: Development and validation of a machine learning-based postpartum depression prediction model: A nationwide cohort study. Depression and Anxiety. 2021, 38:400-411. 10.1002/da.23123

16. Wang Y, Yan P, Wang G, et al.: Trajectory on postpartum depression of Chinese women and the risk prediction models: A machine-learning based three-wave follow-up research. Journal of Affective Disorders. 2024, 365:185-192. 10.1016/j.jad.2024.08.074

17. Matsuo S, Ushida T, Emoto R, et al.: Machine learning prediction models for postpartum depression: A multicenter study in Japan. Journal of Obstetrics and Gynaecology Research. 2022, 48:1775-1785. 10.1111/jog.15266

18. Raisa JF, Kaiser MS, Mahmud M: A machine learning approach for early detection of postpartum depression in Bangladesh. Brain Informatics. Mahmud M, He J, Vassanelli S, van Zundert A, Zhong N (ed): Springer, Cham; 2022. 13406:241-252. 10.1007/978-3-031-15037-1_20

19. Kim H: Predictive analysis of postpartum depression using machine learning. Healthcare. 2025, 13:897. 10.3390/healthcare13080897

20. Zhang Y, Joly R, Beecy AN, et al.: Implementation of a machine learning risk prediction model for postpartum depression in the electronic health records. AMIA Summits on Translational Science Proceedings. 2024, 2024:1057-1066.

21. Prabhashwaree THKR, Wagarachchi NM: Towards machine learning approaches for predicting risk level of postpartum depression. 2022 6th SLAAI International Conference on Artificial Intelligence (SLAAI-ICAI), Colombo, Sri Lanka. 2022, 1-6. 10.1109/SLAAI-ICAI56923.2022.10002477

22. Shivaprasad S, Chadaga K, Sampathila N, Prabhu S, Rajagopala Chadaga P, Swathi KS: Explainable machine learning methods to predict postpartum depression risk. Systems Science & Control Engineering. 2024, 12:2427033. 10.1080/21642583.2024.2427033

23. Anand Y, Sharma V, Jain S, Tarwani S: Ensemble machine learning model for predicting postpartum depression disorder. 2024 IEEE Region 10 Symposium (TENSYMP), New Delhi, India. 2024, 1-6. 10.1109/TENSYMP61132.2024.10752305

24. Suganthi D, Geetha A: Predicting postpartum depression with aid of social media texts using optimized machine learning model. International Journal of Intelligent Engineering & Systems. 2024, 17:417-427. 10.22266/ijies2024.0630.33

25. Nasim S, Al-Shamayleh AS, Thalji N, et al.: Novel meta learning approach for detecting postpartum depression disorder using questionnaire data. IEEE Access. 2024, 12:101247-101259. 10.1109/ACCESS.2024.3427685

26. Fischbein R, Cook HL, Baughman K, Díaz SR: Using machine learning to predict help-seeking among 2016-2018 Pregnancy Risk Assessment Monitoring System participants with postpartum depression symptoms. Women's Health. 2022, 18:1-15. 10.1177/17455057221139664

27. Lin S, Wang P, Yue C, et al.: Development and validation of an interpretable machine learning model and online web-based calculator based on social-ecosystem theory for early prediction of postpartum depression: a longitudinal study. Frontiers in Public Health. 2025, 13:1685305. 10.3389/fpubh.2025.1685305

28. PostPartum Depression. (2023). https://doi.org/10.34740/KAGGLE/DS/2830731.

29. Stewart DE, Vigod S: Postpartum depression. New England Journal of Medicine. 2016, 375:2177-2186. 10.1056/NEJMcp1607649

30. Pearlstein T, Howard M, Salisbury A, Zlotnick C: Postpartum depression. American Journal of Obstetrics and Gynecology. 2009, 200:357-364. 10.1016/j.ajog.2008.11.033

31. O'Hara MW: Postpartum depression: what we know. Journal of Clinical Psychology. 2009, 65:1258-1269. 10.1002/jclp.20644

How to cite this article:
Roy Chowdhury T, Sarkar D, Chatterjee M (April 13, 2026) Assessing the Association Between Maternal Age and Postpartum Depression Using Statistical Analysis and Explainable Artificial Intelligence. Cureus J Comput Sci 3 : es44389-026-00057-3. DOI https://doi.org/10.7759/s44389-026-00057-3