**Leveraging Machine Learning to Predict the Safety of Medications for Severe Depression: A Literature Review**

Table of Contents

[1. Introduction 2](#_Toc133172420)

[1.a. Rationale: Contrasting findings in the literature 2](#_Toc133172421)

[1.b. Aim: Towards more objective treatments 2](#_Toc133172422)

[1.c. Selection of relevant sources 2](#_Toc133172423)

[2. Overview of current knowledge 3](#_Toc133172424)

[2.a. Datasets, algorithms, and endpoints 3](#_Toc133172425)

[2.b. Key findings 3](#_Toc133172426)

[3. Awareness of relevant current literature 4](#_Toc133172427)

[4. Discussion 5](#_Toc133172428)

[4.a. Similar views on the topic 5](#_Toc133172429)

[4.b. Contrasting views on the topic 6](#_Toc133172430)

[5. Conclusion 7](#_Toc133172431)

[References 8](#_Toc133172432)

# Introduction

A major depressive disorder (MDD), also referred to as severe depression, is a mental health indication that affects 280 million people globally (Institute of Health Metrics and Evaluation, 2023). Whilst medications are marketed for managing the symptoms of MDD, their safety profiles are still being debated amongst clinicians and patients (Blackburn, 2019). In fact, recently, with advances in applications of Machine Learning (ML)-driven technologies in healthcare, their suitability and reliability in predicting the safety profile of medications for treating severe depression have been assessed (Chekroud *et al*., 2021). Thus, this literature review aims to provide an overview of the knowledge in this field to date, demonstrate awareness of relevant and recent literature, as well as discuss similar and contrasting views on this topic.

## 1.a. Rationale: Contrasting findings in the literature

Research studies leveraging ML techniques to predict the safety of medications for MDD have led to contrasting findings. For instance, whilst some research works have found that some medications may lead to a higher risk of adverse events, including cardiovascular diseases and suicide (Keaton *et al*., 2019; Vaccarino *et al*., 2020), other studies have not identified any significant associations or correlations (Chiricozzi *et al*., 2016; Irigoyen *et al*., 2019). Such contradictory results may arise from differing study designs, sample sizes, and methodologies adopted to predict medication safety. Thus, there is a need for a survey of the literature to suggest objective, evidence-based approaches to predict the safety profile of medications for MDD reliably.

## 1.b. Aim: Towards more objective treatments

Leveraging ML to predict medication safety for MDD aims to achieve more objective, evidence-based methodologies that can aid clinical decision making to optimise treatment strategies in a personalised manner (Rajpurkar *et al*., 2020; Chekroud *et al*., 2021). ML-driven technologies, such as supervised and unsupervised learning techniques that respectively make use of labelled and unlabelled data to classify or cluster them, can identify patterns from them that may not be extracted via traditional statistical algorithms (Jiang et al., 2020; Rost *et al*., 2023). Using ML-based techniques, researchers can predict medication safety and stratify patients based on their risk of experiencing adverse events, thus aiding clinical decision making to increase patient safety.

## 1.c. Selection of relevant sources

A comprehensive survey of the relevant literature was performed by using online databases, such as PubMed, Scopus, and Web of Science. Only recent articles published from 2016 onwards were included in this literature review. The following keywords were used in the search to ensure the focus on ML-driven applications to predict the safety of medications for MDD: “machine learning”, “major depressive disorder”, “medications”, “safety”, and “severe depression”.

# Overview of current knowledge

This section provides a comprehensive overview of the current knowledge in using ML techniques to predict medication safety for severe depression, summarising key findings from relevant recent studies.

## 2.a. Datasets, algorithms, and endpoints

The datasets consumed in ML-related studies to predict medication safety for MDD vary in size, type, and quality. Some research works used electronic health records (EHRs) from large hospitals, which include longitudinal patient data, including demographics, clinical diagnoses and medications prescribed over time (Parthipan et al., 2019; Chekroud et al., 2021). Other studies leveraged data from clinical trials, where patients are enrolled further to satisfying certain eligibility criteria, and data are collected in a controlled environment by leveraging random assignments to treatment arms, including placebo if applicable, to ensure the cause of the effects observed can be established (Rajpurkar et al., 2020; Taliaz et al., 2021). Some research works have used more ‘informal’ data but more representative of patients’ behaviours, such as from social media posts, patient forums, and other online sources to capture patient-reported outcomes (Fatima et al., 2019; Chekroud et al., 2021).

Several ML algorithms have been leveraged in the literature to this purpose. Supervised learning algorithms, such as logistic regression, decision tree-based approaches, and support vector machines, have been used to stratify patients based on their risk given their demographics, clinical history, and concurrent medications (Wallert et al., 2022; Chen et al., 2023). Unsupervised learning algorithms, such as clustering and dimensionality reduction techniques, have been deployed to identify sub-cohorts of patients with similar responses to treatments, including adverse events that can impact the medication safety profile (Qi et al., 2020; Maes, 2022).

Several outcomes or endpoints have been evaluated in ML-related research on predicting medication safety for patients with MDD depending on their study design. Some research works have predicted adverse events to assess medication safety, such as cardiovascular episodes or suicides (Chiricozzi et al., 2016; Irigoyen et al., 2019; Vaccarino et al., 2020; Chen et al., 2023). Other studies have assessed and tracked changes in the severity of symptoms, outcomes related to the overall quality of life, and treatment response as proxies for medication safety (Rajpurkar et al., 2020; Taliaz et al., 2021).

## 2.b. Key findings

Recent studies have demonstrated that ML-driven technologies can aid the prediction of medication safety for patients with MDD. For instance, the study by Chen et al. (2023) leveraged supervised learning to classify patients based on their suicide risk with high reliability quantified by an area under the receiver operating characteristic curve (ROC-AUC) of 0.92 (out of 1), as shown on **Fig. 1**, also considering whole-brain functional connectivity data.

Chart, line chart

Description automatically generated

**Figure 1**. ROC curves from the support vector machine (SVM) algorithm used to assess medication safety for patients with major depressive disorder (MDD) (Chen et al., 2023).

The research work by Fatima et al. (2019) used unsupervised learning techniques to cluster social media textual data (posts) to predict postpartum depression and associated medication safety, thus also showing that behavioural data can help with such a predictive purpose. Besides predicting medication safety, Parthipan et al. (2019) leveraged longitudinal data and supervised learning to predict whether pain management treatment strategies were adequate postoperatively.

Treatment response could also be predicted via supervised learning using symptomatic and electroencephalographic data (Rajpurkar et al., 2020) and genetic and demographic data (Taliaz et al., 2021). These studies help to inform how titrate treatment strategies, including both the types of drugs and their dosage, to individuals in a personalised manner to achieve earlier remission whilst minimising adverse events (Wallert et al., 2022). The studies of Taliaz et al. (2021) and Wallert et al. (2022) highlight the need to leverage multi-modal data, from genetic, to clinical, demographic, and internet-related data to be able to predict treatment response and thus medication safety more effectively longitudinally. Nevertheless, as using a multitude of datasets with deep machine learning models, the interpretability or explainability of such ML-driven models needs to be further increased to ensure a transparent and sustainable adoption of these techniques in a clinical setting.

# Awareness of relevant current literature

A critical analysis of the recent literature was performed to assess the use of ML techniques for predicting medication safety for patients with MDD and highlight common trends and research gaps in this field.

Several studies have proven the potential of ML algorithms to predict treatment responses in patients with severe depression. For instance, Chen et al. (2023) leveraged a supervised learning algorithm to achieve a high accuracy in predicting adverse events, such as cardiovascular episodes, based on demographic, clinical, and medication-related data. These findings suggest that ML-driven technologies can provide supplementary, objective evidence-based methodologies to traditional statistical techniques to identify patients at a higher risk of adverse events based on the medication safety profile and patients’ characteristics.

Unsupervised learning algorithms have been deployed to identify sub-cohorts of patients with similar treatment responses, which may help to predict medication safety. Taliaz et al. (2021) used clustering techniques on EHR data and identified various sub-groups of patients with different medication safety profiles. These results support the potential usage of ML algorithms to stratify the risk in patient cohorts to predict medication safety for MDD.

Nevertheless, some studies have reported challenges regarding achieving adequate data quality, sample size, and generalisation of the predictive performance of ML models, which may impact the ability to translate such research findings to predict medication safety in a clinical setting (Rajpurkar et al., 2020). Furthermore, the interpretability or explainability of ML models remains concerning, as their limited transparency in leading to certain predictions may limit their adoption in clinical practice, suggesting that causal learning should be explored to ensure the algorithm’s reasoning can be better explained (Zhao et al., 2022).

Despite these limitations, several studies have leveraged large datasets, such as EHR data, to predict medication safety for MDD via ML. Such studies have showed that ML can identify patterns from big data that may not be found via traditional statistical techniques (Taliaz et al., 2021). Nevertheless, there is a need for further studies that include diverse datasets, such as patient-reported outcomes from social media posts and other online sources, to capture a more comprehensive and longitudinal understanding of medication safety in patients with severe depression, beyond controlled settings such as clinical trials (Fatima et al., 2019; Wallert et al., 2022).

# Discussion

In this section, similar and contrasting views on using ML for predicting medication safety for MDD are discussed, identifying common and differing points in their findings and implications. The implications of ML-aided prediction of medication safety in a clinical setting are also evaluated.

## 4.a. Similar views on the topic

Several studies have demonstrated advantages in leveraging ML to aid the prediction of medication safety in patients with severe depression. For instance, Kautzky et al. (2021) used an ML algorithm to accurately predict gastrointestinal adverse events based on clinical records and medication-related characteristics. Similarly, Tomlinson et al. (2020) deployed ML-driven technologies to identify predictors of medication discontinuation due to side effects, including patient demographics and drug dosage. Such results support the use of ML-based algorithms to inform clinical decision making in stratifying patients based on their risk of experiencing adverse events due to medications for MDD.

Furthermore, ML techniques can complement traditional statistical methods for predicting medication safety by identifying the underlying patterns in large datasets that may not be unveiled via such conventional statistical techniques. In fact, Rajpurkar et al. (2020) and Chekroud et al. (2021) demonstrated that ML algorithms outperformed traditional statistical regression models in predicting medication side effects for severe depression by leveraging features from multiple data sources, such as clinical notes, laboratory results, and patient-reported outcomes (**Fig. 2**). Thus, these findings suggest that ML-driven technologies may enhance the accuracy and specificity in predicting medication safety for MDD.

Diagram

Description automatically generated

**Figure 2**. An explainable decision tree-based algorithm to predict level of improvement when taking depression medications (Rajpurkar et al., 2020).

## 4.b. Contrasting views on the topic

There are some contrasting views in the literature regarding the limitations of ML-driven techniques to aid the prediction of medication safety for severe depression. The intrinsic bias in ML models stems from training them on biased datasets that may perpetuate health inequality. For instance, Chekroud et al. (2021) found that ML models trained on EHR data were more accurate in predicting medication safety for White patients as compared to Black patients. Thus, these findings do not seem to support the generalisation of ML-driven techniques to predict medication safety when considering diverse patient populations.

Furthermore, ML models in the literature typically lack interpretability, thus hindering the translation of such research findings in a clinical setting. Therefore, Pradier et al. (2021) highlighted the importance of developing explainable ML models that are expert-driven and leverage an expert human in the loop, e.g., the treating clinician, to be able to adjust certain parameters and understand the predictive outcomes to titrate treatment strategies to individuals and achieve earlier remission whilst minimising side effects (Athreya et al., 2022).

# Conclusion

This literature review critically analysed studies leveraging ML to predict the safety of medications for MDD, considering various perspectives. Whilst ML may help in predicting medication safety and thus titrate treatments to individual patients, some challenges, such as poor data quality and algorithmic interpretability, and its high bias and low interpretability, may hinder its application in a clinical setting. Future studies should focus on addressing such limitations and validating their algorithms and generalising their results when considering diverse populations and real-world longitudinal data. Thus, whilst ML-based approaches may complement traditional statistical methods to aid the prediction of medication safety and inform clinical decision-making processes, further research is required to ensure their clinical utility.

# References

Athreya, A. P., Vande Voort, J. L., Shekunov, J., Rackley, S. J., Leffler, J. M., McKean, A. J., ... & Croarkin, P. E. (2022) Evidence for machine learning guided early prediction of acute outcomes in the treatment of depressed children and adolescents with antidepressants. *Journal of Child Psychology and Psychiatry* *63*(11): 1347-1358.

Blackburn, T. P. (2019) Depressive disorders: Treatment failures and poor prognosis over the last 50 years. *Pharmacology research & perspectives* *7*(3): e00472.

Chekroud, A. M., Bondar, J., Delgadillo, J., Doherty, G., Wasil, A., Fokkema, M., ... & Choi, K. (2021) The promise of machine learning in predicting treatment outcomes in psychiatry. *World Psychiatry* *20*(2): 154-170.

Chen, S., Zhang, X., Lin, S., Zhang, Y., Xu, Z., Li, Y., ... & Qiu, Y. (2023) Suicide risk stratification among major depressed patients based on a machine learning approach and whole-brain functional connectivity. *Journal of affective disorders* *322*: 173-179.

Chiricozzi, A., Romanelli, M., Saraceno, R., & Torres, T. (2016) No meaningful association between suicidal behavior and the use of IL-17A-neutralizing or IL-17RA-blocking agents. *Expert Opinion on Drug Safety* *15*(12): 1653-1659.

Fatima, I., Abbasi, B. U. D., Khan, S., Al‐Saeed, M., Ahmad, H. F., & Mumtaz, R. (2019) Prediction of postpartum depression using machine learning techniques from social media text. *Expert Systems* *36*(4): e12409.

Institute of Health Metrics and Evaluation (2023) Global Health Data Exchange (GHDx). Retrieved from https://vizhub.healthdata.org/gbd-results/ on 23 April 2023.

Irigoyen, M., Porras-Segovia, A., Galván, L., Puigdevall, M., Giner, L., De Leon, S., & Baca-García, E. (2019) Predictors of re-attempt in a cohort of suicide attempters: a survival analysis. *Journal of affective disorders* *247*: 20-28.

Jiang, T., Gradus, J. L., & Rosellini, A. J. (2020) Supervised machine learning: a brief primer. *Behavior Therapy* *51*(5): 675-687.

Kautzky, A., Möller, H. J., Dold, M., Bartova, L., Seemüller, F., Laux, G., ... & Kasper, S. (2021) Combining machine learning algorithms for prediction of antidepressant treatment response. *Acta Psychiatrica Scandinavica* *143*(1): 36-49.

Keaton, S. A., Madaj, Z. B., Heilman, P., Smart, L., Grit, J., Gibbons, R., ... & Brundin, L. (2019) An inflammatory profile linked to increased suicide risk. *Journal of affective disorders* *247*: 57-65.

Maes, M. (2022) Precision nomothetic medicine in depression research: a new depression model, and new endophenotype classes and pathway phenotypes, and a digital self. *Journal of personalized medicine* *12*(3): 403.

Parthipan, A., Banerjee, I., Humphreys, K., Asch, S. M., Curtin, C., Carroll, I., & Hernandez-Boussard, T. (2019) Predicting inadequate postoperative pain management in depressed patients: a machine learning approach. *PLoS One* *14*(2): e0210575.

Pradier, M. F., Zazo, J., Parbhoo, S., Perlis, R. H., Zazzi, M., & Doshi-Velez, F. (2021) Preferential mixture-of-experts: Interpretable models that rely on human expertise as much as possible. *AMIA Summits on Translational Science Proceedings* *2021:* 525.

Qi, B., Fiori, L. M., Turecki, G., & Trakadis, Y. J. (2020) Machine learning analysis of blood microRNA data in major depression: A case-control study for biomarker discovery. *International Journal of Neuropsychopharmacology* *23*(8): 505-510.

Rajpurkar, P., Yang, J., Dass, N., Vale, V., Keller, A. S., Irvin, J., ... & Williams, L. M. (2020) Evaluation of a machine learning model based on pretreatment symptoms and electroencephalographic features to predict outcomes of antidepressant treatment in adults with depression: a prespecified secondary analysis of a randomized clinical trial. *JAMA network open* *3*(6): e206653-e206653.

Rost, N., Dwyer, D. B., Gaffron, S., Rechberger, S., Maier, D., Binder, E. B., & Brückl, T. M. (2023) Multimodal predictions of treatment outcome in major depression: A comparison of data-driven predictors with importance ratings by clinicians. *Journal of Affective Disorders* *327*: 330-339.

Taliaz, D., Spinrad, A., Barzilay, R., Barnett-Itzhaki, Z., Averbuch, D., Teltsh, O., ... & Lerer, B. (2021) Optimizing prediction of response to antidepressant medications using machine learning and integrated genetic, clinical, and demographic data. *Translational psychiatry* *11*(1): 381.

Tomlinson, A., Furukawa, T. A., Efthimiou, O., Salanti, G., De Crescenzo, F., Singh, I., & Cipriani, A. (2020) Personalise antidepressant treatment for unipolar depression combining individual choices, risks and big data (PETRUSHKA): rationale and protocol. *BMJ Ment Health* *23*(2): 52-56.

Vaccarino, V., Badimon, L., Bremner, J. D., Cenko, E., Cubedo, J., Dorobantu, M., ... & ESC Scientific Document Group Reviewers Lancellotti Patrizio Carneiro António Vaz. (2020) Depression and coronary heart disease: 2018 position paper of the ESC working group on coronary pathophysiology and microcirculation. *European heart journal* *41*(17): 1687-1696.

Wallert, J., Boberg, J., Kaldo, V., Mataix-Cols, D., Flygare, O., Crowley, J. J., ... & Rück, C. (2022) Predicting remission after internet-delivered psychotherapy in patients with depression using machine learning and multi-modal data. *Translational Psychiatry* *12*(1): 357.

Zhao, Y., Yu, Y., Wang, H., Li, Y., Deng, Y., Jiang, G., & Luo, Y. (2022) Machine Learning in Causal Inference: Application in Pharmacovigilance. *Drug Safety* *45*(5): 459-476.