Review

Personalized Diagnosis and Treatment for Neuroimaging in Depressive Disorders

Jongha Lee [1], Suhyuk Chi [2] and Moon-Soo Lee [2,3,*]

1 Department of Psychiatry, Korea University Ansan Hospital, Ansan 15355, Korea
2 Department of Psychiatry, Korea University Guro Hospital, Seoul 08308, Korea
3 Department of Life Sciences, Korea University, Seoul 02841, Korea
* Correspondence: npboard@korea.ac.kr; Tel.: +82-2-2626-3163; Fax: +82-2-852-1937

Abstract: Depressive disorders are highly heterogeneous in nature. Previous studies have not been useful for the clinical diagnosis and prediction of outcomes of major depressive disorder (MDD) at the individual level, although they provide many meaningful insights. To make inferences beyond group-level analyses, machine learning (ML) techniques can be used for the diagnosis of subtypes of MDD and the prediction of treatment responses. We searched PubMed for relevant studies published until December 2021 that included depressive disorders and applied ML algorithms in neuroimaging fields for depressive disorders. We divided these studies into two sections, namely diagnosis and treatment outcomes, for the application of prediction using ML. Structural and functional magnetic resonance imaging studies using ML algorithms were included. Thirty studies were summarized for the prediction of an MDD diagnosis. In addition, 19 studies on the prediction of treatment outcomes for MDD were reviewed. We summarized and discussed the results of previous studies. For future research results to be useful in clinical practice, ML enabling individual inferences is important. At the same time, there are important challenges to be addressed in the future.

Keywords: major depressive disorder; resting-state functional connectivity; machine learning; classification; neuroimaging; diagnosis; prediction

1. Introduction

Psychiatric disorders such as depressive disorders are highly heterogeneous. Depressed mood and markedly diminished interest in previously enjoyed activities are key characteristics of major depressive disorder (MDD), but symptoms show numerous heterogeneous constellations. This indicates that the characteristics of MDD are not confined to certain single parameters but may associate with multiple bio–psycho–social dimensions. Thus, there is a need to integrate various data forms for better knowledge on the disease. Many past studies have identified a range of possible biomarkers for depressive disorders. However, these results have not yet been successfully integrated into clinical practice for diagnosis and treatment. In addition, there is heterogeneity in the clinical presentation of MDD and its responses to treatment. Accordingly, identifying applicable biomarkers that can predict diagnosis and treatment outcomes would be useful in clinical practice. Currently, depressive disorders are diagnosed by trained professionals using their clinical judgement. However, this process is time-consuming and depends on the subjective judgement of the clinician. In clinical practice, it is common not to receive additional help from brain imaging, although brain imaging provides a noninvasive evaluation of brain structure and function and provides a deeper understanding of the neuropathophysiology of MDD. Brain imaging studies should be introduced in the clinical management of depression. In this process, both objective biomarkers and subjective impressions of clinicians are required for accurate classification and treatment decisions. There is a need for objective biomarkers for more effective diagnosis and treatment of depression. From this perspective, we would

like to review the status of brain imaging studies by dividing them into two large groups: diagnostic biomarkers and prediction of treatment outcomes. For a more specific search, we focused on MDD, the most representative disease in depression.

Currently, there are many neuroimaging methodologies available. For example, magnetic resonance imaging (MRI) comprises many methods, including structural MRI (sMRI) and functional MRI (fMRI). sMRI has been conventionally used in patients with clinical depression, whereas the fMRI technique can provide other kinds of useful information, including brain functional connectivity, which provides useful insights into the pathophysiology of depression in relation to functional connectivity in the prefrontal–limbic and prefrontal–striatum systems in patients with MDD. However, such studies are not useful for the clinical diagnosis and prediction of MDD outcomes at an individual level [1]. There are also many clinical studies using neuroimaging in MDD, and most of them reported differences between patient and control populations, but clinicians need to make inferences at an individual level in most clinical settings [2]. The use of machine learning (ML) techniques is a possible way to make inferences beyond population-level analyses, both for the diagnosis of more specific subtypes of MDD and the prediction of treatment responses [3]. ML algorithms are powerful analytical tools that enable us to integrate neuroimaging and non-imaging data, helping to make decisions on the diagnosis and treatment outcomes of individual patients. ML can generalize patterns from the input data to generate a classification based on new data. Recently, deep learning, a particular branch of ML, has been increasingly used as it can even more effectively integrate neuroimaging data and non-imaging multimodal data [4].

In this paper, we summarize the regular and distinct findings of ML in neuroimaging studies of MDD (mainly focused on MRI) among mood disorders. We comprehensively reviewed the application of ML algorithms for neuroimaging in depressive disorders. A bibliographic search of PubMed and Google Scholar was conducted in December 2021. We searched PubMed through December 2021 for relevant studies that included depressive disorders and applied an ML algorithm in neuroimaging fields for depressive disorders. We divided these into diagnosis and treatment outcomes for the purpose of prediction using ML. In addition, we attempted to classify the studies according to the characteristics of the methods used, namely structural and functional studies. By doing this, we were able to follow the existing research results based on ML algorithms more intuitively.

2. Considerations for ML in Neuroimaging in Depressive Disorders: Diagnosis (Table 1)

Table 1. Selection of studies investigating machine learning methods for the prediction of diagnosis in depression.

References (Year)	Subjects (Mean Age)	Features	Machine Learning Method	Cross-Validation	Accuracy *	Comments
Foland-Ross et al., 2015 [5]	Baseline 33 adolescents (follow-up: 18 MDD and 15 HC)	Cortical thickness	SVM	stratified 10-fold cross validation	Average accuracy, 69.7%	Girls with an onset of MDD show baseline thinner right medial orbitofrontal cortex and thicker left insula
Kim et al., 2019 [6]	27 HC (15.96 ± 1.02) and 27 MDD (15.48 ± 1.72)	Cortical thickness	SVM	Double LOOCV	94.4% (sensitivity, 92.6% and specificity, 96.3%)	TreeBagging, RF, MLP, AdaBoost, and GBM were used, but they showed lower accuracies than SVM
Qiu et al., 2014 [7]	32 HC (35.0 ± 11.2) and 32 MDD (34.9 ± 11.1)	High-resolution T1-weighted imaging (morphometric parameters)	multivariate SVM	LOOCV	cortical thickness of right hemisphere, 78% ($p \leq 0.001$)	First-episode, medication-naïve MDD without any psychiatric comorbidities
Qin et al., 2014 [8]	30 HC (35.57 ± 11.73) and 29 MDD (38.97 ± 9.95)	DTI data	SVM with RBF kernel	LOOCV	83.05%	Hubs including the bilateral dorsolateral part of the superior frontal gyrus, the left middle frontal gyrus, the bilateral middle temporal gyrus, and the bilateral inferior temporal gyrus played an important role in diagnosing MDD
Patel et al., 2015 [9]	35 HC and 33 MDD	DTI data, structural imaging, functional imaging	Decision tree	LOOCV	87.3%	The optimal ADTree model selected MMSE score, age, whole brain atrophy, and fluid-attenuated inversion recovery Global WM hyperintensity count for predicting depression diagnosis
Wise et al., 2018 [10]	39 MDD (30.67 ± 8.71) and 8 BPD (29.50 ± 6.21)	High-resolution T1-weighted structural imaging	SVM	LOOCV		Greater gray volume predicted higher MADRS scores
Fung et al., 2015 [11]	19 MDD (30.0 ± 8.9), 16 BPD (26.3 ± 7.9) and HC (27.1 ± 8.4)	T1-weighted structural imaging (Cortical thickness, subcortical volume)	SVM	10-fold cross validation	74.3% (sensitivity, 62.5% and specificity, 84.2%)	Limitation: Effects of medication and chronicity of conditions in BPD and MDD on brain morphological alterations were not estimated
Deng et al., 2018 [12]	36 MDD (29.5 ± 8.6) and 31 BPD (26.3 ± 8.2)	DTI data (FA)	SVM	LOOCV	Left ATR, 68.33% ($p = 0.018$) Right SLF, 66.67% ($p = 0.029$)	RD profile (accuracy) Left CC, 65.57% ($p = 0.043$), Right SLF, 68.25% ($p = 0.024$) Right AF, 72.34% ($p = 0.008$)

Table 1. *Cont.*

References (Year)	Subjects (Mean Age)	Features	Machine Learning Method	Cross-Validation	Accuracy *	Comments
Fu et al., 2008 [13]	19 MDD (43.2 ± 8.8) and 19 HC (42.8 ± 6.7)	fMRI data	SVM	LOOCV	86% (sensitivity 84% and specificity 89%)	Lateral temporal cortex, amygdala, and visual processing networks contributed most
Cao et al., 2014 [14]	39 MDD (27.99 ± 7.49) and 37 HC (28.22 ± 6.47)	fMRI data	SVM	LOOCV	84%	Inferior orbitofrontal, supramarginal gyrus, inferior parietal lobule-posterior cingulated gyrus, and middle temporal gyrus-inferior temporal gyrus contributed most
Mourao-Miranda et al., 2011 [15]	19 MDD (43.2 ± 8.8) and 19 HC (42.8 ± 6.7)	fMRI data	SVM	Nested LOOCV	52%	Patients were identified as outliers during facial recognition, with 30% of outliers responding to antidepressants, whereas 89% of non-outliers responded
Zeng et al., 2012 [16]	24 MDD (31.83 ± 10.99) and 29 HC (33.62 ± 10.29)	fMRI data	SVM	LOOCV	94.3%	550 discriminating functional connections; 100% accuracy for patients, 89.7% for controls
Guo et al., 2018 [17]	59 MDD and 31 HC, 29 MDD and 24 HC	fMRI data	SVM	LOOCV	92.22% and 90.57%	Voxel-mirrored homotopic connectivity (VMHC) alterations examined for two separate samples
Wei et al., 2013 [18]	20 MDD (34.3 ± 8.2 and 20 HC (30.8 ± 8.7)	fMRI data	SVM	LOOCV	90% (sensitivity 95% and specificity 85%)	Right fronto-parietal and default mode networks showed deficits, while the left fronto-parietal, ventromedial prefrontal, and salience network were excess networks
He et al., 2021 [19]	40 MDD (40.05 ± 12.32) and 34 HC (34.44 ± 11.76)	fMRI data, peripheral blood	SVM	LOOCV	85.1%	MicroRNA-9, thought to be a neural substrate of childhood maltreatment, integrated into analysis
Ramasubbu et al., 2016 [20]	45 MDD (37 ± 11) and 19 HC (33 ± 10)	fMRI data	SVM	5-fold cross validation	66%	Patients grouped by severity. Mild to moderate (58%) and severe (52%) groups showed lower accuracies
Ramasubbu et al., 2019 [21]	22 MDD (27.36 ± 7.5) and 22 HC (28.09 ± 2.71)	fMRI data	SVM	Nested LOOCV	77.3% (sensitivity 75% and specificity 80%)	Arterial spin labeling MRI was used to measure cerebral blood flow (CBF). Regional CBF of cortical, limbic, and paralimbic regions contributed to classification.

Table 1. *Cont.*

References (Year)	Subjects (Mean Age)	Features	Machine Learning Method	Cross-Validation	Accuracy *	Comments
Yamasita et al., 2020 [22]	149 MDD and 564 HC from four sites, 185 MDD and 264 HC from five sites	fMRI data	LASSO	Nested cross validation	70%	Functional connectivity differences were identified in multisite data, which were applied for classification on another multisite dataset for validation.
Nouretdinov et al., 2011 [23]	19 MDD and 19 HC	fMRI data	TCP	Conformal prediction	89.5% and 92.1% at 90% confidence	Two sad-face recognition tasks used to classify patients using the TCP method; prediction accuracy at least 90% at 90% confidence level
Hahn et al., 2011 [24]	30 patients (MDD, BPD) and 30 HC	fMRI data	GP classification	LOOCV	60%	Sad face, happy face, anxious face, neutral face, anticipation of no reward, anticipation of large reward, anticipation of no loss, and avoiding small loss were significant classifiers
Rosa et al., 2015 [25]	30 patients (MDD, BPD) and 30 HC	fMRI data by Hahn et al. (2011)	Linear L1-norm regularized SVM	Nested cross validation	85%	A novel sparse network based discriminative modeling framework was applied on existing data. Higher accuracies were reached
Shi et al., 2021 [26]	92 MDD, 460 MDD, and 470 HC	fMRI data	Relevance vector regression, eXtreme Gradient Boosting classification	LOOCV, 10-fold cross validation	86.3%	Gray matter density and fractional amplitude of low-frequency fluctuation predicted sleep disturbance in patients. The model was applied to a multicenter dataset for validation.
Guo et al., 2017 [27]	38 MDD (28.4 ± 9.68) and 28 HC (26.6 ± 9.4)	fMRI data	Multikernel SVM		97.54%	A method generating a high order minimum spanning tree functional connectivity network was used to reduce computing consumption and produce a scale conducive to subsequent network analysis
Sato et al., 2015 [28]	25 MDD and 21 HC	fMRI data	Maximum entropy linear discriminant analysis	LOOCV	78.3% (sensitivity 72.0%, specificity 85.7%)	Guilt selective connections used for classification
Han et al., 2019 [29]	25 MDD and 21 schizophrenia	fMRI data	Nonnegative matrix factorization	LOOCV	82.6%	"Triple network" (default mode, salience, central executive) used to distinguish MDD patients from schizophrenia patients

Table 1. *Cont.*

References (Year)	Subjects (Mean Age)	Features	Machine Learning Method	Cross-Validation	Accuracy *	Comments
Yu et al., 2013 [30]	19 MDD (26.65 ± 7.62), 32 schizophrenia (24 ± 5.66), and 38 HC (24.44 ± 4.45)	fMRI data	SVM	LOOCV	80.9% (84.2% for MDD, 81.3% for schizophrenia, 78.9% for HC)	Altered connections in medial prefrontal, anterior cingulate, thalamus, hippocampus, and cerebellum for both patient groups; differences in prefrontal, amygdala, and temporal poles
Grotegerd et al., 2013 [31]	10 MDD (36.8 ± 10.1) and 100 BPD (36.8 ± 8.5)	fMRI data	SVM	LOOCV	90%	Medial prefrontal, orbitofrontal regions contributed to classifying unipolar and bipolar depression
He et al., 2020 [32]	63 MDD (35.35 ± 11.02) and 63 HC (31.78 ± 10.56)	fMRI data	SVR	LOOCV		Left and right amygdala/hippocampus predicted trait sadness; medical prefrontal/anterior cingulate and amygdala/parahippocampal gyrus predicted state anhedonia scores
Maglanoc et al., 2020 [33]	170 MDD (38.7 ± 13.3) and 71 HC (41.8 ± 13.1)	fMRI data	Shrinkage discriminant analysis	10-fold cross validation		Low model performance for classification of depression or anxiety symptoms
Sundermann et al., 2017 [34]	Two subsets of 180 MDD and 180 HC	fMRI data	SVM	LOOCV	56.1%	The subgroup with a higher symptom severity showed a higher classification accuracy (61.7%).

TreeBagging, tree-based bagging; RF, random forest; MLP, multilayer perception; AdaBoost, adaptive boosting; GBM, gradient boosting machine; SVM, support vector machine; SVR, support vector regression; RBF, Gaussian radial basis; ADTree, alternating decision tree; LASSO, least absolute shrinkage and selection operator; TCP, transductive conformal predictor; GP, Gaussian process; LOOCV, leave-one-out cross-validation; FA, fractional anisotropy; RD, radial diffusivity; GM. gray matter; WM. white matter; ATR, anterior thalamic radiation; SLF, superior longitudinal fasciculus; AF, arcuate fasciculus; CC, cingulum cingulate; MADRS, Montgomery–Asberg Depression Rating Scale; HDRS, Hamilton Depression Rating Scale; MDD, major depressive disorder; BPD, bipolar disorder; HC, healthy controls; Accuracy *, highest accuracies presented.

2.1. Structural Characteristics for the Assessment of Depression

2.1.1. Structural Neuroimaging Studies for Diagnosis

Several MRI scan sequences have been used in this study, among which high-resolution T1-weighted imaging confirms gray matter thickness in volume and changes in brain morphology. Previous sMRI studies have suggested variable results of structural changes in a depressed patient group compared with a normal healthy group [35–45]. Thus, sMRI may be the most feasible method for clinical practice. Conventional structural neuroimaging studies have mainly focused on regional volume alterations in gray matter. However, this is insufficient, as morphometric alterations also include changes in shapes and geometric features. Differences in cortical thickness, gray matter volume, and white matter integrity were investigated. Alterations in the cortical thickness have been suggested in several brain regions. Some studies have reported that in patients with MDD, cortical thickness increases in the orbitofrontal cortex [37,38], superior frontal gyrus [36], cingulate cortex [37,39,40], and occipital cortex [40]. Other studies have reported decreased cortical thickness in the orbitofrontal cortex [41,42], insular cortex [43,44], bilateral fusiform gyrus [39], and left occipital area [45]. In a recent meta-analysis of medication-free patients with MDD, Li et al. showed a complex pattern of increased cortical thickness in the posterior cingulate cortex, ventromedial prefrontal cortex, and anterior cingulate cortex and decreased cortical thickness in the gyrus rectus, orbital segment of the superior frontal gyrus, and middle temporal gyrus [46].

Diffusion tensor imaging (DTI) has been used to investigate white matter connectivity and abnormalities in the brain [47]. Using eigenvalues and eigenvectors of water molecules in the brain white matter obtained by MRI, fractional anisotropy (FA), mean diffusivity, axial diffusivity, and radial diffusivity were calculated, and alterations in white matter were confirmed through changes in these values [48]. Previous DTI studies showed relatively consistent results, in that the MDD group had lower FA values than the healthy group in brain regions, including the uncinated fasciculus (UF) [49,50], superior longitudinal fasciculus [51,52], anterior limb of the internal capsule [53,54], corpus callosum (CC) [55–57], and inferior fronto-occipital fasciculus [50,56]. In a recent meta-analysis, adolescents and young adults with MDD showed lower FA values in the CC and frontal-subcortical circuits, which may contribute to the pathogenesis of MDD [58]. Decreased FA values in patients with MDD were associated with the severity of depressive symptoms and duration of illness [53,54,59]. Zhu et al. reported that FA values in the left anterior limb of the internal capsule were negatively correlated with the severity of depressive symptoms [53]. Longer illness duration, the number of previous depressive episodes, and treatment response were related to lower FA values [54,59,60]. The reduction of FA values in treatment-resistant/chronic MMD was significant when compared to that in first-episode MDD and healthy controls [59]. Zheng et al. reported that reduced FA values of the UF in MDD patients returned to normal FA values in healthy controls after 8-week antidepressant treatment [60]. Similarly, reduced white matter connectivity in the anterior cingulum and CC may represent a biomarker of risk for developing MDD [61], and an alteration in white matter microarchitecture has been suggested as a predictor of the treatment outcome in MDD [62,63].

Existing neuroimaging studies have confirmed brain changes in patients with MDD; however, these results have not been applied in current clinical practice. Most of the studies were comparative studies of MDD and healthy control groups, and there was a limitation in investigation of individual-level comparisons. ML has been presented as a method to compensate for these limitations.

2.1.2. ML in Structural Studies for Diagnosing MDD

In this section, we introduce the performance, including accuracy, sensitivity (true positive rate), and specificity (true negative rate), of ML models used in previous sMRI studies for diagnosing MDD. ML studies for diagnosing and predicting the onset of MDD have been steadily increasing, and among them, studies comparing MDD and healthy control groups have been most actively conducted [3]. Foland-Ross et al. reported that baseline cortical thickness predicted the first-onset of MDD with an overall accuracy of 70% in a five-year

follow-up study on adolescent girls aged 10–15 years [5]. Lower baseline thickness of the right medial orbitofrontal cortex and thicker left insula were associated with a higher risk of developing MDD. In a ML diagnostic study, medication-naïve adolescents with first-onset MDD showed increased thickness of the superior segment of the circular sulcus of the insula compared to the healthy control group [6]. In this study, the support vector machine (SVM) method yielded the highest performance with an accuracy of 94.4% (sensitivity, 92.6% and specificity, 96.3%). ML studies using sMRI to diagnose MDD in adults with MDD have also been reported. Qiu et al. found that alteration of the cortical thickness in the right hemisphere could differentiate first-onset MDD patients from healthy controls, providing an accuracy of 78% [7]. They suggested that morphological alterations in the right hemisphere were more evident than those in the left hemisphere in diagnosing MDD. An ML study for diagnosing MDD using DTI data from 29 MDD patients and 30 healthy controls showed an accuracy of 83.05% [8]. In this study, Qin et al. showed that frontoparietal network dysfunction was associated with adult MDD and suggested alterations in this area as a diagnostic measure for MDD. A previous ML study comparing MDD patients and control groups reported that combinations of multimodal imaging and non-imaging measures may help predict late-life depression diagnoses [9]. A learning method called "alternating decision tree" showed the highest accuracy (87.27%) in predicting the diagnosis of late-life depression; poor cognitive ability and whole brain atrophy were found to be associated with late-life depression.

Depression severity was predicted by gray matter volume in patients with bipolar and unipolar depression. Depressive severity was predicted based on the gray matter volume of the bilateral insula, but hypomanic symptom severity was not able to distinguish between unipolar and bipolar depression [10]. In contrast to the previous result that insula volume was smaller in patients with MDD than in healthy controls [43], increased volume was associated with higher symptom severity in mood disorders. These results are likely due to the influence of the bipolar disorder group among participants. In a study comparing bipolar disorder, MDD, and healthy controls, larger volumes of subcortical regions were found in the bipolar disorder group, suggesting potentially varying neuropathological processes in these two conditions [11]. In ML using DTI data, the diagnoses of bipolar disorder and MDD were predicted at the individual level [12]. The FA tract profile of the left anterior thalamic radiation was used to discriminate between bipolar disorder and MDD with an accuracy of 68.33%. These results suggest that the effects of MDD and bipolar disorder on brain structural abnormalities are different, and the accuracy of diagnostic prediction can be improved through a better understanding of the neuropathophysiology.

ML studies for the diagnosis of depression are rapidly increasing, but most studies have a limitation of a small sample size [5–7,64]. In a comparative study of bipolar disorder and MDD, it was difficult to identify the characteristics of bipolar I and II because the bipolar disorder subtypes were not classified in the bipolar disorder patient group. Since this was a cross-sectional study, there is a limit to predicting future changes in bipolar disorder or the onset of comorbid psychiatric disorders in the MDD patient group. Therefore, follow-up studies are required, and when comparing the unipolar depression group with the group that changed from depression to bipolar disorder, a better understanding of the brain structural alterations of the two disorders may be possible.

2.2. Functional Characteristics for the Assessment of Depression

Functional neuroimaging refers to the use of neuroimaging technologies to measure brain function. It is different from structural imaging techniques as it uses various ways to examine the activation and interaction of and between brain regions. Commonly used methods include positron emission tomography, single-photon emission computed tomography, and functional ultrasound imaging; however, the most widely used method is fMRI. Therefore, we mainly focused on recent fMRI studies on depressive disorders that incorporated ML for enhanced results. The use of ML in fMRI studies dates to the late 2000s. Fu et al. applied the SVM method to fMRI data of depressed patients and healthy controls during facial recognition tasks [13]. The authors hypothesized that there would

be stronger contributions from regions that process facial expressions, such as the lateral temporal cortex, amygdala, and visual-processing networks. The ML process resulted in 74% of the patient group and 63% of the control group being correctly classified, yielding an accuracy of 68%. Post-treatment analysis resulted in 75% of partial responders and 62% of full responders being classified correctly. Cao et al. investigated resting-state functional connectivity (rsFC) in 39 MDD patients and 37 matched healthy controls [14]. The altered functional connections were identified and applied to the SVM classification, resulting in an accuracy of 84%. The modules with the highest contribution were the inferior orbitofrontal gyrus, supramarginal gyrus, inferior parietal lobule-posterior cingulated gyrus, and middle temporal gyrus-inferior temporal gyrus.

A study by Mourao-Miranda et al. also investigated the response to sad faces in depressed patients using an SVM [15]. The brain patterns of healthy controls while responding to the stimulus were analyzed, and the patterns of patients with depression were hypothesized to be outliers when compared to controls. Of the patients, 52% were correctly identified as outliers, and 79% of controls were detected as non-outliers. Additional analyses revealed that only 30% of outlier patients responded to antidepressant treatment, whereas 89% of non-outlier patients showed a response.

Zeng et al. analyzed resting-state functional connectivity in MDD patients [16]. Consensus functional connections from previous literature were identified, and many were diminished in the patients. Discriminative power was calculated using the SVM method, and the results showed that the amygdala exhibited the highest discriminative power, showing altered connectivity between the prefrontal lobe, visual cortex, cerebellum, and other limbic areas.

Many non-ML studies have shown that functional connectivity changes are inconsistent in patients with MDD. Guo et al. examined voxel-mirrored homotopic connectivity (VMHC) alterations to obtain more consistent results [17]. Two individual samples of 59 MDD patients and 31 controls and 29 MDD patients and 24 controls were included in the fMRI data acquisition. VMHC was computed using REST software. The overlap of brain clusters showing significant differences between patients and controls was generated using a mask. LIBSVM (A Library for Support Vector Machines) software (http://www.csie.ntu.edu.tw/~cjlin/libsvm/), an integrated software for support vector classification, regression, and distribution estimation, was then used to identify the prediction model. The results showed that the VMHC values in the posterior cingulate cortex and cuneus were able to differentiate MDD patients with an accuracy of 92.22% and 90.57% in each sample, respectively.

Wei et al. concentrated on the "long-term memory" of the temporal dynamics of brain activity [18]. The Hurst exponent has been reported to describe brain activity well in terms of scale-free dynamics. SVM studies involving the Hurst exponent of brain networks of 20 MDD patients and 20 healthy controls revealed a successful discrimination with an accuracy of 90%. The results showed that the right frontoparietal and default mode networks had deficits (lower memory), whereas the left frontoparietal, ventromedial prefrontal, and salience networks were excess networks (longer memory) in patients with MDD.

He et al. examined the role of microRNA-9 in the link between childhood maltreatment and MDD [19]. MicroRNA-9 is thought to be a neural substrate for childhood maltreatment. Forty patients with MDD and 34 healthy controls completed laboratory tests and underwent fMRI, resulting in higher microRNA-9 levels in patients with MDD. SVM models integrating microRNA-9 levels, childhood maltreatment severity, and intrinsic amygdala functional connectivity showed an accuracy of 85.1% in differentiating MDD patients.

Ramasubbu et al. investigated the possible effect of severity on the accuracy of machine-learned classifications [20]. Patients with MDD were divided into groups based on their Hamilton Depression Rating Scale (HRDS) scores, which were classified as mild to moderate, severe, and very severe. fMRI data from 45 patients and 19 controls were collected during the resting state and during an emotional-face matching task. Linear SVM classifiers were used to distinguish patients from controls. The very severe depression group showed an accuracy of 66%, the mild to moderate group showed an accuracy of 58%, and the severe group showed

an accuracy of only 52%. The authors suggested that machine-learned patient classification using fMRI data might be limited to less severe depression.

Ramasubbu also examined patients with MDD using arterial spin labeling (ASL) MRI [21]. Of the 22 MDD patients and 22 healthy controls who underwent pseudo-continuous 3D-ASL imaging to determine regional cerebral blood flow, which was then used in combination with sex and age as SVM classifiers for detecting patients, the resulting classification had an accuracy of 77.3%, with the highest contributing features being sex and the cerebral blood flow in cortical, limbic, and paralimbic regions.

Yamasita et al. acknowledged the difficulty of neuroimaging studies owing to the differences between various study sites and their fMRI products [22]. They used a harmonization method to remove such differences in a dataset of 713 participants from four imaging sites (564 healthy controls and 149 MDD patients). The dataset was then analyzed using an ML algorithm called the least absolute shrinkage and selection operator. It was shown that the functional "under" connectivity (more negative) between the right and left insula was the largest difference between MDD patients and healthy controls. A total of 25 functional connectivities were identified for classifying MDD, of which 19 were more negative and six were more positive than healthy controls. These classifiers were used on another dataset with 521 participants from five imaging sites (264 healthy controls and 185 MDD patients) for validation, resulting in a diagnostic accuracy of 70%.

Nouretdinov et al. applied the transductive conformal predictor (TCP) method to MRI, which generated confidence measures for imaging-based predictions [23]. In fact, this study validated the accuracy of the TCP method compared to more conventional methods such as the SVM. Using sad face recognition as a predictor, the authors found diagnostic and prognostic accuracies comparable to those of the conventional methods. Patients reacted more sensitively to sad faces, and such sensitive individuals tended to respond worse to treatment, in line with the findings of previous studies.

Hahn et al. conducted a study analyzing probable diagnostic biomarkers using Gaussian process classifiers (GPC) [24]. Of the 15 conditions used as classifiers, eight were revealed to be significantly accurate in correctly identifying patients with a median accuracy of 60%: sad face, happy face, anxious face, neutral face, anticipation of no reward, anticipation of large reward, anticipation of no loss, and avoiding small loss. GPC showed a higher accuracy than the conventional SVM method in most cases. The authors also stated that a decision tree algorithm led to an accuracy of 83%, which is an improvement of 11% compared with the best GPC (anticipation of no loss).

Rosa et al. used the same dataset as Fu et al. and Hahn et al. to expand on the research [13,24,25]. The authors used a connectivity-based framework to classify the existing data. These analyses resulted in higher accuracies than the original studies (77% vs. 79% for Fu et al., and 70% vs. 85% for Hahn et al.). It should be noted that the sensitivity and specificity were lower than those in the original reports, thus emphasizing that the new framework might not necessarily be superior.

A multicenter analysis by Shi et al. used a multivariable regression algorithm named relevance vector regression to identify sleep-related MRI indicators in patients with MDD [26]. The analysis of 92 patients with MDD identified 50 MRI features distributed through the subcortical system and frontoparietal and visual networks that showed abnormal metabolism. These findings were validated using a multicenter dataset of 460 patients and 470 controls, indicating that sleep disturbance-related MRI features may be possible biomarkers of MDD.

Guo et al. suggested that traditional methods for processing functional connectivity data are highly limited in interpretation and thus proposed a novel high-order minimum spanning tree network for better analysis [27]. The results showed a classification accuracy of up to 97.54% when comparing MDD patients with healthy controls.

Sato et al. used an ML algorithm to assess not the depression itself but vulnerability to it [28]. Subjects were selected from past depression patients who had remitted at least 1 year previously. Functional connectivity was assessed while the subjects and controls

looked at statements about social and moral values. They were later asked to describe their feelings as guilt, disgust, shame, or anger toward themselves or others. A specific ML algorithm, the Maximum Entropy Linear Discriminant analysis, showed that guilt-related functional connectivity changes in the anterior temporal lobe area discriminated previous depression patients from healthy controls with an accuracy of 78.26%, suggesting it as a possible biomarker of patients' vulnerability to depression.

Han et al. examined the so-called triple network of the brain, consisting of the default mode, salience, and central executive network, to distinguish schizophrenia patients from MDD patients [29]. Twenty-one schizophrenia patients and 25 MDD patients were assessed using sMRI and fMRI, and the data were processed using supervised convex nonnegative matrix factorization. This approach was proposed to extract low-rank network patterns in latent space. The middle cingulate cortex, inferior parietal lobule, and cingulate cortex were the most discriminative between the two disorders in terms of functional connectivity, with an accuracy of 82.6%.

Yu et al. also investigated differences in functional connectivity between MDD and schizophrenia [30]. Thirty-two patients with schizophrenia, 19 patients with MDD, and 38 controls underwent fMRI scans, with the results analyzed using an SVM with intrinsic discriminant analysis. Both groups showed altered connections in the medial prefrontal cortex, anterior cingulate cortex, thalamus, hippocampus, and cerebellum. However, the groups also showed differences in the prefrontal cortex, amygdala, and temporal poles. Patient discrimination achieved an accuracy of 80.9% (84.2% for MDD, 81.3% for schizophrenia, and 78.9% for controls). The connections with the highest discriminative powers were found within the default mode network and cerebellum.

Grotegerd et al. attempted to discriminate between unipolar and bipolar depression using an fMRI pattern classification [31]. Twenty participants (10 bipolar and 10 unipolar) were asked to look at happy, negative, and neutral emotional faces during fMRI scans. The contrasts between negative and happy versus neutral faces were used as classifiers. Both the SVM and GPC algorithms were used for classification. SVM classification showed that the happy versus neutral contrast reached an accuracy of 90%, and the negative versus neutral contrast reached an accuracy of 75% for discriminating unipolar from bipolar depression. GPC classification on the other hand showed both happy versus neutral and negative versus neutral as achieving an accuracy of 70%.

He et al. examined the possibility of predicting the specific characteristics of patients with MDD using fMRI [32]. Sixty-three MDD patients and 63 matched controls underwent rsFC imaging. Their trait characteristics were measured using the Affective Neuroscience Personality Scale (ANPS), and state anhedonia was measured using the Snaith–Hamilton Pleasure Scale. SVM regression was used to predict trait and state characteristics based on changes in rsFC. Abnormal connectivity between the left amygdala/hippocampus and right amygdala/hippocampus predicted sadness scores of the ANPS, while connectivity between the medial prefrontal cortex/anterior cingulate gyrus and amygdala/parahippocampal gyrus predicted a state of anhedonia.

Not all ML approaches yield valuable positive outcomes. Maglanoc et al. implemented a ML approach to assess the relationships between clinical variables and structural and functional brain components [33]. Overall, the models showed low predictive values for depression and anxiety symptoms.

Sundermann et al. suggested that most ML approaches using fMRI results as classifiers were successful only for small samples [34]. The authors selected two subsets of 180 patients with MDD and 180 healthy controls from the BiDirect study. The first subset was analyzed using SVM to identify classifiers for the diagnosis of MDD, and the second subset was used to validate the resulting model. Accuracies ranged from 45.0% to 56.1% for the whole group and from 60.8% to 61.7% for the subgroup with higher depression severity. This resulted in the conclusion that classification models did not translate well in a large realistic population.

3. Considerations for ML in Neuroimaging in Depressive Disorders—Treatment Outcomes (Table 2)

Table 2. Selection of studies investigating machine learning methods for the prediction of treatment outcomes in depression.

References (Year)	Subjects (Mean Age)	Features	Machine Learning Method	Cross-Validation	Accuracy *	Comments
Patel et al., 2015 [9]	11 MDD responders and 13 MDD non-responders	DTI data, structural imaging, functional imaging	Decision tree	LOOCV	89.5%	The optimal ADTree model selected MMSE score, age, whole brain atrophy, and fluid-attenuated inversion recovery. Global WM hyperintensity count for predicting depression diagnosis
Gong et al., 2011 [65]	22 non-refractory MDD (39.17 ± 12.88) and 23 refractory MDD (40.43 ± 12.58)	GM and WM	SVM	LOOCV	69.6% (GM) and 65.22% (WM)	Participants were treated with one of three classes of antidepressants: tricyclic, serotonin–norepinephrine reuptake inhibitor, and selective serotonin reuptake inhibitor
Korgaonkar et al., 2015 [66]	54 remitted MDD and 103 non-remitted MDD	GM volume and DTI data (FA)	Decision tree	Hold-out	85.0% (GM volume) and 84.0% (FA)	Participants were randomized to receive flexibly-dosed escitalopram, sertraline, or venlafaxine-ER for 8 weeks
Johnston et al., 2015 [67]	20 treatment-refractory MDD (51.80 ± 11.23) and 21 HC (46.14 ± 13.97)	T1-weighted brain imaging (GM)	SVM	LOOCV	85% (sensitivity, 85% and specificity, 86%)	MDD participants had experienced lifetime and/or current chronic episodes of depression, not necessarily meeting criteria for MDD at time of scanning
Bartlett et al., 2018 [68]	63 remitters (34.59 ± 12.23) and 121 non-remitters (38.40 ± 13.69)	T1-weighted brain imaging (cortical thickness)	RF, PLR	10 repetitions of 5-fold cross-validation	63.9% (sensitivity, 22.6% and specificity, 85.8%)	Patients with early onset MDD (before age 30) and chronic (episode duration >2 years) or recurrent MDD (≥2 recurrences) were enrolled. Remission status was predicted more accurately with RF than PLR
Redlich et al., 2016 [69]	23 ECT-treated MDD (45.7 ± 9.8), with 13 responders and 10 non-responders	High-resolution T1-weighted structural imaging (GM volume)	SVM, SVR	LOOCV	78.3% (sensitivity, 100% and specificity, 50%)	Brief-pulse ECT was conducted three times a week with antidepressants (mean number of sessions, 14)
Cao et al., 2018 [70]	24 severe MDD (31.3 ± 10.8), with 12 remitters and 12 non-remitters	T1-weighted structural imaging (GM volume)	SVR	LOOCV	Overall, 83.3% (sensitivity, 91.7% and specificity, 75%)	All the patients were under severe unipolar depression and received eight sessions of modified ECT

Table 2. *Cont.*

References (Year)	Subjects (Mean Age)	Features	Machine Learning Method	Cross-Validation	Accuracy *	Comments
Gaertner et al., 2021 [71]	39 responders (50.23 ± 17.53) and 32 non-responders (51.31 ± 18.09)	Structural MRI	SVM with a linear kernel	LOOCV	69% (sensitivity, 67% and specificity, 72%)	Schizoaffective disorder (4%) and BD (13%) were included. Twelve sessions of ECT were administered, and patients with partial response had extra ECT-sessions (mean no. sessions: 13.61 ± 4.34)
Takamiya et al., 2020 [72]	20 remitters and seven non-remitters	High-resolution T1-weighted structural imaging (GM volume) and clinical variables	SVM, SVR	LOOCV	90% (sensitivity, 100% and specificity, 71%)	Clinical variables included age, sex, diagnosis, psychotic features, family history of mood disorder, duration of episode, illness duration, previous ECT, and the score of each item of HDRS-17
Tymofiyeva et al., 2019 [73]	30 MDD (16.0 ± 1.3)	DTI data	Decision tree (J48)	10-fold cross validation	83% (sensitivity, 82% and specificity, 84%)	All patients underwent CBT, and six patients received antidepressants with CBT; 19 improvers and 11 non-improvers were included
Marquand et al., 2008 [74]	20 MDD (43.7 ± 8.6) and 20 HC (43.7 ± 8.3)	fMRI data	SVM	LOOCV		Statistical significance for response prediction not achieved
Frassle et al., 2020 [75]	85 MDD	fMRI data	SVM	LOOCV	79% (chronic vs. fast remission), 61% (gradual improvement vs. fast remission)	Data from the Netherlands Study of Depression and Anxiety were used to classify chronic patients, gradual improvement, and fast remission
Tian et al., 2020 [76]	106 MDD and 109 HC	fMRI data	SVM	LOOCV	79.4%	Multicenter data analyzed while assuming an HDRS score reduction of at least 50% as response after escitalopram monotherapy
Liu et al., 2020 [77]	57 MDD (31 amisulpride, 26 placebo) and 28 HC	fMRI data	Elastic net regularization	Nested cross validation	77% (MDD vs. HC), 59% (amisulpride vs. placebo)	Striatal network functional connectivity changes were most predictive for classification, suggesting a dopaminergic role in treatment outcome
Osuch et al., 2018 [78]	34 MDD (19.7 ± 2.6), 32 BPD (21.3 ± 2.9), and 33 HC (20.2 ± 2.0)	fMRI data	SVM	Nested cross validation	92.4% (MDD vs. BPD), 92% (medication class response prediction)	Diagnostic classification also succeeded in predicting the optimal medication class of response, where BPD patients responded to mood stabilizers, and MDD patients responded better to antidepressants.

Table 2. *Cont.*

References (Year)	Subjects (Mean Age)	Features	Machine Learning Method	Cross-Validation	Accuracy *	Comments
Hopman et al., 2021 [79]	70 MDD (41.93 ± 11.67)	fMRI data	SVM	5-fold cross validation	95.35%	Medication resistant patients were treated with rTMS and analyzed to predict short term and long-term treatment response. Sustained response was associated with stronger anterior cingulate/occipital cortex connectivity
Cash et al., 2019 [80]	47 MDD (43 ± 12) and 29 HC (39 ± 15)	fMRI data	SVM	LOOCV	85~95%	Reduced connectivity in default mode and affective network was associated with better rTMS response
Wang et al., 2018 [81]	23 MDD (38.74 ± 11.02) and 25 HC (39.52 ± 8.07)	fMRI data	SVM	LOOCV	72.92%	Local functional connectivity density of left pre/postcentral gyri, both superior temporal gyri were predictive of ECT treatment response
Pei et al., 2020 [82]	98 MDD	fMRI data, venous blood	SVM	LOOCV	86%	fMRI data were combined with genetic data on selected single nucleotide polymorphisms for classification of responders and non-responders to medication, resulting in higher accuracy than fMRI data alone (61%)

RF, random forest; SVM, support vector machine; SVR, support vector regression; PLR, penalized logistic regression; ADTree, alternating decision tree; LOOCV, leave-one-out cross-validation; FA, fractional anisotropy; GM. gray matter; WM. white matter; MADRS, Montgomery–Asberg Depression Rating Scale; ECT, electroconvulsive therapy; HDRS, Hamilton Depression Rating Scale; rTMS, repeated transcranial magnetic stimulation; MDD, major depressive disorder; BPD, bipolar disorder; HC, healthy controls. Accuracy *, highest accuracies presented.

3.1. Structural Characteristics Related with Depression Treatment Outcomes

The treatment for MDD is determined according to clinical symptoms, and treatments include pharmacotherapy, psychotherapy, and electroconvulsive therapy (ECT). Antidepressants are used as the first-line treatment for depression, and fewer than 50% of patients do not achieve remission [83]. Approximately two-thirds of all patients respond to pharmacotherapy and/or psychotherapy [84], but the remaining one-third are resistant to treatment. The prolonged duration of unremitted MDD increases an individual's functional loss and overall mental healthcare burden. Therefore, it is very important to predict a patient's response to particular treatments and to design treatment strategies early at the onset of MDD. Studies have been conducted to identify biomarkers that can predict treatment response, and ML studies are also increasing [9,65–72,85,86].

Gong et al. distinguished between patients with refractory depression and those with non-refractory depression through ML using sMRI data [65]. In this study, the refractory group was defined as MDD patients with a poor response whose Hamilton Depression Rating Scale (HDRS) score did not decrease by more than 50% even after 6 weeks of treatment with two different classes of antidepressants. SVM was applied, and gray matter distinguished between the refractory and non-refractory groups with an accuracy of 69.57%, and white matter distinguished between them with an accuracy of 65.22%. Compared to pre-treatment white matter images, gray matter images showed higher accuracy in predicting the response to antidepressants in patients with MDD. Korgaonkar et al. explored both gray matter volume and FA in 157 patients with MDD, including 103 non-remitters and 54 remitters [66]. Patients received treatment with antidepressants, including escitalopram, sertraline, and venlafaxine, for 8 weeks, and approximately 35% of all participants achieved remission. Using an ML method (decision tree), this study revealed that gray matter volume (smaller left middle frontal gyrus and greater right angular gyrus) and structural connectivity (lower FA values of the left cingulum bundle, right superior fronto-occipital fasciculus, and right superior longitudinal fasciculus) predicted nonremission. It suggested that pre-treatment MRI measures could predict MDD patients who did not respond to antidepressant treatment. Similarly, high accuracy has been reported in ML for predicting treatment response in late-life depression. Patel reported that the optical ADTree model, including measures of structural and functional connectivity, showed an accuracy of 89.47% in a study of 24 patients with depressive disorders (11 responders and 13 non-responders) [9]. A study comparing patients with treatment-refractory depression and healthy controls reported that the patient group and the healthy control group could be discriminated using sMRI, even if MDD patients did not meet the criteria for depressive episodes at the time of MRI scanning [67]. Johnston et al. reported that gray matter reductions in the caudate, insula, and periventricular gray matter supported individual prediction with an accuracy of 85%. Similar to the results of previous sMRI studies [85,86], they suggested an association between reduced volume of the insula and slower recovery/poor prognosis of MDD in ML using sMRI. The result that early treatment cortical thickness (one week into treatment) was more associated with the selective serotonin reuptake inhibitor (SSRI) treatment response than pre-treatment cortical thickness was presented in the Clinical Trial Establishing Moderators and Biosignatures of Antidepressant Response in Clinical Care [68,87]. Bartlett et al. used two methods of random forest (RF) and penalized logistic regression for predicting SSRI treatment response, and psychometric data, demographic data, pre-treatment cortical thickness/volume, and one-week treatment change in cortical thickness/volume were included [68]. RF predicted the remission status more accurately with an accuracy of 63.9%, and they found that frontal lobe structural alterations in the first week of treatment may be associated with long-term treatment efficacy.

Several ML studies have been reported to predict the response to ECT in MDD [69–71]. In a study conducted by Redlich et al., 23 MDD patients received ECT, and they comprised 13 responders and 10 non-responders based on the reduction in their HDRS score (50%) [69]. Structural images obtained before treatment predicted the treatment response with an ac-

curacy of 78.3% (100% sensitivity, 13 of 13 responders). The results of support vector regression (SVR) showed a positive association between predicted and true individual percentages of change in the HDRS score. This study suggests that a higher pre-treatment subgenual cingulate gyrus gray matter volume is associated with a better clinical response. A previous Chinese study using linear kernel SVR also reported that pre-treatment hippocampal subfield volumes predicted whether a patient could achieve remission after ECT and the degree of alleviation of depressive symptoms through the use of ECT [70]. They found that MDD patients with baseline smaller hippocampal subfields had better outcomes, and baseline hippocampal subfield volumes were used to predict the change in depressive symptoms with an overall accuracy of 83.3%. A study was conducted to predict ECT treatment response in a group of patients with depressive disorders and other psychiatric disorders [71,72]. Gärtner et al. predicted the treatment response (percentage of depressive symptom reduction) after ECT in a retrospective study including patients with depressive disorder, bipolar disorder, and schizoaffective disorder [71]. The results showed that the ML method discriminated between responders and non-responders with an accuracy of 69%; gray matter volume in the right parahippocampal gyrus provided the most informative contribution. In a Japanese study, 25 variables along with sMRI data were used as candidate features to predict remission and reduction of depressive symptoms [72]. Compared to the model using only clinical variables, the model including sMRI data showed higher predictability accuracy (70.4% and 92.6%, respectively), and the volumes of the regions including the gyrus rectus, right anterior lateral temporal lobe, cuneus, and third ventricle predicted ECT treatment response. The model including both clinical variables and sMRI data showed the same predictive value as the model using only sMRI data. Previous studies have suggested that pre-treatment sMRI is predictive of ECT treatment, although there are limitations in these studies in that they have a relatively small sample size and include a heterogenous patient group [71,72].

A study to predict the improvement of depressive symptoms in adolescents receiving non-pharmacological treatment was also conducted in the United States. Tymofiyeva et al. predicted the treatment response of three months of cognitive behavioral therapy (CBT) using MRI-based structural connectome data [73]. They predicted improvement of depressive symptoms with an accuracy of 83% using J48 classification and right thalamus, left middle frontal gyrus, and baseline depression severity, which were associated with the prediction. Although this study had limitations in that the sample size was small and CBT treatment protocols were heterogeneous, it suggests the possibility of predicting the effect of CBT through brain imaging findings.

Studies on the search for objective indicators to predict pharmacological and non-pharmacological treatments of depression are increasing, and several ML methods are being used. Current ML studies have suggested the possibility of predicting treatment response through pre-treatment sMRI, but there are some limitations. Most previous studies included relatively small sample sizes and heterogeneous patient groups [71,72]. In addition, classes and dosages of antidepressants have not been strictly determined. Since the pharmacological profile and medication dosage of antidepressants affect brain structure, the effects of these variables cannot be excluded [88]. Furthermore, there is a difference in the prediction accuracy (63–93%) according to the design of the study, and the characteristics of participants and variables associated with predicting the treatment response are inconsistent. The symptoms of depressive disorders are heterogeneous, and the causes of onset are diverse. Multiple types of data, rather than a single type of data, may be helpful in increasing prediction accuracy [89]. To reduce the duration of the untreated period of MDD, high prediction accuracy of the response to each treatment method is essential to plan the treatment strategy. For example, ECT is an effective and well-established treatment for refractory depression and is not generally considered as a first-line treatment in clinical practice for various reasons, such as the potential adverse effect, stigma associated with the treatment, and uncertainty of the treatment mechanism [90]. A more

accurate prediction of treatment response may help psychiatrists in clinical decision-making regarding first-line treatment for the management of MDD.

3.2. Functional Characteristics Related with Depression Treatment Outcomes

Functional imaging is usually used to identify possible diagnostic markers and refine diagnostic accuracy in patients with depressive disorder, but this is not the only application for this technological advancement. The change in functional connectivity can serve as an indicator for evaluating treatment outcomes and perhaps even the fit between a patient and a certain treatment regimen.

Marquand et al. analyzed task-related fMRI data using the SVM method to examine verbal working memory as a possible biomarker for patients with depression [74]. The brain activity of correlated areas was closer to statistical significance as task difficulty increased, but actual significance was not achieved. Analysis of the treatment response revealed that the most difficult tasks were significantly accurate in predicting the response to 8 weeks of fluoxetine treatment.

A European research team used an ML strategy called generative embedding, which combines models with classifiers, to predict treatment outcomes in patients with MDD at the single-patient level [75]. Neuroimaging data acquired from the Netherlands Study of Depression and Anxiety were used for supervised learning [91]. The team predicted a given patient's recovery to be fast or chronic with an accuracy of 79% and fast or gradual with an accuracy of 61%.

Tian et al. compared the rsFC of 106 patients with MDD and 109 controls to predict treatment outcomes of the antidepressant escitalopram [76]. A linear soft-threshold SVM model discriminated responders from non-responders using a reduction of at least 50% in the HDRS as reference. The anterior cingulate cortex seemed to be the hub for connections for the various interconnections that discriminated responders from non-responders, predicting treatment response with an accuracy of 79.41%.

Liu et al. used an ML technique for model selection in a whole-brain analysis to differentiate MDD patients from healthy controls and further distinguish MDD patients taking amisulpride from those taking placebo to assess the therapeutic effect of dopaminergic enhancement in MDD [77]. The results indicated that the activation and connectivity of reward-related striatal networks were the most predictive, suggesting a possible route by which dopaminergic agents affect treatment outcomes in patients with MDD.

Osuch et al. conducted practical research on the prediction of a medication-class response in patients with mood disorders [78]. A total of 99 subjects (32 with bipolar I disorder, 34 with MDD, and 33 healthy controls) underwent resting-state fMRI, which was used to train a predictive algorithm and construct SVM classifiers. The bipolar disorder group was hypothesized to respond better to mood stabilizers, whereas the MDD group was thought to respond better to antidepressants. This classification resulted in an accuracy of 92.4% in the known-diagnosis group. This method was applied to 12 patients and all had complicated diagnoses. The suggested optimal medication class led to recovery in 11 of 12 cases (approximately 92%).

Hopman et al. predicted the functional connectivity between the left dorsolateral prefrontal cortex (DLPFC) and subgenual cingulate cortex (sgACC) to serve as a biomarker for a repetitive transcranial magnetic stimulation treatment response [79]. Supervised ML analyses of fMRI data from 70 patients with MDD revealed that this was not true. Instead, non-responders showed poor connections between the sgACC and other locations (frontal pole, superior parietal lobule, and occipital cortex) and between the DLPFC and central opercular cortex. These new observations predicted rTMS treatment results with an accuracy of 95.35%.

Cash et al. also investigated possible neuroimaging biomarkers for rTMS treatment outcomes [80]. Data of 47 patients and 29 controls were analyzed using fMRI, resulting in lower activation in the caudate, prefrontal cortex, and thalamic areas in the patient group. Reduced functional connectivity in the default mode and affective networks in patients

was also associated with a better treatment response. These features were used to train SVMs, resulting in an rTMS treatment outcome prediction with 85–95% accuracy.

Wang et al. sought to identify biomarkers for the treatment response to ECT, another non-pharmacological treatment option for MDD [81]. They focused on the functional connectivity density (FCD) and rsFC in 23 patients before and after ECT. Neuroimaging data analyses showed that local FCD but not long-range FCD of the left pre-and postcentral gyri and both superior temporal gyri were predictive of changes in HDRS scores after treatment. The SVM-based classification resulted in a prediction accuracy of 72.92%.

Pei et al. combined neuroimaging data with genetic data for more precise modeling of predicting outcomes [82]. The participants were divided into treatment responders and non-responders based on the HDRS score changes after 2 weeks of treatment. Functional connectivity between 14 selected regions of interest and genomic data on selected single nucleotide polymorphisms were acquired. Using SVM with a combination of both datasets resulted in a higher prediction accuracy than when using only one dataset (61% to 86%).

Patel et al. researched patients with late-life depression to find an alternative learning method to the traditional SVM for predicting the diagnosis of and treatment response to depression [9]. They combined various clinical variables with structural and functional neuroimaging data using alternating decision trees. The model showed a diagnostic accuracy of 87.27% using age, mini-mental state examination scores, and structural imaging data as variables, and it showed a treatment response accuracy of 89.47% using structural and functional connectivity data as variables. The best functional connectivity predictors were lower resting-state connections within the dorsal default mode network.

There are many studies trying to find biomarkers for treatment response in MDD patients. In the field of neuroimaging, functional connectivity is a possible candidate since it has resulted in prediction rates with accuracies over 90%, depending on the treatment regimen and the included clinical variables. As an MRI apparatus is for diagnostic purposes, its high costs and low accessibility remain a challenge for its everyday use as a biomarker, but these insights will aid future methods that are more affordable and available.

4. Further Considerations in ML for Depressive Disorder

These models may seem close to being clinically applicable. However, it is unknown whether the models can maintain these accuracies when applied to brain images acquired using different scanners and in different populations. In addition, we summarize the issues in the future application of ML to depressive disorders as follows.

4.1. Sample Sizes

When training our model, it was impossible to use information about the entire population. Instead, we could use only a small finite sample. Larger sample sizes are required to use these algorithms. It is quite difficult to acquire sufficient sample data from neuroimaging studies. Small sample sizes and the complexity of the model result in overfitting. In addition, simply increasing the amount of data can worsen overfitting if the number of dimensions also increases. These problems limit the generalizability of the model to clinical settings. Both dimensionality reduction and an increased sample size are required.

4.2. Type of Data including Imaging Modality and Selection of Features from Those Data

No single feature consistently predicted the diagnosis or treatment response across different studies. This reflects the heterogeneity of depression. In addition, neuroimaging data by themselves have limitations in the information they contain. Many features have been hypothesized to be useful for predicting results regarding the diagnosis and treatment outcome. These include sociodemographic, clinical, psychological, neuroimaging, genetic, immune, and endocrine data. It is necessary to include other depression-related clinical variables, such as the clinical characteristics of depression, sex, number of episodes, and multimodal data, as variables for prediction. This further increases the prediction accuracy.

As the predictive power of these variables is quite different across different clinical populations, this could limit the generalizability of the study results. In addition, as mentioned for sample sizes, simply increasing the number of features leads to an increase in dimensionality. This also leads to overfitting with respect to the sample size. Researchers must decide how to combine prediction models from different dimensions to achieve accuracy [92].

4.3. Training Algorithms and Types of Validation

Many different algorithms have been used, although the most commonly used algorithms are SVM algorithms. Most studies have used supervised prediction and classification algorithms capable of modeling linear and nonlinear relationships to construct predictive models. As there is no one way to reliably integrate all the variables with different modalities into one model, and certain algorithms are more suitable for different combinations of features [93], the proper choice of algorithms and further improvement of algorithms are needed.

4.4. Clinical Applicability from Results

The heterogeneity of the samples in studies using ML in relation to clinical characteristics and medication status may limit the generalizability of the results [94]. For example, it is common to use medication-naïve samples in MDD neuroimaging studies. As the effects of medication on neuroimaging findings need to be controlled, many researchers have attempted to compare medication-naïve patients and control groups [7]. MDD is a chronic disease in clinical practice, and many patients suffer from chronic impairments caused by MDD itself. They are also influenced by the medications used to treat MDD. The study results from artificially selected medication-naïve patient groups for comparability issues with healthy controls might be limited in their generalizability and clinical use in real practice, involving a high proportion of chronically depressed patients. The high cost and low availability of neuroimaging facilities in the general population is also another limitation for this field.

Although there are many challenges, it is thought that these ML techniques will eventually integrate the various data to enable individual-level clinical inferences that are applicable to actual clinical practice. This is also expected to be related to personalized precision medicine in the future.

References

1. Shi, Y.; Zhang, L.; Wang, Z.; Lu, X.; Wang, T.; Zhou, D.; Zhang, Z. Multivariate machine learning analyses in identification of major depressive disorder using resting-state functional connectivity: A multicentral study. *ACS Chem. Neurosci.* **2021**, *12*, 2878–2886. [CrossRef] [PubMed]
2. Orrù, G.; Pettersson-Yeo, W.; Marquand, A.F.; Sartori, G.; Mechelli, A. Using support vector machine to identify imaging biomarkers of neurological and psychiatric disease: A critical review. *Neurosci. Biobehav. Rev.* **2012**, *36*, 1140–1152. [CrossRef] [PubMed]
3. Gao, S.; Calhoun, V.D.; Sui, J. Machine learning in major depression: From classification to treatment outcome prediction. *CNS Neurosci. Ther.* **2018**, *24*, 1037–1052. [CrossRef] [PubMed]

4. Squarcina, L.; Villa, F.M.; Nobile, M.; Grisan, E.; Brambilla, P. Deep learning for the prediction of treatment response in depression. *J. Affect. Disord.* **2021**, *281*, 618–622. [CrossRef]

5. Foland-Ross, L.C.; Sacchet, M.D.; Prasad, G.; Gilbert, B.; Thompson, P.M.; Gotlib, I.H. Cortical thickness predicts the first onset of major depression in adolescence. *Int. J. Dev. Neurosci.* **2015**, *46*, 125–131. [CrossRef]

6. Kim, D.H.; Kang, P.S.; Kim, J.H.; Kim, C.Y.; Lee, J.H.; Suh, S.I.; Lee, M.S. Machine learning classification of first-onset drug-naïve MDD using structural MRI. *IEEE Access* **2019**, *7*, 153977–153985. [CrossRef]

7. Qiu, L.; Huang, X.; Zhang, J.; Wang, Y.; Kuang, W.; Li, J.; Wang, X.; Wang, L.; Yang, X.; Lui, S.; et al. Characterization of major depressive disorder using a multiparametric classification approach based on high resolution structural images. *J. Psychiatry Neurosci.* **2014**, *39*, 78–86. [CrossRef]

8. Qin, J.; Wei, M.; Liu, H.; Chen, J.; Yan, R.; Hua, L.; Zhao, K.; Yao, Z.; Lu, Q. Abnormal hubs of white matter networks in the frontal-parieto circuit contribute to depression discrimination via pattern classification. *Magn. Reson. Imaging* **2014**, *32*, 1314–1320. [CrossRef]

9. Patel, M.J.; Andreescu, C.; Price, J.C.; Edelman, K.L.; Reynolds, C.F., 3rd; Aizenstein, H.J. Machine learning approaches for integrating clinical and imaging features in late-life depression classification and response prediction. *Int. J. Geriatr. Psychiatry* **2015**, *30*, 1056–1067. [CrossRef]

10. Wise, T.; Marwood, L.; Perkins, A.M.; Herane-Vives, A.; Williams, S.C.R.; Young, A.H.; Cleare, A.J.; Arnone, D. A morphometric signature of depressive symptoms in unmedicated patients with mood disorders. *Acta Psychiatr. Scand.* **2018**, *138*, 73–82. [CrossRef]

11. Fung, G.; Deng, Y.; Zhao, Q.; Li, Z.; Qu, M.; Li, K.; Zeng, Y.W.; Jin, Z.; Ma, Y.T.; Yu, X.; et al. Distinguishing bipolar and major depressive disorders by brain structural morphometry: A pilot study. *BMC Psychiatry* **2015**, *15*, 298. [CrossRef]

12. Deng, F.; Wang, Y.; Huang, H.; Niu, M.; Zhong, S.; Zhao, L.; Qi, Z.; Wu, X.; Sun, Y.; Niu, C.; et al. Abnormal segments of right uncinate fasciculus and left anterior thalamic radiation in major and bipolar depression. *Prog. Neuro-Psychopharmacol. Biol. Psychiatry* **2018**, *81*, 340–349. [CrossRef]

13. Fu, C.H.; Mourao-Miranda, J.; Costafreda, S.G.; Khanna, A.; Marquand, A.F.; Williams, S.C.; Brammer, M.J. Pattern classification of sad facial processing: Toward the development of neurobiological markers in depression. *Biol. Psychiatry* **2008**, *63*, 656–662. [CrossRef]

14. Cao, L.; Guo, S.; Xue, Z.; Hu, Y.; Liu, H.; Mwansisya, T.E.; Pu, W.; Yang, B.; Liu, C.; Feng, J.; et al. Aberrant functional connectivity for diagnosis of major depressive disorder: A discriminant analysis. *Psychiatry Clin. Neurosci.* **2014**, *68*, 110–119. [CrossRef]

15. Mourão-Miranda, J.; Hardoon, D.R.; Hahn, T.; Marquand, A.F.; Williams, S.C.; Shawe-Taylor, J.; Brammer, M. Patient classification as an outlier detection problem: An application of the one-class support vector machine. *Neuroimage* **2011**, *58*, 793–804. [CrossRef]

16. Zeng, L.L.; Shen, H.; Liu, L.; Wang, L.; Li, B.; Fang, P.; Zhou, Z.; Li, Y.; Hu, D. Identifying major depression using whole-brain functional connectivity: A multivariate pattern analysis. *Brain* **2012**, *135*, 1498–1507. [CrossRef]

17. Guo, W.; Cui, X.; Liu, F.; Chen, J.; Xie, G.; Wu, R.; Zhang, Z.; Chen, H.; Zhang, X.; Zhao, J. Decreased interhemispheric coordination in the posterior default-mode network and visual regions as trait alterations in first-episode, drug-naive major depressive disorder. *Brain Imaging Behav.* **2018**, *12*, 1251–1258. [CrossRef]

18. Wei, M.; Qin, J.; Yan, R.; Li, H.; Yao, Z.; Lu, Q. Identifying major depressive disorder using Hurst exponent of resting-state brain networks. *Psychiatry Res.* **2013**, *214*, 306–312. [CrossRef]

19. He, C.; Bai, Y.; Wang, Z.; Fan, D.; Wang, Q.; Liu, X.; Zhang, H.; Zhang, H.; Zhang, Z.; Yao, H.; et al. Identification of microRNA-9 linking the effects of childhood maltreatment on depression using amygdala connectivity. *Neuroimage* **2021**, *224*, 117428. [CrossRef]

20. Ramasubbu, R.; Brown, M.R.; Cortese, F.; Gaxiola, I.; Goodyear, B.; Greenshaw, A.J.; Dursun, S.M.; Greiner, R. Accuracy of automated classification of major depressive disorder as a function of symptom severity. *Neuroimage Clin.* **2016**, *12*, 320–331. [CrossRef]

21. Ramasubbu, R.; Brown, E.C.; Marcil, L.D.; Talai, A.S.; Forkert, N.D. Automatic classification of major depression disorder using arterial spin labeling MRI perfusion measurements. *Psychiatry Clin. Neurosci.* **2019**, *73*, 486–493. [CrossRef]

22. Yamashita, A.; Sakai, Y.; Yamada, T.; Yahata, N.; Kunimatsu, A.; Okada, N.; Itahashi, T.; Hashimoto, R.; Mizuta, H.; Ichikawa, N.; et al. Generalizable brain network markers of major depressive disorder across multiple imaging sites. *PLoS Biol.* **2020**, *18*, e3000966. [CrossRef]

23. Nouretdinov, I.; Costafreda, S.G.; Gammerman, A.; Chervonenkis, A.; Vovk, V.; Vapnik, V.; Fu, C.H. Machine learning classification with confidence: Application of transductive conformal predictors to MRI-based diagnostic and prognostic markers in depression. *Neuroimage* **2011**, *56*, 809–813. [CrossRef]

24. Hahn, T.; Marquand, A.F.; Ehlis, A.C.; Dresler, T.; Kittel-Schneider, S.; Jarczok, T.A.; Lesch, K.P.; Jakob, P.M.; Mourao-Miranda, J.; Brammer, M.J.; et al. Integrating neurobiological markers of depression. *Arch. Gen. Psychiatry* **2011**, *68*, 361–368. [CrossRef]

25. Rosa, M.J.; Portugal, L.; Hahn, T.; Fallgatter, A.J.; Garrido, M.I.; Shawe-Taylor, J.; Mourao-Miranda, J. Sparse network-based models for patient classification using fMRI. *Neuroimage* **2015**, *105*, 493–506. [CrossRef]

26. Shi, Y.; Zhang, L.; He, C.; Yin, Y.; Song, R.; Chen, S.; Fan, D.; Zhou, D.; Yuan, Y.; Xie, C.; et al. Sleep disturbance-related neuroimaging features as potential biomarkers for the diagnosis of major depressive disorder: A multicenter study based on machine learning. *J. Affect. Disord.* **2021**, *295*, 148–155. [CrossRef]

27. Guo, H.; Qin, M.; Chen, J.; Xu, Y.; Xiang, J. Machine-learning classifier for patients with major depressive disorder: Multifeature approach based on a high-order minimum spanning tree functional brain network. *Comput. Math. Methods Med.* **2017**, *2017*, 4820935. [CrossRef]

28. Sato, J.R.; Moll, J.; Green, S.; Deakin, J.F.; Thomaz, C.E.; Zahn, R. Machine learning algorithm accurately detects fMRI signature of vulnerability to major depression. *Psychiatry Res.* **2015**, *233*, 289–291. [CrossRef]

29. Han, W.; Sorg, C.; Zheng, C.; Yang, Q.; Zhang, X.; Ternblom, A.; Mawuli, C.B.; Gao, L.; Luo, C.; Yao, D.; et al. Low-rank network signatures in the triple network separate schizophrenia and major depressive disorder. *Neuroimage Clin.* **2019**, *22*, 101725. [CrossRef]

30. Yu, Y.; Shen, H.; Zeng, L.L.; Ma, Q.; Hu, D. Convergent and divergent functional connectivity patterns in schizophrenia and depression. *PLoS ONE* **2013**, *8*, e68250. [CrossRef]

31. Grotegerd, D.; Suslow, T.; Bauer, J.; Ohrmann, P.; Arolt, V.; Stuhrmann, A.; Heindel, W.; Kugel, H.; Dannlowski, U. Discriminating unipolar and bipolar depression by means of fMRI and pattern classification: A pilot study. *Eur. Arch. Psychiatry Clin. Neurosci.* **2013**, *263*, 119–131. [CrossRef] [PubMed]

32. He, Z.; Lu, F.; Sheng, W.; Han, S.; Pang, Y.; Chen, Y.; Tang, Q.; Yang, Y.; Luo, W.; Yu, Y.; et al. Abnormal functional connectivity as neural biological substrate of trait and state characteristics in major depressive disorder. *Prog. Neuro-Psychopharmacol. Biol. Psychiatry* **2020**, *102*, 109949. [CrossRef] [PubMed]

33. Maglanoc, L.A.; Kaufmann, T.; Jonassen, R.; Hilland, E.; Beck, D.; Landrø, N.I.; Westlye, L.T. Multimodal fusion of structural and functional brain imaging in depression using linked independent component analysis. *Hum. Brain Mapp.* **2020**, *41*, 241–255. [CrossRef] [PubMed]

34. Sundermann, B.; Feder, S.; Wersching, H.; Teuber, A.; Schwindt, W.; Kugel, H.; Heindel, W.; Arolt, V.; Berger, K.; Pfleiderer, B. Diagnostic classification of unipolar depression based on resting-state functional connectivity MRI: Effects of generalization to a diverse sample. *J. Neural Transm.* **2017**, *124*, 589–605. [CrossRef]

35. Schmaal, L.; Hibar, D.P.; Sämann, P.G.; Hall, G.B.; Baune, B.T.; Jahanshad, N.; Cheung, J.W.; van Erp, T.G.M.; Bos, D.; Ikram, M.A.; et al. Cortical abnormalities in adults and adolescents with major depression based on brain scans from 20 cohorts worldwide in the ENIGMA major depressive disorder working group. *Mol. Psychiatry* **2017**, *22*, 900–909. [CrossRef]

36. Grieve, S.M.; Korgaonkar, M.S.; Koslow, S.H.; Gordon, E.; Williams, L.M. Widespread reductions in gray matter volume in depression. *Neuroimage Clin.* **2013**, *3*, 332–339. [CrossRef]

37. Qiu, L.; Lui, S.; Kuang, W.; Huang, X.; Li, J.; Li, J.; Zhang, J.; Chen, H.; Sweeney, J.A.; Gong, Q. Regional increases of cortical thickness in untreated, first-episode major depressive disorder. *Transl. Psychiatry* **2014**, *4*, e378. [CrossRef]

38. Tu, P.C.; Chen, L.F.; Hsieh, J.C.; Bai, Y.M.; Li, C.T.; Su, T.P. Regional cortical thinning in patients with major depressive disorder: A surface-based morphometry study. *Psychiatry Res.* **2012**, *202*, 206–213. [CrossRef]

39. Lee, J.S.; Kang, W.; Kang, Y.; Kim, A.; Han, K.M.; Tae, W.S.; Ham, B.J. Alterations in the occipital cortex of drug-naïve adults with major depressive disorder: A surface-based analysis of surface area and cortical thickness. *Psychiatry Investig.* **2021**, *18*, 1025–1033. [CrossRef]

40. Peng, D.; Shi, F.; Li, G.; Fralick, D.; Shen, T.; Qiu, M.; Liu, J.; Jiang, K.; Shen, D.; Fang, Y. Surface vulnerability of cerebral cortex to major depressive disorder. *PLoS ONE* **2015**, *10*, e0120704. [CrossRef]

41. Na, K.S.; Won, E.; Kang, J.; Chang, H.S.; Yoon, H.K.; Tae, W.S.; Kim, Y.K.; Lee, M.S.; Joe, S.H.; Kim, H.; et al. Brain-derived neurotrophic factor promoter methylation and cortical thickness in recurrent major depressive disorder. *Sci. Rep.* **2016**, *6*, 21089. [CrossRef]

42. Liu, X.; Kakeda, S.; Watanabe, K.; Yoshimura, R.; Abe, O.; Ide, S.; Hayashi, K.; Katsuki, A.; Umeno-Nakano, W.; Watanabe, R.; et al. Relationship between the cortical thickness and serum cortisol levels in drug-naïve, first-episode patients with major depressive disorder: A surface-based morphometric study. *Depress. Anxiety* **2015**, *32*, 702–708. [CrossRef]

43. Wise, T.; Radua, J.; Via, E.; Cardoner, N.; Abe, O.; Adams, T.M.; Amico, F.; Cheng, Y.; Cole, J.H.; de Azevedo Marques Périco, C.; et al. Common and distinct patterns of grey-matter volume alteration in major depression and bipolar disorder: Evidence from voxel-based meta-analysis. *Mol. Psychiatry* **2017**, *22*, 1455–1463. [CrossRef]

44. Järnum, H.; Eskildsen, S.F.; Steffensen, E.G.; Lundbye-Christensen, S.; Simonsen, C.W.; Thomsen, I.S.; Fründ, E.T.; Théberge, J.; Larsson, E.M. Longitudinal MRI study of cortical thickness, perfusion, and metabolite levels in major depressive disorder. *Acta Psychiatr. Scand.* **2011**, *124*, 435–446. [CrossRef]

45. Kim, J.H.; Suh, S.I.; Lee, H.J.; Lee, J.H.; Lee, M.S. Cortical and subcortical gray matter alterations in first-episode drug-naïve adolescents with major depressive disorder. *Neuroreport* **2019**, *30*, 1172–1178. [CrossRef]

46. Li, Q.; Zhao, Y.; Chen, Z.; Long, J.; Dai, J.; Huang, X.; Lui, S.; Radua, J.; Vieta, E.; Kemp, G.J.; et al. Meta-analysis of cortical thickness abnormalities in medication-free patients with major depressive disorder. *Neuropsychopharmacology* **2020**, *45*, 703–712. [CrossRef]

47. Bracht, T.; Linden, D.; Keedwell, P. A review of white matter microstructure alterations of pathways of the reward circuit in depression. *J. Affect. Disord.* **2015**, *187*, 45–53. [CrossRef]

48. Beaulieu, C. The basis of anisotropic water diffusion in the nervous system-a technical review. *NMR Biomed.* **2002**, *15*, 435–455. [CrossRef]

49. Zhang, A.; Leow, A.; Ajilore, O.; Lamar, M.; Yang, S.; Joseph, J.; Medina, J.; Zhan, L.; Kumar, A. Quantitative tract-specific measures of uncinate and cingulum in major depression using diffusion tensor imaging. *Neuropsychopharmacology* **2012**, *37*, 959–967. [CrossRef]

50. Manelis, A.; Soehner, A.; Halchenko, Y.O.; Satz, S.; Ragozzino, R.; Lucero, M.; Swartz, H.A.; Phillips, M.L.; Versace, A. White matter abnormalities in adults with bipolar disorder type-II and unipolar depression. *Sci. Rep.* **2021**, *11*, 7541. [CrossRef]

51. Ota, M.; Noda, T.; Sato, N.; Hattori, K.; Hori, H.; Sasayama, D.; Teraishi, T.; Nagashima, A.; Obu, S.; Higuchi, T.; et al. White matter abnormalities in major depressive disorder with melancholic and atypical features: A diffusion tensor imaging study. *Psychiatry Clin. Neurosci.* **2015**, *69*, 360–368. [CrossRef] [PubMed]

52. Zou, K.; Huang, X.; Li, T.; Gong, Q.; Li, Z.; Ou-yang, L.; Deng, W.; Chen, Q.; Li, C.; Ding, Y.; et al. Alterations of white matter integrity in adults with major depressive disorder: A magnetic resonance imaging study. *J. Psychiatry Neurosci.* **2008**, *33*, 525–530. [PubMed]

53. Zhu, X.; Wang, X.; Xiao, J.; Zhong, M.; Liao, J.; Yao, S. Altered white matter integrity in first-episode, treatment-naive young adults with major depressive disorder: A tract-based spatial statistics study. *Brain Res.* **2011**, *1369*, 223–229. [CrossRef] [PubMed]

54. Chen, G.; Hu, X.; Li, L.; Huang, X.; Lui, S.; Kuang, W.; Ai, H.; Bi, F.; Gu, Z.; Gong, Q. Disorganization of white matter architecture in major depressive disorder: A meta-analysis of diffusion tensor imaging with tract-based spatial statistics. *Sci. Rep.* **2016**, *6*, 21825. [CrossRef]

55. Kieseppä, T.; Eerola, M.; Mäntylä, R.; Neuvonen, T.; Poutanen, V.P.; Luoma, K.; Tuulio-Henriksson, A.; Jylhä, P.; Mantere, O.; Melartin, T.; et al. Major depressive disorder and white matter abnormalities: A diffusion tensor imaging study with tract-based spatial statistics. *J. Affect. Disord.* **2010**, *120*, 240–244. [CrossRef]

56. Sugimoto, K.; Kakeda, S.; Watanabe, K.; Katsuki, A.; Ueda, I.; Igata, N.; Igata, R.; Abe, O.; Yoshimura, R.; Korogi, Y. Relationship between white matter integrity and serum inflammatory cytokine levels in drug-naive patients with major depressive disorder: Diffusion tensor imaging study using tract-based spatial statistics. *Transl. Psychiatry* **2018**, *8*, 141. [CrossRef]

57. Han, K.M.; Choi, S.; Jung, J.; Na, K.S.; Yoon, H.K.; Lee, M.S.; Ham, B.J. Cortical thickness, cortical and subcortical volume, and white matter integrity in patients with their first episode of major depression. *J. Affect. Disord.* **2014**, *155*, 42–48. [CrossRef]

58. Zhou, L.; Wang, L.; Wang, M.; Dai, G.; Xiao, Y.; Feng, Z.; Wang, S.; Chen, G. Alterations in white matter microarchitecture in adolescents and young adults with major depressive disorder: A voxel-based meta-analysis of diffusion tensor imaging. *Psychiatry Res. Neuroimaging* **2022**, *323*, 111482. [CrossRef]

59. de Diego-Adeliño, J.; Pires, P.; Gómez-Ansón, B.; Serra-Blasco, M.; Vives-Gilabert, Y.; Puigdemont, D.; Martín-Blanco, A.; Alvarez, E.; Pérez, V.; Portella, M.J. Microstructural white-matter abnormalities associated with treatment resistance, severity and duration of illness in major depression. *Psychol. Med.* **2014**, *44*, 1171–1182. [CrossRef]

60. Zheng, K.Z.; Wang, H.N.; Liu, J.; Xi, Y.B.; Li, L.; Zhang, X.; Li, J.M.; Yin, H.; Tan, Q.R.; Lu, H.B.; et al. Incapacity to control emotion in major depression may arise from disrupted white matter integrity and OFC-amygdala inhibition. *CNS Neurosci. Ther.* **2018**, *24*, 1053–1062. [CrossRef]

61. Uchida, M.; Hung, Y.; Green, A.; Kelberman, C.; Capella, J.; Gaillard, S.L.; Gabrieli, J.D.E.; Biederman, J. Association between frontal cortico-limbic white-matter microstructure and risk for pediatric depression. *Psychiatry Res. Neuroimaging* **2021**, *318*, 111396. [CrossRef]

62. Korgaonkar, M.S.; Williams, L.M.; Song, Y.J.; Usherwood, T.; Grieve, S.M. Diffusion tensor imaging predictors of treatment outcomes in major depressive disorder. *Br. J. Psychiatry* **2014**, *205*, 321–328. [CrossRef]

63. Grieve, S.M.; Korgaonkar, M.S.; Gordon, E.; Williams, L.M.; Rush, A.J. Prediction of nonremission to antidepressant therapy using diffusion tensor imaging. *J. Clin. Psychiatry* **2016**, *77*, e436–e443. [CrossRef]

64. Mwangi, B.; Ebmeier, K.P.; Matthews, K.; Steele, J.D. Multi-centre diagnostic classification of individual structural neuroimaging scans from patients with major depressive disorder. *Brain* **2012**, *135*, 1508–1521. [CrossRef]

65. Gong, Q.; Wu, Q.; Scarpazza, C.; Lui, S.; Jia, Z.; Marquand, A.; Huang, X.; McGuire, P.; Mechelli, A. Prognostic prediction of therapeutic response in depression using high-field MR imaging. *Neuroimage* **2011**, *55*, 1497–1503. [CrossRef]

66. Korgaonkar, M.S.; Rekshan, W.; Gordon, E.; Rush, A.J.; Williams, L.M.; Blasey, C.; Grieve, S.M. Magnetic resonance imaging measures of brain structure to predict antidepressant treatment outcome in major depressive disorder. *EBioMedicine* **2015**, *2*, 37–45. [CrossRef]

67. Johnston, B.A.; Steele, J.D.; Tolomeo, S.; Christmas, D.; Matthews, K. Structural MRI-based predictions in patients with treatment-refractory depression (TRD). *PLoS ONE* **2015**, *10*, e0132958. [CrossRef]

68. Bartlett, E.A.; DeLorenzo, C.; Sharma, P.; Yang, J.; Zhang, M.; Petkova, E.; Weissman, M.; McGrath, P.J.; Fava, M.; Ogden, R.T.; et al. Pretreatment and early-treatment cortical thickness is associated with SSRI treatment response in major depressive disorder. *Neuropsychopharmacology* **2018**, *43*, 2221–2230. [CrossRef]

69. Redlich, R.; Opel, N.; Grotegerd, D.; Dohm, K.; Zaremba, D.; Bürger, C.; Münker, S.; Mühlmann, L.; Wahl, P.; Heindel, W.; et al. Prediction of individual response to electroconvulsive therapy via machine learning on structural magnetic resonance imaging data. *JAMA Psychiatry* **2016**, *73*, 557–564. [CrossRef]

70. Cao, B.; Luo, Q.; Fu, Y.; Du, L.; Qiu, T.; Yang, X.; Chen, X.; Chen, Q.; Soares, J.C.; Cho, R.Y.; et al. Predicting individual responses to the electroconvulsive therapy with hippocampal subfield volumes in major depression disorder. *Sci. Rep.* **2018**, *8*, 5434. [CrossRef]

71. Gärtner, M.; Ghisu, E.; Herrera-Melendez, A.L.; Koslowski, M.; Aust, S.; Asbach, P.; Otte, C.; Regen, F.; Heuser, I.; Borgwardt, K.; et al. Using routine MRI data of depressed patients to predict individual responses to electroconvulsive therapy. *Exp. Neurol.* **2021**, *335*, 113505. [CrossRef]

72. Takamiya, A.; Liang, K.C.; Nishikata, S.; Tarumi, R.; Sawada, K.; Kurokawa, S.; Hirano, J.; Yamagata, B.; Mimura, M.; Kishimoto, T. Predicting individual remission after electroconvulsive therapy based on structural magnetic resonance imaging: A machine learning approach. *J. ECT* **2020**, *36*, 205–210. [CrossRef]

73. Tymofiyeva, O.; Yuan, J.P.; Huang, C.Y.; Connolly, C.G.; Henje Blom, E.; Xu, D.; Yang, T.T. Application of machine learning to structural connectome to predict symptom reduction in depressed adolescents with cognitive behavioral therapy (CBT). *Neuroimage Clin.* **2019**, *23*, 101914. [CrossRef]

74. Marquand, A.F.; Mourão-Miranda, J.; Brammer, M.J.; Cleare, A.J.; Fu, C.H. Neuroanatomy of verbal working memory as a diagnostic biomarker for depression. *Neuroreport* **2008**, *19*, 1507–1511. [CrossRef]

75. Frässle, S.; Marquand, A.F.; Schmaal, L.; Dinga, R.; Veltman, D.J.; van der Wee, N.J.A.; van Tol, M.J.; Schöbi, D.; Penninx, B.; Stephan, K.E. Predicting individual clinical trajectories of depression with generative embedding. *Neuroimage Clin.* **2020**, *26*, 102213. [CrossRef]

76. Tian, S.; Sun, Y.; Shao, J.; Zhang, S.; Mo, Z.; Liu, X.; Wang, Q.; Wang, L.; Zhao, P.; Chattun, M.R.; et al. Predicting escitalopram monotherapy response in depression: The role of anterior cingulate cortex. *Hum. Brain Mapp.* **2020**, *41*, 1249–1260. [CrossRef]

77. Liu, Y.; Admon, R.; Mellem, M.S.; Belleau, E.L.; Kaiser, R.H.; Clegg, R.; Beltzer, M.; Goer, F.; Vitaliano, G.; Ahammad, P.; et al. Machine learning identifies large-scale reward-related activity modulated by dopaminergic enhancement in major depression. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* **2020**, *5*, 163–172. [CrossRef]

78. Osuch, E.; Gao, S.; Wammes, M.; Théberge, J.; Willimason, P.; Neufeld, R.J.; Du, Y.; Sui, J.; Calhoun, V. Complexity in mood disorder diagnosis: fMRI connectivity networks predicted medication-class of response in complex patients. *Acta Psychiatr. Scand.* **2018**, *138*, 472–482. [CrossRef]

79. Hopman, H.J.; Chan, S.M.S.; Chu, W.C.W.; Lu, H.; Tse, C.Y.; Chau, S.W.H.; Lam, L.C.W.; Mak, A.D.P.; Neggers, S.F.W. Personalized prediction of transcranial magnetic stimulation clinical response in patients with treatment-refractory depression using neuroimaging biomarkers and machine learning. *J. Affect. Disord.* **2021**, *290*, 261–271. [CrossRef]

80. Cash, R.F.H.; Cocchi, L.; Anderson, R.; Rogachov, A.; Kucyi, A.; Barnett, A.J.; Zalesky, A.; Fitzgerald, P.B. A multivariate neuroimaging biomarker of individual outcome to transcranial magnetic stimulation in depression. *Hum. Brain Mapp.* **2019**, *40*, 4618–4629. [CrossRef]

81. Wang, J.; Wei, Q.; Yuan, X.; Jiang, X.; Xu, J.; Zhou, X.; Tian, Y.; Wang, K. Local functional connectivity density is closely associated with the response of electroconvulsive therapy in major depressive disorder. *J. Affect. Disord.* **2018**, *225*, 658–664. [CrossRef] [PubMed]

82. Pei, C.; Sun, Y.; Zhu, J.; Wang, X.; Zhang, Y.; Zhang, S.; Yao, Z.; Lu, Q. Ensemble learning for early-response prediction of antidepressant treatment in major depressive disorder. *J. Magn. Reson. Imaging* **2020**, *52*, 161–171. [CrossRef] [PubMed]

83. Gartlehner, G.; Thaler, K.; Hill, S.; Hansen, R.A. How should primary care doctors select which antidepressants to administer? *Curr. Psychiatry Rep.* **2012**, *14*, 360–369. [CrossRef] [PubMed]

84. Rush, A.J.; Trivedi, M.H.; Wisniewski, S.R.; Nierenberg, A.A.; Stewart, J.W.; Warden, D.; Niederehe, G.; Thase, M.E.; Lavori, P.W.; Lebowitz, B.D.; et al. Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: A STAR*D report. *Am. J. Psychiatry* **2006**, *163*, 1905–1917. [CrossRef]

85. Soriano-Mas, C.; Hernández-Ribas, R.; Pujol, J.; Urretavizcaya, M.; Deus, J.; Harrison, B.J.; Ortiz, H.; López-Solà, M.; Menchón, J.M.; Cardoner, N. Cross-sectional and longitudinal assessment of structural brain alterations in melancholic depression. *Biol. Psychiatry* **2011**, *69*, 318–325. [CrossRef]

86. Takahashi, T.; Yücel, M.; Lorenzetti, V.; Tanino, R.; Whittle, S.; Suzuki, M.; Walterfang, M.; Pantelis, C.; Allen, N.B. Volumetric MRI study of the insular cortex in individuals with current and past major depression. *J. Affect. Disord.* **2010**, *121*, 231–238. [CrossRef]

87. Trivedi, M.H.; McGrath, P.J.; Fava, M.; Parsey, R.V.; Kurian, B.T.; Phillips, M.L.; Oquendo, M.A.; Bruder, G.; Pizzagalli, D.; Toups, M.; et al. Establishing moderators and biosignatures of antidepressant response in clinical care (EMBARC): Rationale and design. *J. Psychiatr. Res.* **2016**, *78*, 11–23. [CrossRef]

88. Bellani, M.; Dusi, N.; Yeh, P.H.; Soares, J.C.; Brambilla, P. The effects of antidepressants on human brain as detected by imaging studies. Focus on major depression. *Prog. Neuro-Psychopharmacol. Biol. Psychiatry* **2011**, *35*, 1544–1552. [CrossRef]

89. Lee, Y.; Ragguett, R.M.; Mansur, R.B.; Boutilier, J.J.; Rosenblat, J.D.; Trevizol, A.; Brietzke, E.; Lin, K.; Pan, Z.; Subramaniapillai, M.; et al. Applications of machine learning algorithms to predict therapeutic outcomes in depression: A meta-analysis and systematic review. *J. Affect. Disord.* **2018**, *241*, 519–532. [CrossRef]

90. Aoki, Y.; Yamaguchi, S.; Ando, S.; Sasaki, N.; Bernick, P.J.; Akiyama, T. The experience of electroconvulsive therapy and its impact on associated stigma: A meta-analysis. *Int. J. Soc. Psychiatry* **2016**, *62*, 708–718. [CrossRef]

91. Penninx, B.W.; Beekman, A.T.; Smit, J.H.; Zitman, F.G.; Nolen, W.A.; Spinhoven, P.; Cuijpers, P.; De Jong, P.J.; Van Marwijk, H.W.; Assendelft, W.J.; et al. The Netherlands Study of Depression and Anxiety (NESDA): Rationale, objectives and methods. *Int. J. Methods Psychiatr. Res.* **2008**, *17*, 121–140. [CrossRef]

92. Gillett, G.; Tomlinson, A.; Efthimiou, O.; Cipriani, A. Predicting treatment effects in unipolar depression: A meta-review. *Pharmacol. Ther.* **2020**, *212*, 107557. [CrossRef]

93.	Janssen, R.J.; Mourão-Miranda, J.; Schnack, H.G. Making individual prognoses in psychiatry using neuroimaging and machine learning. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* **2018**, *3*, 798–808. [CrossRef]

94.	Kim, Y.K.; Na, K.S. Application of machine learning classification for structural brain MRI in mood disorders: Critical review from a clinical perspective. *Prog. Neuro-Psychopharmacol. Biol. Psychiatry* **2018**, *80*, 71–80. [CrossRef]