GenoMed Connect
ISSN: 3105-3645
2025, Article ID. 2025.0017, Cite as: https://doi.org/10.69709/GenomC.2025.100122

SCIFINITI PUBLISHING

Research Article

OPEN ACCESS

AI-Driven Machine Learning Analysis in Major Depressive Disorder (MDD): Sex-Based Variations in Oxytocin and Clinical Profiles

Hayder Abdul-Amir Makki Al-Hindy

Department of Pharmacology and Toxicology, College of Pharmacy, University of Babylon, Babylon 51001, Iraq

Article History

Submitted: August 07, 2025 Accepted: December 09, 2025 Published: December 18, 2025

Abstract

Background: Major depressive disorder (MDD) exhibits significant sex-specific differences in psychiatric manifestations, biological characteristics, and clinical responses to treatment. The current breakthroughs in machine learning and artificial intelligence (AI) offer novel ways to analyze multidimensional clinical and biomarker datasets in psychiatric research. However, sex-related differences in oxytocin, inflammatory markers, and clinical profiles in MDD remain understudied, especially among people in the Middle East. **Objectives:** This research aimed at examining sex-based variations in oxytocin, clinical, and inflammatory measures in adults with MDD. In addition to assessing the predictive value of these variables for depression severity, we tested them using machine learning. **Methods:** A cross-sectional study among 198 adults diagnosed with MDD was conducted at Merjan Medical City, Babylon, Iraq (2022–2023). Sociodemographic and medical data, in addition to laboratory data—including plasma oxytocin (measured using ELISA; Enzo Life Sciences, CA, USA), hemoglobin (Hb), leukocyte counts (WBCs), and body mass index (BMI)—were collected. The severity of MDD presentation was evaluated through the "Patient Health Questionnaire-9 (i.e., PHQ-9)". The statistical analyses encompassed *t*-tests, MANOVA, and logistic regression. AI-machine learning analyses comprised principal component analysis (PCA), Random Forest classification, and K-means clustering. **Results:** Females had higher depression severity ($p = 0.010$) and higher BMI ($p = 0.0002$), whereas males had higher hemoglobin levels ($p < 0.001$). There was no difference in Oxytocin or WBC. MANOVA established that there are significant multivariate effects of sex on the severity of depressive symptoms, oxytocin, and WBCs ($p = 0.001$). Logistic regression indicated that no single biomarker was significantly associated with severe depression. The Random Forest model correctly classified non-severe cases but not severe ones; moreover, the importance of individual features was negligible. K-means clustering showed moderate sex-based separation with partial overlap. **Conclusion:** Clinical and biological differences between sexes are also observable in adult patients with MDD in Iraq, though single biomarkers, such as oxytocin, are only partially predictive of the severity of depression. Despite data limitations, AI-based analyses helped highlight the complexity of depression and the need for multifaceted, sex-sensitive risk stratification and personalized care.

Keywords:
sociodemographic; major depressive disorder (MDD); oxytocin; sex-based variations; leukocyte counts

1. Introduction

Major depressive disorder (MDD) is a prevalent psychiatric illness and a leading cause of disability, affecting over 300 million people worldwide and substantially impairing social and professional functioning [1,2]. The World Health Organization (WHO) recognizes MDD as a significant cause of the worldwide burden of depression, with increasing rates of frequency being experienced in both developed and developing nations. Some of the notable situations that the WHO Mental Health Gap Action Program (mhGAP) deals with include depression and self-harm/suicide [3]. MDD is defined by low mood persistence, cognitive impairment, anhedonia, and a set of other

Corresponding Author:
Hayder Abdul-Amir Makki Al-Hindy, Department of Pharmacology and Toxicology, College of Pharmacy, University of Babylon, Babylon 51001, Iraq, phar.hayder.abdul@uobabylon.edu.iq

symptoms, which frequently lead to a significant level of distress and a decrease in quality of life [4].

One of the most salient aspects of MDD is that it has sex-based differences [5]. The disability adjusted life-years (DALYs) health burden was far higher in women than in men [6]. Such variations do not only end at prevalence, but also at symptom pattern, course, and response to treatment. Women who have MDD tend to complain more of somatic complaints, fatigue, and insomnia, whereas men tend to show more impulsive substance abuse and atypical affective performances [7]. Several neuroimaging investigations have shown sex-related changes in the brain tissues and function among MDD patients, with alterations in the brain cortical thickness and/or gray matter size, which could explain such phenotypic differences [8].

An increasing number of studies are examining the biological processes that underlie these sex differences.[9]. Inflammation, which is specified by an elevated C-reactive protein (CRP) [10] and changed white blood cells (WBCs) counts are becoming more and more an acknowledged part of depressive pathophysiology [5]. Depression risk and severity have also been linked to hemoglobin (Hb) and body mass index (BMI), whereas women tend to have higher inflammatory levels and BMI, and men have higher levels of hemoglobin [5,7]. These results indicate that there can be sex-specific physiological reactions that can affect the probability and the expression of depression.

Oxytocin is another neuroendocrine factor that has been identified. Oxytocin is a neuropeptide hormone that is formed in the hypothalamus, which is associated with stress regulation, social bonding, and emotional expressions [11]. Current research indicates that oxytocin could regulate depressive symptomatology and could be involved in the interaction with inflammatory pathways, which might have sex-specific effects on moods and behavior. It is still uncertain, though, as some reports show that in depressed patients, there is a decrease in oxytocin, and some studies do not see any significant difference [11]. Little is known about the relationship between oxytocin and inflammation and clinical aspects of MDD, in particular among populations in Middle Eastern countries.

The demographic and socioeconomic factors also influence the clinical course of MDD. Depression risk, severity of the symptoms, as well as treatment outcomes, are affected by education level, marital status, and living in a city versus rural areas in the country [12]. As an illustration, educational levels can protect against depression in certain situations, whereas marital status and social support have consistently been associated with favorable outcomes of mental health [11,13]. Such factors are especially relevant in Iraq and other low- and middle-income nations because of a current sociopolitical crisis, limited mental health providers, and cultural stigma against psychiatric disease.

In this study, the author aimed to investigate sex-based variations in clinical and biological indicators related to depression, including oxytocin levels, body mass index (BMI), hemoglobin, leukocyte counts (WBCs), severity, and duration of depression. To identify patterns significant with respect to sex, we employed combined statistical analyses, and machine learning methods were used to assess the predictive power of these markers for severe forms of depression. The study also examined the data structure underlying using K-means clustering to reveal possible subgroups. Finally, it aimed to offer a multivariate and holistic insight into sex-differentiated interactions between biological and clinical factors in MDD.

2. Materials and Methods

2.1. Study Design and Plan

This observational cross-sectional study was conducted at the outpatient psychiatric clinics of Merjan Medical City, Babylon, Iraq, in collaboration with the College of Pharmacy, University of Babylon, and the Babil Health Directorate. The research was done between August 2022 and August 2023.

2.2. Patient Selection

A total of 198 participants diagnosed with MDD were registered. The study inclusion criteria were: Aged $\geq$ 18 years, diagnosis of MDD based on DSM-5 criteria (version 5), verified by the "Mini International Neuropsychiatric Interview." [9–11], currently on antidepressants for the last 3 months, as well as regular attendance at follow-up schedules throughout the research period.

Exclusion criteria comprised a history of degenerative or traumatic illnesses of the brain, convulsions, use of steroids within the past three months, drug addiction, or failure to complete the questionnaire formula independently or with family help.

2.3. Severity Assessment of Depression

Psychiatric evaluation supported the clinical diagnosis, with additional information provided by family members as needed. The clinical diagnosis of MDD was established according to the "Diagnostic and Statistical Manual of Mental Disorders, 5th Edition (DSM-5)" [14], confirmed by the "Mini International Neuropsychiatric Interview (MINI)" [9–11]. The "Patient Health Questionnaire-9 (PHQ-9)", which is an authenticated questionnaire, was used to determine the severity of depression by consist-

2025, Vol. 2, Article ID. 2025.0017
https://doi.org/10.69709/GenomC.2025.100122

ing of nine items that measured the presence of depression symptoms in the past week [9–11]. Items are scored on a 4-point scale from 0 (not at all) to 3 (nearly every day), separately. Scores were added to yield an overall score range of 0 to 27. Severity of MDD was categorized as follows: (0–4) No depression, (5–9) Mild depression, (10–14) Moderate depression, (15–20) Moderately severe depression, and (21–27) Severe depression.

2.4. Chemical and Physiological Measurements

Blood samples from all participants were collected at standardized times. Laboratory investigations included: Hb and WBCs, which were assessed by a fully automated hematology analyzer (Sysmex XN-550, Kobe, Japan). Oxytocin levels were measured using a commercially available enzyme-linked immunosorbent assay kit (Enzo Life Sciences, NY, USA). The latter ELISA has been rigorously validated for measuring oxytocin in human blood. Height and weight were recorded to calculate body mass index (BMI, kg/m²).

2.5. Sociodemographic Data

The required information on age, sex, marital status, education, and rural/urban residence was recorded using structured questionnaire interviews and a review of medical archives.

2.6. Statistical Analyses

Data were analyzed with SPSS V-28 (IBM, Cal, USA) [15] and JASP software version 0.18.3.0 [16]. The descriptive metrics (means, standard deviation [SD], and frequencies) presented counts for demographic or clinical variables. Group comparisons were performed using independent-samples t-tests for continuous data and chi-square tests for categorical data. The effect size was measured by using Cohen's d.

MANOVA was used to evaluate the multivariate effect of sex on oxytocin levels, depression severity scores, and WBC count. Logistic regression analysis was used to estimate the predictive power of biomarkers for severe depression, with odds ratios (ORs) and 95% confidence intervals (CIs) reported. Receiver operating characteristic (ROC) curve analysis evaluates sensitivity and specificity.

Priori sample size calculation was completed using GPower computer software (version 132 3.1.9.7). For an independent-samples t-test (two-tailed), with an effect size of d = 0.4 (moderate), an alpha error probability of 0.05, and a desired power of 0.80, the total sample size needed was around 200 applicants. Our final sample of 198 participants is therefore considered adequate to detect moderate effects for the primary group comparisons.

2.7. AI and Machine Learning Tools

The Scikit-learn Python library (version 1.2.2) was used to conduct machine learning analyses, including Principal Component Analysis (PCA), Random Forest classification, and K-means clustering. The specified AI-based tools were applied to the tasks of data dimensionality reduction (PCA), categorization of depression severity (Random Forest), and unsupervised segmentation of patient groups (K-means). These tools were used to analyze and model the exploratory data analysis in this work. Machine learning analysis involved Random Forest classification [17], K-means clustering [18], in addition to the principal component analyses (PCA) applied for data visualization. In the Random Forest classification, the data were split into training and test sets. It is important to note that the problem of class imbalance in the outcome variable (severe vs. non-severe depression) was not considered through the use of techniques such as stratified sampling or class weighting in the initial analysis. Additionally, the model's performance was evaluated on a single test split rather than cross-validation. Although these methodological decisions are typical in exploratory analyses, they reduce the strength and applicability of the machine learning findings and may also have contributed to the model's low predictive accuracy in severe cases.

The GPower 3.1 post-hoc power analysis showed that the 198-sample size had a power greater than 80 to identify medium-sized effect sizes (f 2 = 0.15) in multivariate analyses at an alpha = 0.05.

2.8. Patient Data Security and Ethical Issues

All participant-reported data were anonymized and deposited securely, consistent with official guiding principles. Written informed consent was obtained from all applicants. The protocol was approved by the Babil Health Directorate (reference ID: 374-2; 23 June 2022) and by the Ethics Committee of the College of Pharmacy, University of Babylon. The current study protocol adhered to the principles of the Declaration of Helsinki.

3. Results

The results of Table 1 show significant differences across the sexes in multiple variables, as indicated by descriptive statistics. The mean age of females was somewhat higher (37.6 ± 14.6 years) than that of males (35.6 ± 12.9 years);

however, the difference was not statistically significant (p = 0.166).

Oxytocin concentrations were similar across the sexes (males: 25.6 ± 15.2, females: 24.2 ± 16.6; p = 0.407), as were WBCs (males: 7.9 ± 1.3, females: 8.08 ± 1.4; p = 0.326). Nevertheless, men had much higher Hb measures (14.1 ± 1.09) compared to the women (13.0 ± 0.9; p = 0.001), and the effect size (Cohen's d = −1.107) was big. The BMI distributions were very different and more varied among males (20.7–177.3) than among females (20.8–54.7), and non-parametric results demonstrated significant differences (p = 0.0002).

Sex-related variation (Table 2) was also observed with the measures of depression (Table 1). Women reported greater depression severity (3.83 ± 0.43) than men (3.69 ± 0.56; p = 0.010), with an effect size of small to moderate magnitude (d = 0.305). On the other hand, males exhibited a greater average depression symptom (46.4 ± 72.2 months) than females (24.05 ± 60.1 months; p = 0.003), which is a modest effect size (d = −0.347). The use of MANOVA (Table 3) revealed multivariate significant sex effects on the aggregate dependent variables of severe depression, oxytocin concentration, and WBCs (Wilks Lambda = 0.677, F (49, 330) = 3.21, p = 0.001).

Table 1: Descriptive statistics and group comparisons by sex.

Variables	Females	Males	Mann-Whitney U p-Value
Age	37.62 (14.65)	35.58 (12.95)	0.255
	15.00–82.00	12.00–75.00	
Oxytocin Levels	24.20 (16.63)	25.61 (15.18)	0.3214
	0.42–64.19	1.06–57.38	
BMI	31.18 (4.85)	31.36 (18.73)	0.0002
	20.80–54.70	20.73–177.29	
WBCs	8.08 (1.35)	7.94 (1.34)	0.6787
	4.80–14.00	4.00–13.50	
Hemoglobin	13.04 (0.91)	14.12 (1.09)	0.0
	9.00–14.70	13.00–17.40	
Depression Severity	3.83 (0.43)	3.69 (0.56)	0.0036
	2.00–4.00	2.00–4.00	
Depression Duration	24.05 (60.12)	46.41 (72.21)	0.0003
	0.00–360.00	0.00–210.00	

Table 2: Effect sizes (Cohen's d) between the males and females.

Parameters	Cohen's d
Age	0.145
Body Mass Index	−0.015
Oxytocin Level	−0.088
Hemoglobin Level	−1.107
White Blood Cells	0.106
Depression Duration	−0.347
Depression Severity	0.305

Table 3: Effect of sex on depression severity, oxytocin, and WBCs using MANOVA analyses.

Wilks' Lambda	F-Value	Denominator DF	Nominator DF	Significance
0.677	3.207	330.0	49	$p < 0.001$

2025, Vol. 2, Article ID. 2025.0017
https://doi.org/10.69709/GenomC.2025.100122

The logistic regression results have provided no significant predictors of the severity of depression among the studied variables, including oxytocin levels, BMI, white blood cells, and hemoglobin (all $p > 0.05$), (Table 4). ORs were near 1.0, and their wide confidence intervals indicated low predictive ability. The values of sensitivity and specificity were between 50 and 70 percent, which further supports the moderate usefulness of these parameters as a risk factor for depression.

The Random Forest model, intended to predict depression severity, achieved perfect categorization for non-severe cases; however, it was unable to identify any severe cases, as demonstrated by the classification report (Table 5) and confusion matrix (Table 6). The feature importance study yielded null values for all predictors, including duration of depressive symptoms and oxytocin concentrations, which were anticipated to be significant based on prior literature. This proposes potential restrictions in the feature selection or the model's training data (Table 2).

As revealed in Table 7, the feature importance analysis yields null values for all predictors. The random forest model performed well in non-severe cases, but did not classify severe cases, likely due to class imbalance. To overcome this, stratified sampling or class weighting should be introduced in the future. Its use is suggested to be improved through cross-validation and other performance metrics (e.g., AUC, precision, recall), thereby enhancing the model's generalizability and clinical applicability.

The K-means clustering analysis is presented in a two-dimensional PCA space showing modest separation between males and females, with incomplete overlap in the dominant area of the curve plot (Figure 1). Principal Component 1 (horizontal axis) seemed to explain the majority of the data's variation (probably due to depression-related variables). In contrast, Principal Component 2 (vertical axis) had less variance, but was still substantial. It is also important to note that female subjects tended to be more concentrated in a few spots, especially on the upper-left-hand side, compared to male subjects, who were more widely spread across the plot. This trend follows our previous observation of sex-based disparities in the severity and duration of depression (Table 3). However, the overlap is quite significant; thus, the variables considered may not be sufficient to provide complete differentiation between sexes within clinical groups. The clustering findings complement the MANOVA findings (Table 3) and provide a more detailed illustration of the interactions among these variables in multivariate space. The presence of some outlier data points across the groups is something to be investigated in forthcoming studies involving larger sample sizes.

Table 4: Predictive performance of the logistic regression models that test relationships between the depression severity and biomarkers.

Variables	Odds Ratio	CI 95%	Significance	Specificity	Sensitivity
BMI	1.04	[0.99, 1.07]	0.12	64%	55%
Oxytocin	0.98	[0.95, 1.01]	0.25	58%	62%
White Blood Cell Counts	1.15	[0.97, 1.36]	0.09	52%	70%
Hemoglobin	1.08	[0.89, 1.31]	0.42	50%	65%

Table 5: Results of the classification of the random forest machine learning analyses predicting severity of depression based on clinical and biological features.

Classes	F1-Score	Recall	Precision
0	1.00	1.00	1.00
Macro avg	1.00	1.00	1.00
Weighted avg	1.00	1.00	1.00

Table 6: Confusion matrix of the random forest machine learning analyses predicting the severity of depression based on clinical and biological features.

	Predicted 1	Predicted 0
Actual 1	0	0
Actual 0	0	114

Table 7: Feature importance of the random forest machine learning analyses predicting the severity of depression based on clinical and biological features.

Feature	Importance
Sex	0.000
Age	0.000
Body Mass Index	0.000
Oxytocin Level	0.000
Hemoglobin Level	0.000
White Blood Cells	0.000
Depression Duration	0.000

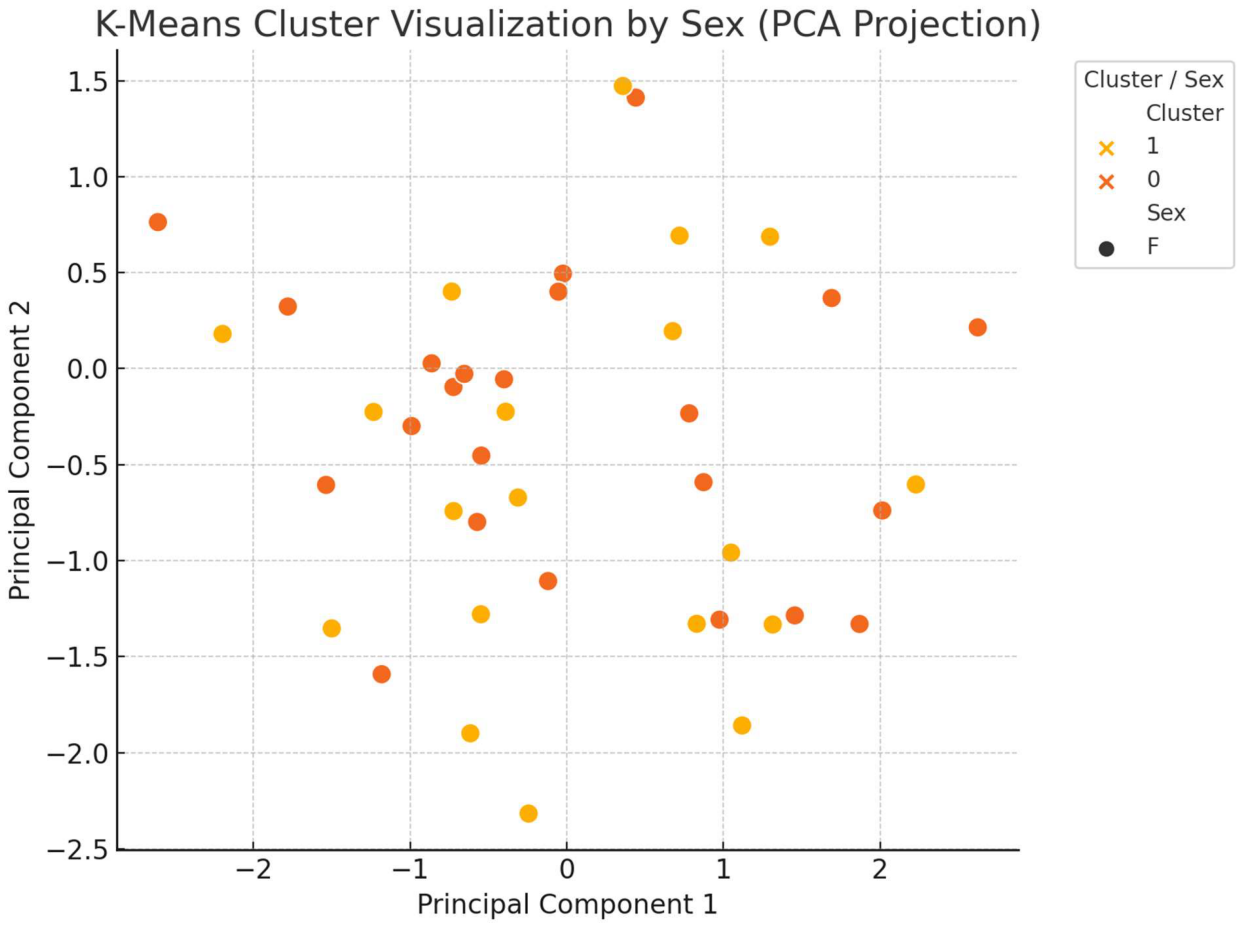

Figure 1: Two-dimensional projection of K-Means clustering analysis of stratified by sex participants with Principal Component Analysis (PCA). This is a unique character developed in this manuscript. The figure was created using the scikit-learn and matplotlib python libraries.

The MANOVA revealed significant multivariate effects of sex on the combined dependent variables—oxytocin levels, depression severity scores, and WBC counts (Wilks' Lambda = 0.677, F(49, 330) = 3.2, $p < 0.001$).

2025, Vol. 2, Article ID. 2025.0017
https://doi.org/10.69709/GenomC.2025.100122

4. Discussion

4.1. Sex Differences in Physiological and Clinical Profiles

The study results extend and support the increasing body of evidence that substantial sex-based variations in clinical appearance and primary biomarkers characterize major depression. Consistent with epidemiologic evidence, women in this cohort showed greater depression severity than men, despite the latter showing a longer median duration of depression. This supports preceding research signifying that females not only suffer depressive symptoms more often but also tend to experience much symptom load and/or functional loss [5–8,12,13].

4.1.1. Biological Biomarkers: Hemoglobin, WBC, BMI, and Oxytocin

Another important finding was a sex difference in hemoglobin, with males having higher values that were significantly greater than those of females ($p < 0.001$). This is consistent with known physiological parameters, because testosterone increases erythropoiesis [19], although it also indicates that depressive states can have sex-specific effects on hematological parameters.

The clinical implications are also interesting: low hemoglobin levels in depressed females may increase fatigue and cognitive symptoms, which in turn could increase the total burden of the disease. There were also sex-based cells, differences in white blood cell levels, and BMI. WBC counts and BMI were slightly higher in females, which may reflect increased inflammatory responses or greater somatic symptomatology in women with MDD. The results are in line with the literature that correlates inflammation due to a high level of CRP and the number of leukocytes with depressive pathophysiology, especially in women [11]. The BMI role can be complicated, as, on the one hand, high BMI is a risk factor of depression. Still, on the other hand, depressive symptoms may be caused by the low level of physical activity, emotional eating, etc. The connection between BMI and inflammation can also be the mediator of depressive symptoms [20], and its relation with systemic inflammation may further facilitate depressive manifestations.

The neuropeptide Oxytocin, a hormone involved in social bonding and stress regulation, was not significantly different between sexes in the cohort we studied. To some degree, it is surprising, considering that sex-specific oxytocinergic influences on mood and social cognition have been reported [5,21]. Nevertheless, subtle differences might be obscured by broad inter-individual differences and the possible effects of acute stressors, medication, or sampling timing. There is justification to conduct future studies that involve larger populations and allow a longitudinal design to bring out a clear understanding of the role of oxytocin in the sex-specific depressive features.

4.1.2. Demographic and Socioeconomic Influences

Socioeconomic factors (education, marital status, residence) were associated with clinical and biological profiles. In regression models, higher education level emerged as a significant predictor of lower depression severity scores ($\beta = -0.24$, $p < 0.05$), particularly among female participants. We find that females with a higher education level had greater BMI in comparison to their male counterparts, and here we see that sociocultural and biological influence is intertwined to form depressive symptomatology. Marital status also influenced physical and psychological health, likely due to differences in age and BMI profiles between married and non-married participants. These outcomes were consistent with the earlier studies that have shown the beneficial effect of social support and steady relations on mental health, especially in female patients [1,7,12,13].

Urban–rural residence was associated with BMI and may relate to differences in depression risk and course. The elevated risk of obesity and depression appears to be associated with urban populations, where higher psychosocial stress and lower levels of physical activity are prevalent, with this association especially strong in women. These observations highlight the importance of context-sensitive measures of depression and interventions that are responsive to the socioeconomic reality.

4.2. Predictive Modeling: Logistic Regression and Machine Learning

The failure of the Random Forest model to identify severe depression cases and the null values of features of importance underscore the challenge of the task of class imbalance and predicting complex psychiatry using a small number of biomarkers. Resampling, cross-validation, and an enriched feature set should be used in future studies to improve the model's performance.

Although group differences were observed, logistic regression analysis indicated that none of the investigated independent variables (oxytocin, BMI, WBC, hemoglobin) significantly predicted depression severity. Odds ratio values were near unity, and the predictive performances (sensitivity/specificity) were low. This underscores the multifactorial landscape of depression, where a single variable lacks satisfactory discriminatory power for clinical applications.

The results of this study in an Iraqi cohort provide valuable regional insights into the literature on MDD. However, it aligns with a recent pattern of global literature that indicates greater development of depression among women [5,9–11]. The sociocultural peculiarities of a particular region, such as conflict, displacement, and gender roles, might determine specific manifestations of these gender disparities. The direct comparison with other Middle Eastern studies is limited by the scarcity of comparative research on biomarkers in the region.

Modern AI research has attained much evolution in enhancing the diagnosis of MDD and the neurological characteristics of sex differences in depression. The 2025 studies developed explainable AI models with polysomnographic phenotypes and more advanced machine learning (e.g., random forests, XGBoost), which were highly accurate (~85) in predicting depression. However, demographic characteristics, such as sex, were also considered. While sex-based differences in AI model performance or feature importance were not among the dominant findings, they remain essential variables [22].

A 2023 systematic review summarized sex differences in the brain in terms of depression, with various neuroimaging differences between males and females being reported, with differences in the limbic and frontal circuits. These neurobiological distinctions underlie clinical sex differences and suggest that incorporating sex-specific brain characteristics could enhance the accuracy of future AI models [23]. Therefore, the researchers did not find evidence of sex-based variances in the inflammatory marker associations with MDD among these Middle Eastern refugees.

A different study on explainable AI (XGBoost, SHAP, LIME) applied to wearable actigraphy data to predict depression and measure its severity also identified demographic variables such as age, but not specifically sex, as essential predictors, and found circadian disruption to play an important role [24].

The Random Forest machine learning model was applied and further highlighted these challenges. Though the algorithm perfectly categorized non-severe MDD patients (Tables 5 and 6), it failed to categorize severe depressive form. Meanwhile, feature importance analysis generated null values for all prognosticators (Table 7). This raises the question of whether the features described here are not informative enough, or whether more advanced modeling, including any of the following, is needed to make the prediction robust: incorporating more clinical, genetic, or neuroimaging data. It also highlights the importance of large, diverse datasets for training credible models.

K-means clustering and principal component analysis have provided a fine-grained visual representation of the multivariate association among variables. Women respondents were more concentrated, especially in the upper-left area of the PCA plot, as compared to males. Such a trend is indicative of more homogeneity of depression display in females and the effect of sex on the interaction of clinical and physiological variables. Nevertheless, the substantial overlap between sexes suggests that these markers cannot fully account for the individual differences observed in MDD.

The potential benefits and challenges of AI and machine learning in psychiatry are illustrated by the use of these high-tech tools in the current study, although the investigation remains exploratory. ML algorithms like Random Forest have the capability to capture high-order and non-linear interactions between variables compared to traditional statistical models (e.g., regression), which assume a linear relationship between variables [25]. This is particularly suitable in a heterogeneous disorder like MDD, whereby etiology is multifactorial. Even though limited sample size and the method used to assess class imbalance, our models provided a valuable hint: that simple linear combinations of a few biomarkers are inadequate for prediction. It underlines the necessity to use more sophisticated, data-driven approaches so that they can detect the hidden subgroups and effects of interactions that otherwise would remain overlooked in traditional methods. In the future, it will be essential to carry out bigger and more complicated studies that will assist in utilizing the maximum potential of AI to rank the risks individually when MDD.

4.3. Clinical and Research Implications

The findings of this study carry significant implications. To begin with, they confirm the necessity of sex-specific assessment and treatment approaches to MDD. The clinicians need to be sensitive to the high symptom burden and inflammatory patterns in women, under-recognition of depressive presentations in men, who may have unusual manifestations, including substance use and impulsivity. Second, the lack of predictive power of specific markers underscores the importance of multidimensional evaluation, including clinical, laboratory, and psychosocial domains.

The study's findings demonstrate the importance of considering sex as a key biological factor at every level of the study design, investigation, and reporting. Forthcoming research must consider longitudinal changes in physiological biomarkers, the impact of hormonal changes (e.g., menstrual cycle, menopause), and the possibility of specific interventions based on metabolic and inflammatory levels.

2025, Vol. 2, Article ID. 2025.0017
https://doi.org/10.69709/GenomC.2025.100122

Lastly, the overlap of psychological, biological, and social elements of health in MDD necessitates a transtheoretical approach and cultural sensitivity in care, especially in the Middle East and the low-resource regions, where stigma and barriers to access can hinder care.

5. Limitations

Some drawbacks should be mentioned. The cross-sectional type does not allow causal inference, and the sample size is relatively small, which might be insufficient to generalize. The paper relies on self-report measures of the severity of depression, and this is likely to be affected by reporting bias. The study was sufficiently powered to carry out the primary comparisons. Yet, some of the sex-based differences in the study (e.g., the severity of depression) are of a relatively small scale, which is an indication that larger sample sizes may be needed to expand upon the more subtle effects, particularly when it comes to machine learning, and that there was an imbalance in the representation of the classes in the depression severity outcome.

A more detailed inflammatory profile should be included in future research to help elucidate sex-specific discrepancies in MDD. Also, the sample of patients with comorbid neurological or substance use diseases was excluded, but this is required due to the homogeneity of the sample, which could limit the generalizability of the results to more complex clinical groups.

Further studies to continue this research need to investigate the longitudinal changes in oxytocin, inflammatory markers, and clinical symptoms in major depressive disorder. A combination of genetic, neuroimaging, and hormonal evidence can contribute to a deeper understanding of how the observed sex differences are implemented and a better comprehension of how social and demographic factors influence the biological components. Larger and more diverse cohorts will be necessary to confirm these findings and to develop robust predictive models. Finally, such initiatives may lead to more accurate and personalized research on the diagnosis and treatment of depression, especially when it involves seeking sex-specific biological and psychosocial factors.

6. Conclusions

The study reveals sex-specific differences in clinical features, inflammatory markers, and oxytocin levels among patients with major depressive disorder. Though oxytocin levels and WBC counts were comparable between the sexes, males had higher hemoglobin concentrations, whereas females exhibited more severe depressive symptoms. BMI also differed: males showed a wider distribution, while females, especially those with higher education, tended toward elevated values. Interestingly, none of the biomarkers—oxytocin, BMI, WBC, or hemoglobin—predicted severe depression, as confirmed by logistic regression and machine learning analyses, highlighting the multifactorial nature of MDD. The multivariate model demonstrated a significant combined effect of sex, oxytocin, and WBC on the severity of depression. Furthermore, cluster analysis identified overlapping yet distinct profiles across sexes. These outcomes underscore the need to consider sex-specific biological and socioeconomic factors in the assessment and management of MDD. The use of AI-driven approaches, such as Random Forests and clustering, provides a robust framework for disentangling complex interactions in heterogeneous disorders like MDD, facilitating personalized interventions that may enhance therapeutic success.

List of Abbreviations

AI	Artificial Intelligence
AUC	Area Under the Curve
BMI	Body Mass Index
CI	Confidence Interval
CRP	C-Reactive Protein
DALYs	Disability-Adjusted Life Years
DSM-5	Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition
ELISA	Enzyme-Linked Immunosorbent Assay
Hb	Hemoglobin
MANOVA	Multivariate Analysis of Variance
MDD	Major Depressive Disorder
MINI	Mini International Neuropsychiatric Interview
OR	Odds Ratio
PCA	Principal Component Analysis
PHQ-9	Patient Health Questionnaire-9
ROC	Receiver Operating Characteristic
SD	Standard Deviation
WBC	White Blood Cell
WHO	World Health Organization

Author Contributions

The author confirms that he is solely responsible for Conceptualization, Methodology, Validation, Formal Analysis, Investigation, Resources, Data Curation, Writing---Original Draft Preparation, Writing---Review & Editing, Visualization, and Project Administration. The author has read and agreed to the published version of the manuscript.

Availability of Data and Materials

The data used to evaluate the study can be provided on request.

Ethics Committee Approval and Consent to Participate

The Ethics Committee of the College of Pharmacy, University of Babylon (reference number 374-2, June 23, 2022) approved the study protocol, and the local health authority approved it. Informed consent was written and issued to all the participants.

Human Rights

The current study protocol adhered to the principles of the Helsinki Declaration.

Consent for Publication

No consent for publication is required, as the manuscript does not involve any individual personal data, images, videos, or other materials that would necessitate consent.

Conflicts of Interest

The author declares no conflicts of interest.

Funding

The study did not receive any external funding and was conducted using only institutional resources.

Acknowledgments

The author acknowledges the use of the scikit-learn Python library for the machine learning analyses conducted in this study. The authors confirm that no AI tools were used to generate any content of this manuscript.

References

[1] Cui, L.; Li, S.; Wang, S.; Wu, X.; Liu, Y.; Yu, W.; Li, B. Major Depressive Disorder: Hypothesis, Mechanism, Prevention and Treatment. *Signal Transduct. Target. Ther.* **2024**, *9*, 30. [CrossRef]

[2] Proudman, D.; Greenberg, P.; Nellesen, D. The Growing Burden of Major Depressive Disorders (MDD): Implications for Researchers and Policy Makers. *Pharmacoeconomics* **2021**, *39*, 619–625. [CrossRef]

[3] World Health Organization. Depressive Disorder (Depression). Key Facts. 2024. Available online: https://www.who.int/news-room/fact-sheets/detail/depression (accessed on 11 December 2025).

[4] Zhao, H.; Li, L.; Zhang, X.; Shi, J.; Lai, W.; Wang, W.; Lu, C. Global, Regional, and National Burden of Depressive Disorders among Young People Aged 10–24 Years, 2010–2019. *J. Psychiatr. Res.* **2024**, *170*, 47–57. [CrossRef]

[5] Abd, B.A.; Al-Hindy, H.A.-A.M.; Obaid, S.R. Sex-Based Variations in Hemato-Physiological Characteristics of Patients with Major Depressive Disorder: Clinical and Socioeconomic Insights. *Med. J. Babylon* **2024**, *21*, 804–811. [CrossRef]

[6] Li, S.; Zhang, X.; Cai, Y.; Zheng, L.; Pang, H.; Lou, L. Sex Difference in Incidence of Major Depressive Disorder: An Analysis from the Global Burden of Disease Study 2019. *Ann. Gen. Psychiatry* **2023**, *22*, 53. [CrossRef] [PubMed] [PubMed Central]

[7] Hyde, J.S.; Mezulis, A.H. Gender Differences in Depression: Biological, Affective, Cognitive, and Sociocultural Factors. *Harv. Rev. Psychiatry* **2020**, *28*, 4–13. [CrossRef] [PubMed]

[8] Mou, J.; Zheng, T.; Long, Z.; Mei, L.; Wang, Y.; Yuan, Y.; Qiu, L. Sex Differences of Brain Cortical Structure in Major Depressive Disorder. *Psychoradiology* **2023**, *3*, kkad014. [CrossRef]

[9] Li, S.; Zhang, X.; Cai, Y.; Zheng, L.; Pang, H.; Lou, L. Sex Difference in Incidence of Major Depressive Disorder: An Analysis from the Global Burden of Disease Study 2019. *Ann. Gen. Psychiatry* **2023**, *22*, 53. [CrossRef]

[10] Al-Agam, A.N.M.; Obeiad, A.W.; Alzughaibi, M.A.K.; Al-Hindy, Hayder A.-A.M.; Alhaideri, A.F. The Association of Depressive Symptoms with Plasma C-Reactive Protein in Patients with Major Depressive Disorder under Treatment. *Iran. Rehabil. J.* **2021**, *19*, 425–432. [CrossRef]

[11] Alhaideri, A.F.; Al-Agam, A.N.M.; Al-Hindy, H.A.-A.M.; Mousa, M.J.; Kadhum, H.; Hatem, S. Inflammatory Associations of Peripheral Oxytocin, C-Reactive Protein Levels with Depression among Adult Age Group with Major Depressive Disorder. *Clin. Schizophr. Relat. Psychoses* **2021**, *15*, S9. [CrossRef]

[12] Hoveling, L.A.; Liefbroer, A.C.; Schweren, L.J.S.; Bültmann, U.; Smidt, N. Socioeconomic Differences in Major Depressive Disorder Onset among Adults Are Partially Explained by Lifestyle Factors: A Longitudinal Analysis of the Lifelines Cohort Study. *J. Affect. Disord.* **2022**, *314*, 309–317. [CrossRef]

[13] Wei, M.; Qin, Y.; Niu, X.; Niu, S.; Mu, F.; Yang, L.; Liu, Y. Marriage and Postpartum Major Depressive Disorder: A Systematic Review and Meta-Analysis of Cohort Studies. *J. Psychiatr. Res.* **2025**, *182*, 83–91. [CrossRef]

[14] Vahia, V.N. Diagnostic and statistical manual of mental disorders 5: A quick glance. *Indian J. Psychiatry* **2013**, *55*(3), 220–223. [CrossRef] [PubMed] [PubMed Central]

[15] IBM Corp. *IBM SPSS Statistics for Windows,* Version 28.0; IBM Corp: Armonk, NY, USA, 2021. Available online: https://www.ibm.com/support/pages/ibm-spss-statistics.

[16] Love, J.; Selker, R.; Marsman, M.; Jamil, T.; Dropmann, D.; Verhagen, J.; Ly, A.; Gronau, Q.F.; Šmíra,

M.; Epskamp, S.; et al. JASP: Graphical Statistical Software for Common Statistical Designs. *J. Stat. Softw.* **2019**, *88*, 2. [CrossRef]

[17] Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]

[18] Oti, E.; Olusola, M.; Eze, F.; Enogwe, S. Comprehensive Review of K-Means Clustering Algorithms. *Int. J. Adv. Sci. Res. Eng.* **2021**, *07*, 64–69. [CrossRef]

[19] Greenwall, B.; Reeder, K.; Anani, W. Elevated Plasma Testosterone Concentrations from Males on Testosterone Replacement Therapy Are Mitigated with Pathogen Reduction Technology. *Transfusion* **2025**, *65*, 524–530. [CrossRef] [PubMed] [PubMed Central]

[20] Cui, H.; Xiong, Y.; Wang, C.; Ye, J.; Zhao, W. The Relationship between BMI and Depression: A Cross-Sectional Study. *Front. Psychiatry* **2024**, *15*, 1410782. [CrossRef] [PubMed] [PubMed Central]

[21] Rosenfeld-Ganzel, A.; Shalev, H.; Hochman, S.; Zultan, R.; Cohen, N.; Naparstek, S. Oxytocin's Role in the Interaction between Emotion and Cognitive Control. *Biol. Psychol.* **2025**, *196*, 109004. [CrossRef] [PubMed]

[22] Enkhbayar, D.; Ko, J.; Oh, S.; Ferdushi, R.; Kim, J.; Key, J.; Urtnasan, E. Explainable Artificial Intelligence Models for Predicting Depression Based on Polysomnographic Phenotypes. *Bioengineering* **2025**, *12*, 186. [CrossRef]

[23] Mohammadi, S.; Seyedmirzaei, H.; Salehi, M.A.; Jahanshahi, A.; Zakavi, S.S.; Dehghani Firouzabadi, F.; Yousem, D.M. Brain-Based Sex Differences in Depression: A Systematic Review of Neuroimaging Studies. *Brain Imaging Behav.* **2023**, *17*, 541–569. [CrossRef] [PubMed] [PubMed Central]

[24] Ahmed, I.; Brahmacharimayum, A.; Ali, R.H.; Khan, T.A.; Ahmad, M.O. Explainable AI for Depression Detection and Severity Classification from Activity Data: Development and Evaluation Study of an Interpretable Framework. *JMIR Ment. Health* **2025**, *12*, e72038. [CrossRef] [PubMed] [PubMed Central]

[25] Liu, Z.; Lu, Z.; Zhu, W.; Yuan, J.; Cao, Z.; Cao, T.; Zhang, X. Comparison of Machine Learning Methods for Predicting Ground-Level Ozone Pollution in Beijing. *Front. Environ. Sci.* **2025**, *13*, 1561794. [CrossRef]