Predicting Depression using a Biochemistry Profile and Machine Learning for Better Risk Stratification

Tauseef Ibne Mamun
Michigan Technological University
Houghton, Michigan
USA

Lamia Alam
Michigan Technological University
Houghton, Michigan
USA

ABSTRACT

Depression is one of the mental health issues that are responsible for various morbidities if it remains undiagnosed. Previous works that used machine learning to predict depression used mainly qualitative information like socio-economic or socio-demographic data for creating predictive models. But there is a chance of getting biased qualitative data from patients that will hinder any correct prediction. Our goal was to examine if the major form of depression can be predicted with a high accuracy using a patient's biochemistry profile using Machine Learning. We also examined if minor depression can be predicted for a large portion of patients. The result suggests the method we proposed can be used as the first point of screening for depression especially major depression.

General Terms

Machine Learning, Biochemistry, Quantitative Data

Keywords

Depression, Early Detection, Risk Stratification

1. INTRODUCTION

Depression is a type of mood disorder that negatively affects how we think, feel, and act. According to a study conducted in 2017, 3.4% of people worldwide suffer from depression that ranges from 2-6% across different countries [1]. The economic burden of major depression is estimated to be $210.5 billion per year between 2005 and 2010, and 62% of the cost is attributed to different comorbidities [2] in the US. Sadly, less than one-third of adults with depression seek professional treatment [3]. The prevalence of major depression is 2-3 times higher in primary care patients than in the overall population because patients with depression use healthcare resources more frequently [4], [5]. Early detection of depression can help in reducing the mental and financial burden of the disease [6].

The diagnosis of depression is a very complicated and challenging task. Therefore, one needs to be careful to diagnose them with high accuracy. Machine learning (ML) and data mining techniques can play a significant role in early detection. Qualitative data regarding a patient can be a handy tool for early detection but the phrase 'mental help' is still taboo in most of the world [7], [8]. Research shows that depressed individuals who are concerned about stigma may adapt their social behavior to avoid exposure to discrimination [7]. So, a depressed patient may be reluctant to ask for help or, provide wrong or partially correct data due to the patient's social-economic/social-cultural status that will make it difficult for any ML algorithm to predict depression.

We focus on predicting depression based on quantitative information to identify patients who may be suffering from depression. The goal is to get them medical help so that condition of patients does not worsen. If used with other procedures for depression screening, the procedure in this paper can be used during the first point of contact for depression screening by any health care providers or AI applications.

2. RELATED WORKS

2.1 Qualitative data in prediction

Past research found association between depression and socio-economic or socio-demographic factors [9]–[12]. Machine learning algorithms are beneficial as screening tool for depression [13]. Age, sex, literacy, residence, marital status, recent bereavement, employment status, socio-economic status, personal income, history of anxiety, history of depression, family history of depression, any kind of addiction, pain at multiple sites, diabetes, HTN, hearing problem, visual impairment, mobility impairment, and insomnia were used to predict depression and anxiety in a study [14]. Here highest prediction accuracy of 89% was obtained with the Random Forest (RF) classifier. RF was also accurate when used in Depression, Anxiety and Stress Scale questionnaire (DASS 21) [15]. Early detection for depression also can be done by analyzing content in social networks. Decision Tree (DT) was effective in detecting depression for Facebook data collected from an online public source [16]. Early detection of depression using machine learning models was also done using posts and comments from Reddit [17]. So, qualitative information is a strong tool in the early detection of depression.

But self-report information can be influenced by many factors like question-wording, format, and context [18]. If multiple contexts influence a single domain, appropriately measuring something will become difficult after taking heed of every context. A domain can be influenced by multiple contexts like individual, organizational, cultural, and environment. Respondent bias [19] can be an issue for data gatherers for ML algorithms. Prediction models created on biased qualitative data can also affect trust in AI applications. If these models do not perform accurately in critical situations due to the biases, it could even be a source of rejection of models[20], [21].

2.2 Chemicals- Depression

Low levels of certain neurotransmitters can contribute to depression though researchers are still investigating the role of these brain chemicals play in depression [22]. One such chemical that can be used as a predictor for depression is amino-acid neurotransmitter [23]. But predictions based on quantitative data like chemicals in the biochemistry profile

using ML are scarce. [24] reviewed possible links between magnesium and depression and a higher intake of magnesium seemed to be associated with lower depression. Another study [25] found that sodium, potassium, and calcium levels do not seem to be related to the intensity of major depression in contrast of magnesium. For the patients with major depressive disorder (MDD), the level of routine biochemical markers like urea nitrogen (UN), fasting blood-glucose (FBG), high density lipoprotein-cholesterol (HDL-C), total protein (TP), creatinine (Cr), total bilirubin (Tbil), fructosamine (SF) were found to be associated with MDD [26]. There has been also research on looking into the relation between several biochemical markers and depression [27]–[33] separately, but overall the results have been inconsistent and majorly focused on MDD only. Most of them also were focused on only a few of the biochemical markers, so it is difficult to understand the pattern of association of these markers with depression overall. In this paper, we focused on finding a link between depression (both major and minor) and overall biochemistry profiles using a robust data source from National Health and Nutrition Examination Survey (NHANES) that uses a multistage, stratified survey methodology to obtain data representative of the general, noninstitutionalized US population.

3. METHODOLOGY

We used the NHANES dataset from 2011-2018 where almost 19000 people ages 12 years and older submitted their biochemistry profile and mental health-related questionnaire. Since people had the option of not answering mental health-related questions, data may be devoid of any biases. For prediction, we will use Random Forest (RF), Decision Tree (DT), K-nearest neighbor (KNN), Logistic Regression (LR), Stochastic Gradient Descent (SGD), Naive Bayes (NB), and Gradient Boosting (GB) classifiers to find out the best fit results. Synthetic data will be created using the synthpop library in R [34] for balancing positive and negative data (positive data are where participants acknowledged they have depression, negative data are where participants acknowledged they do not have depression). For major and minor depression models, data will be divided into two sets; for minor depression, data, where participants reported they are depressed for several days in the last two weeks, will be counted. For major depression, participants who reported they are depressed for more than half the days or nearly every day in the last two weeks will be counted.

In this paper, 9900 synthetic positive data was added to the original data (where patients with both types of depression are included) because the original data set had 14531 negative data against 4632 positive data. For the minor depression data set, 11000 synthetic positive data was added to the data set because the data set had 14531 negative data against 3197 positive data. For the major depression data set, 12500 synthetic positive data was added to the data set because the data set had 14531 negative data against 1435 positive data.

We will compare the result from the models with traditional qualitative depression prediction models.

The title (Helvetica 18-point bold), authors' names (Helvetica 12-point) and affiliations (Helvetica 10-point) run across the full width of the page – one column wide. We also recommend e-mail address (Helvetica 12-point). See the top of this page for three addresses. If only one address is needed, center all address text. For two addresses, use two centered tabs, and so on. For three authors, you may have to improvise.

4. DATA

From the NHANES 2011-2018 laboratory data, we will use different biochemicals from the human biochemistry profile to predict depression (see Table 1). We will also use refrigerated serums to make the models better (see Table 2).

This set of data has all the important variables on the body's biochemistry that are used in diagnosing different types of diseases [35]. NHANES's datasets are significant since they have been used in many different depression-related studies [36]–[39].

Table 1. Main Biochemical List

Chemical Name	Quantity	Variable Name in the Data Set
Alanine Aminotransferase (ALT)	IU/L	lbxsatsi
Alkaline Phosphatase (ALP)	IU/L	lbxsapsi
Aspartate Aminotransferase (AST)	U/L	lbxsassi
Bicarbonate	mmol/L	lbxsc3si
Blood Urea Nitrogen	g/dL	lbxsbu
Chloride	mmol/L	lbxsclsi
Creatine Phosphokinase (CPK)	IU/L	lbxsck
Globulin	g/dL	lbxsgb
Gamma Glutamyl Transferase (GGT)	U/L	lbxsgtsi
Lactate Dehydrogenase (LDH)	U/L	lbxsldsi
Osmolality	mmol/Kg	lbxsossi
Phosphorus	mg/dL	lbxsph
Potassium	mmol/L	lbxsksi
Sodium	mmol/L	lbxsnasi
Total Bilirubin	mg/dL	lbxstb
Total Calcium	mg/dL	lbxsca
Total Protein	g/dL	lbxstp
Uric Acid	mg/dL	lbxsua

Table 2. Refrigerated Serums

Chemical Name	Quantity	Variable Name in the Data Set
Glucose	mg/dL	lbxsgl
Albumin	g/dL	lbxsal
Creatinine	mg/dL	lbxscr
Iron	ug/dL	lbxsir
Cholesterol	mg/dL	lbxsch
Triglycerides	mg/dL	lbxstr

5. RESULTS

We shuffled each data set and divided them into three parts: test, validation, and training set (70% training data, 15% validation data, 15% test data). During predictions, Decision Tree and K-nearest neighbor did well in the training set, but accuracy dropped significantly during validation. The code for this project can be found in this GitHub repository[1].

5.1 Prediction of Depression

Both Random Forest and Gradient Boosting showed Area Under Curve (AUC) greater than 0.65 (see Fig. 1). Both the classifiers showed an underfitting tendency.
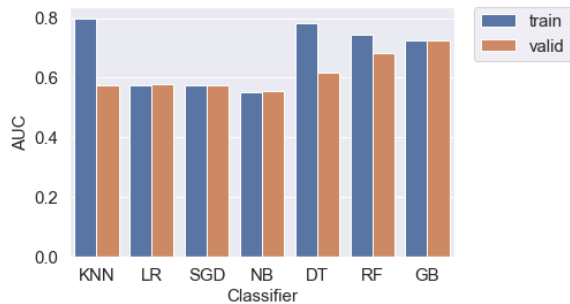
Fig 1: Test and Validation Results for the Classifiers

Hyperparameter tuning did improve the models (Gradient Boosting and Random Forest) especially for Random Forest still the AUC value is low than Gradient Boosting. After Hyperparameter tuning AUC value for the model improves from 0.72 to 0.74 (base to optimized) for Gradient Boosting and 0.68 to 0.71 (base to optimized) for Random Forest.

As Gradient Boosting has the highest validation AUC value, we used Gradient Boosting for testing. AUC value for testing was 0.72 (see Fig.2).

Both Gradient Boosting and Random Forest gave the same important features, lbxsck, lbxsatsi, lbxscr, lbxsua, lbxsassi (importance score greater than 0.06) though in a different order. This means Creatine Phosphokinase, Alanine Aminotransferase, Creatinine, Uric Acid, Aspartate Aminotransferase are good predictors for depression (see Fig. 2).

5.2 Prediction of Minor Depression

Classifiers for minor depression data set showed a similar trend during validation and training phases like Fig.1. After Hyperparameter tuning AUC value for the model improves from 0.73 to 0.74 (base to optimized) for Gradient Boosting and 0.68 to 0.71 (base to optimized) for Random Forest.
Testing AUC for Gradient Boosting was 0.74 (see Fig. 4),which is an improvement from the previous test AUC score for this classifier.
A significant change can be observed in the importance scores for predictors. Though we get the same set of predictors for Gradient Boosting in a different order (see Fig. 3 and Fig. 4), and if we compare the importance scores for the top predictors (importance score greater than 0.06) in Fig.3 and Fig.5, the importance scores are a lot closer in Fig.4 for the top predictors.
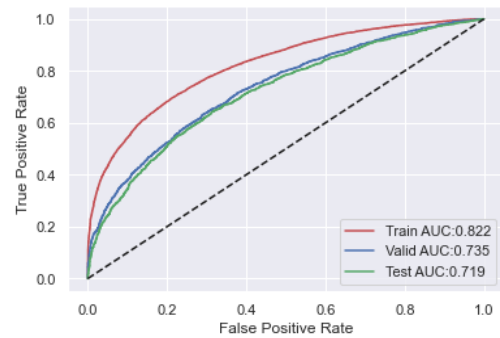
---

[1]https://github.com/tm012/biochemistry_depression_python
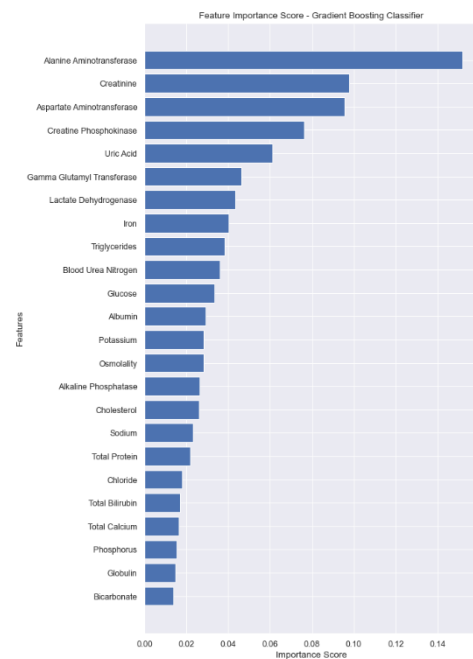

Fig2: Learning Curves for Gradient Boosting – Depression

Fig3: Feature Importance Score - Gradient Boosting (Depression)

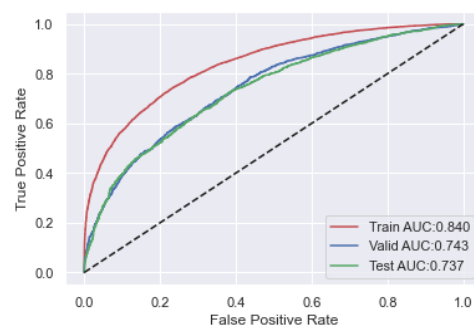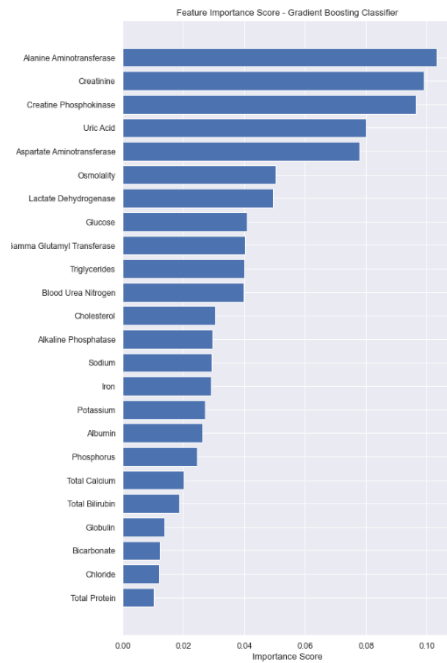**Fig 4: Learning Curves for Gradient Boosting – Minor Depression**

Fig5: Feature Importance Score - Gradient Boosting (Minor Depression)
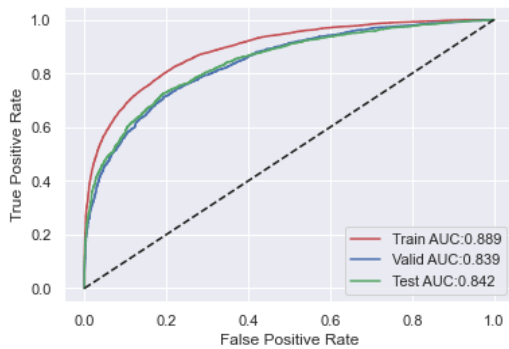

Fig 6: Learning Curves for Gradient Boosting – Major Depression

5.3 Prediction of Major Depression

Both Random Forest and Gradient Boosting Classifier stand out during training and testing phases like earlier classifications. After Hyperparameter tuning AUC value for the model improves from 0.82 to 0.84 (base to optimized) for Gradient Boosting and 0.75 to 0.79 (base to optimized) for Random Forest.

AUC value for testing was 0.84 (see Fig.6) for Gradient Boosting. The AUC test curve is a lot closer to the AUC training curve in this part.

The importance score also gives a different set of predictors for major depression. If we compare the top predictors for major and minor depression where top predictors have an importance score greater than 0.06 (see Figs. 5 and 7), we get lbxsgtsi (Gamma Glutamyl Transferase) instead of lbxsua (Uric Acid) and all other features are the same.

6. DISCUSSION

The result from three types of predictions shows us that prediction targeting minor and major depressions rather than predicting just depression shows better results (see Table 3). Astonishingly, for predicting major depression our best model (Gradient Boosting) shows a precision of 0.79 in test data.

This indicates we can achieve a low false-positive rate from this model. True Negative rate is also high for this model (0.82). Recall is 0.72 for this model meaning we will be able to catch 72% of probable patients who are suffering from major depression. For minor depression, the recall for the best model is 0.65 meaning we will be able to catch 65% of probable patients who are suffering from minor depression. As we discussed earlier, since people are not keen to seek help for depression, our best model can identify a large portion of depressed patients.
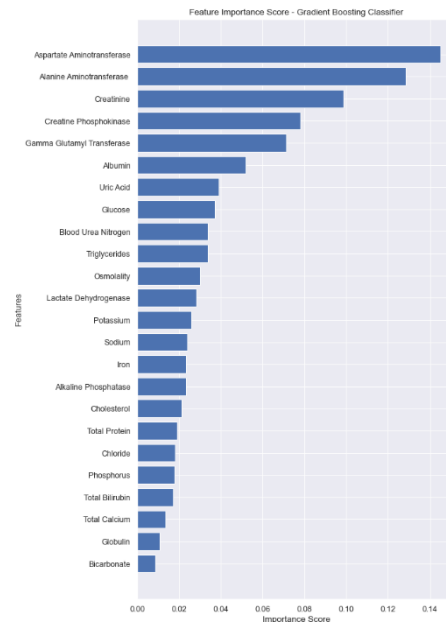

Fig 7: Feature Importance Score – Gradient Boosting (Major Depression)

Reduced features (feature importance score equal, or greater than 0.04) does not change the models' performance by that much from the original models (see Table 3 and 4).

Though the results (for original models) do not show high accuracy of detection like traditional qualitative depression models, the procedure we proposed can be used especially in rural and remote areas for depression screening because this type of area often lacks good mental health care facilities to do any qualitative-based screening or in a society where 'mental help' is still a tabooed phrase. A patient's biochemistry profile can be obtained by a health care provider while a patient comes for a routine/urgent health check-up or part of a depression screening process.

Table 3. Test Results from Three Types of Predictions by the Models (Gradient Boosting)

	AUC	Re-call	Preci-sion	Specifi-city	F-Score
Depression	0.72	0.63	0.68	0.70	0.65
Minor Depression	0.74	0.65	0.67	0.69	0.66
Major Depression	0.84	0.72	0.79	0.82	0.75

Table 4. Test Results from Three Types of Predictions by the Models with reduced features (Gradient Boosting)

	AUC	Recall	Precision	Specificity	F-Score
Depression	0.70	0.57	0.67	0.71	0.62
Minor Depression	0.73	0.64	0.66	0.69	0.65
Major Depression	0.82	0.68	0.76	0.79	0.72

7. LIMITATIONS

In the earlier part of this paper (see Methodology section), we mentioned that a questionnaire was used for mental health data. Though we argued why this data set may have less biased data, there is always a chance of significant biases in qualitative data. Some biochemicals like magnesium that are helpful for detecting depression were absent from this biochemistry profile.

8. CONCLUSION

Healthcare cost for depression is rising because it is a source of many comorbidities. After analyzing NHANES 2011-2018 Hospital Utilization & Access to Care data, for major depression, out of 1449 patients 831 patients seen mental health professional(s) in the past year. This number falls for minor depression (524 out of 3209). So, for better risk stratification, the healthcare providers can use the proposed method; that shows a good screening process based on quantitative data. This way they can identify patients for depression and can begin treatment for depression. This will lower healthcare-related costs by preventing comorbidities and stopping depression to worsen.

9. REFERENCES

[1] H. Ritchie and M. Roser, "Mental Health," *Our World Data*, Jan. 2018, Accessed: Feb. 15, 2021. [Online]. Available: https://ourworldindata.org/mental-health.

[2] P. E. Greenberg, A.-A. Fournier, T. Sisitsky, C. T. Pike, and R. C. Kessler, "The economic burden of adults with major depressive disorder in the United States (2005 and 2010)," *J. Clin. Psychiatry*, vol. 76, no. 2, pp. 155–162, 2015.

[3] L. L. Judd, M. P. Paulus, K. B. Wells, and M. H. Rapaport, "Socioeconomic burden of subsyndromal depressive symptoms and major depression in a sample of the general population.," 1996.

[4] D. A. Regier, W. E. Narrow, D. S. Rae, R. W. Manderscheid, B. Z. Locke, and F. K. Goodwin, "The de facto US mental and addictive disorders service system: Epidemiologic Catchment Area prospective 1-year prevalence rates of disorders and services," *Arch. Gen. Psychiatry*, vol. 50, no. 2, pp. 85–94, 1993.

[5] J. C. Coyne, S. Fechner-Bates, and T. L. Schwenk, "Prevalence, nature, and comorbidity of depressive disorders in primary care," *Gen. Hosp. Psychiatry*, vol. 16, no. 4, pp. 267–276, 1994.

[6] A. Halfin, "Depression: the benefits of early and appropriate treatment," *Am. J. Manag. Care*, vol. 13, no. 4, p. S92, 2007.

[7] A. Lasalvia *et al.*, "Global pattern of experienced and anticipated discrimination reported by people with major depressive disorder: a cross-sectional survey," *The Lancet*, vol. 381, no. 9860, pp. 55–62, 2013.

[8] K. B. Tharaldsen, P. Stallard, P. Cuijpers, E. Bru, and J. F. Bjaastad, "'It's a bit taboo': a qualitative study of Norwegian adolescents' perceptions of mental healthcare services," *Emot. Behav. Difficulties*, vol. 22, no. 2, pp. 111–126, 2017.

[9] R. A. Miech and M. J. Shanahan, "Socioeconomic status and depression over the life course," *J. Health Soc. Behav.*, pp. 162–176, 2000.

[10] B. G. Link, M. C. Lennon, and B. P. Dohrenwend, "Socioeconomic status and depression: The role of occupations involving direction, control, and planning," *Am. J. Sociol.*, vol. 98, no. 6, pp. 1351–1387, 1993.

[11] N. Akhtar-Danesh and J. Landeen, "Relation between depression and sociodemographic factors," *Int. J. Ment. Health Syst.*, vol. 1, no. 1, pp. 1–9, 2007.

[12] A. Barua, M. K. Ghosh, N. Kar, and M. A. Basilio, "Socio-demographic factors of geriatric depression," *Indian J. Psychol. Med.*, vol. 32, no. 2, pp. 87–92, 2010.

[13] E. M. de Souza Filho *et al.*, "Can machine learning be useful as a screening tool for depression in primary care?," *J. Psychiatr. Res.*, vol. 132, pp. 1–6, Jan. 2021, doi: 10.1016/j.jpsychires.2020.09.025.

[14] A. Sau and I. Bhakta, "Predicting anxiety and depression in elderly patients using machine learning technology," *Healthc. Technol. Lett.*, vol. 4, no. 6, pp. 238–243, 2017.

[15] A. Priya, S. Garg, and N. P. Tigga, "Predicting anxiety, depression and stress in modern life using machine learning algorithms," *Procedia Comput. Sci.*, vol. 167, pp. 1258–1267, 2020.

[16] M. R. Islam, M. A. Kabir, A. Ahmed, A. R. M. Kamal, H. Wang, and A. Ulhaq, "Depression detection from social network data using machine learning techniques," *Health Inf. Sci. Syst.*, vol. 6, no. 1, pp. 1–12, 2018.

[17] M. Trotzek, S. Koitka, and C. M. Friedrich, "Utilizing neural networks and linguistic metadata for early detection of depression indications in text sequences," *IEEE Trans. Knowl. Data Eng.*, vol. 32, no. 3, pp. 588–601, 2018.

[18] N. Schwarz, "Self-reports: how the questions shape the answers.," *Am. Psychol.*, vol. 54, no. 2, p. 93, 1999.

[19] T. L. Jones, M. A. J. Baxter, and V. Khanduja, "A quick guide to survey research," *Ann. R. Coll. Surg. Engl.*, vol. 95, no. 1, pp. 5–7, 2013.

[20] L. Alam, "Investigating the Impact of Explanation on Repairing Trust in Ai Diagnostic Systems for Re-Diagnosis," 2020.

[21] Mueller, S.T., Veinott, E.S., Hoffman, R.R., Klein, G., Alam, L., Mamun, T. and Clancey, W.J., "Principles of Explanation in Human-AI Systems." *arXiv preprint arXiv:2102.04972*, 2021.

[22] "What Is the Chemistry Behind Depression?," *Verywell Mind*. https://www.verywellmind.com/the-chemistry-of-depression-1065137 (accessed Feb. 15, 2021).

[23] Salvadore, G., van der Veen, J.W., Zhang, Y., Marenco, S., Machado-Vieira, R., Baumann, J., Ibrahim, L.A., Luckenbaugh, D.A., Shen, J., Drevets, W.C. and Zarate Jr, C.A., "An investigation of amino-acid neurotransmitters as potential predictors of clinical improvement to ketamine in depression," *Int. J. Neuropsychopharmacol.*, vol. 15, no. 8, pp. 1063–1072, 2012.

[24] M.-L. Derom, C. Sayón-Orea, J. M. Martínez-Ortega, and M. A. Martínez-González, "Magnesium and depression: a systematic review," *Nutr. Neurosci.*, vol. 16, no. 5, pp. 191–206, Sep. 2013, doi: 10.1179/1476830512Y.0000000044.

[25] Widmer, J., Mouthon, D., Raffin, Y., Chollet, D., Hilleret, H., Malafosse, A., &Bovier, P.,"Weak Association between Blood Sodium, Potassium, and Calcium and Intensity of Symptoms in Major Depressed Patients," *Neuropsychobiology*, vol. 36, no. 4, pp. 164–171, 1997, doi: 10.1159/000119378.

[26] Y.-F. Peng, Y. Xiang, and Y.-S. Wei, "The significance of routine biochemical markers in patients with major depressive disorder," *Sci. Rep.*, vol. 6, no. 1, Art. no. 1, Sep. 2016, doi: 10.1038/srep34402.

[27] A. Parekh, D. Smeeth, Y. Milner, and S. Thuret, "The role of lipid biomarkers in major depression," in *Healthcare*, 2017, vol. 5, no. 1, p. 5.

[28] D. Rafter, "Biochemical markers of anxiety and depression," *Psychiatry Res.*, vol. 103, no. 1, pp. 93–96, Aug. 2001, doi: 10.1016/S0165-1781(01)00251-7.

[29] Glueck, C. J., Tieger, M., Kunkel, R., Tracy, T., Speirs, J., Streicher, P., &Illig, E., "Improvement in symptoms of depression and in an index of life stressors accompany treatment of severe hypertriglyceridemia," *Biol. Psychiatry*, vol. 34, no. 4, pp. 240–252, 1993.

[30] B. M. Ross, P. Ward, and I. Glen, "Delayed vasodilatory response to methylnicotinate in patients with unipolar depressive disorder," *J. Affect. Disord.*, vol. 82, no. 2, pp. 285–290, 2004.

[31] Maes, M., Vandoolaeghe, E., Neels, H., Demedts, P., Wauters, A., Meltzer, H. Y., Altamura, C. andDesnyder, R., "Lower serum zinc in major depression is a sensitive marker of treatment resistance and of the immune/inflammatory response in that illness," *Biol. Psychiatry*, vol. 42, no. 5, pp. 349–358, 1997.

[32] A.-M. Hvas, S. Juul, P. Bech, and E. Nexø, "Vitamin B6 level is associated with symptoms of depression," *Psychother. Psychosom.*, vol. 73, no. 6, pp. 340–343, 2004.

[33] Y. W.-Y. Yu, T.-J. Chen, Y.-C. Wang, Y.-J. Liou, C.-J. Hong, and S.-J. Tsai, "Association analysis for neuronal nitric oxide synthase gene polymorphism with major depression and fluoxetine response," *Neuropsychobiology*, vol. 47, no. 3, pp. 137–140, 2003.

[34] B. Nowok, G. M. Raab, and C. Dibben, "synthpop: Bespoke creation of synthetic data in R," *J. Stat. Softw.*, vol. 74, no. 11, pp. 1–26, 2016.

[35] https://wwwn.cdc.gov/Nchs/Nhanes/2017-2018/BIOPRO_J.htm#Component_Description (accessed Feb. 16, 2021).

[36] J. Istvan, K. Zavela, and G. Weidner, "Body weight and psychological distress in NHANES I.," *Int. J. Obes. Relat. Metab. Disord. J. Int. Assoc. Study Obes.*, vol. 16, no. 12, pp. 999–1003, 1992.

[37] Jia, Z., Li, X., Yuan, X., Zhang, B., Liu, Y., Zhao, J., and Li, S., "Depression is associated with diabetes status of family members: NHANES (1999–2016)," *J. Affect. Disord.*, vol. 249, pp. 121–126, 2019.

[38] R. S. Shim, P. Baltrus, J. Ye, and G. Rust, "Prevalence, treatment, and control of depressive symptoms in the United States: results from the National Health and Nutrition Examination Survey (NHANES), 2005–2008," *J. Am. Board Fam. Med.*, vol. 24, no. 1, pp. 33–38, 2011.

[39] J. K. Vallance, E. A. Winkler, P. A. Gardiner, G. N. Healy, B. M. Lynch, and N. Owen, "Associations of objectively-assessed physical activity and sedentary time with depression: NHANES (2005–2006)," *Prev. Med.*, vol. 53, no. 4–5, pp. 284–288, 2011.