



Machine Learning and Brain Imaging for Psychiatric Disorders: New Perspectives

Ivan Brossollet, Quentin Gallet, Pauline Favre, and Josselin Houenou

Abstract

Psychiatric disorders include a broad panel of heterogeneous conditions. Among the most severe psychiatric diseases, in intensity and incidence, depression will affect 15–20% of the population in their lifetime, schizophrenia 0.7–1%, and bipolar disorder 1–2.5%. Today, the diagnosis is solely based on clinical evaluation, causing major issues since it is subjective and as different diseases can present similar symptoms. These limitations in diagnosis lead to limitations in the classification of psychiatric diseases and treatments. There is therefore a great need for new biomarkers, usable at an individual level. Among them, magnetic resonance imaging (MRI) allows to measure potential brain abnormalities in patients with psychiatric disorders. This creates datasets with high dimensionality and very subtle variations between healthy subjects and patients, making machine and statistical learning ideal tools to extract biomarkers from these data. Machine learning brings different tools that could be useful to tackle these issues. On the one hand, supervised learning can support automated classification between different psychiatric conditions. On the other hand, unsupervised learning could allow the identification of new homogeneous subgroups of patients, refining our understanding of the classification of these disorders. In this chapter, we will review current research applying machine learning tools to brain imaging in psychiatry, and we will discuss its interest, limitations, and future applications.

Key words Psychiatry, Depression, Schizophrenia, Bipolar disorder, Machine learning, Artificial intelligence, Neuroimaging, MRI, Clustering, Classification

1 Introduction

Major psychiatric conditions affecting adults can be classified into several groups: affective disorders (e.g., bipolar disorders, major depressive disorders), psychotic disorders (e.g., schizophrenia), anxiety disorders (e.g., obsessive-compulsive disorders), neurodevelopmental disorders (e.g., autism), and substance use disorders. We will focus this chapter on the two first categories, as they carry a high individual and societal burden and are highly prevalent throughout the world.

1.1 Major Depressive Disorder

Major depressive disorder (MDD) is defined by the occurrence of one or more major depressive episodes without any manic or hypomanic episodes in the lifetime. Its prevalence can vary significantly according to the studies, but exceeds 15% of the population during their lifetime [1], and affects two women for one man. Depression can affect people at any time during their life [2]. Nowadays, the diagnosis is based on structured interviews, and the clinical criteria are given by, among others, two classification manuals: the International Classification of Diseases [1] and the fifth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) [3]. According to the DSM-5, to meet the criteria for a major depressive episode, five of the nine following symptoms must be present over a 2-week period: depressed mood or anhedonia (loss of interest or pleasure), change in weight or appetite, sleep disturbances (insomnia or hypersomnia), psychomotor retardation or restlessness, loss of energy or fatigue, low self-esteem or guilt, difficulty in concentrating or indecisiveness, and thoughts of death or suicidal thoughts. Patients with MDD are at an increased risk of other comorbid disorders. Most commonly, they may present alcohol abuse or dependence, anxiety disorders such as panic disorder, obsessive-compulsive disorder, and generalized anxiety disorder. Treatment options for MDD include a variable combination of pharmacotherapy (antidepressants such as serotonin selective reuptake inhibitors or tricyclics) and psychotherapy (cognitive behavioral therapy, interpersonal therapy, etc.). Despite considerable progress in its diagnosis and treatment, MDD remains underdiagnosed and underestimated and remains a challenge for healthcare institutions, especially since one of the main risks of mood disorders (BD or MDD) is suicidal behavior.

1.2 Bipolar Disorder

Bipolar disorder (BD) is defined as a chronic mood disorder characterized by episodes of depression and episodes of abnormal excitation (mania, hypomania), separated by periods of “euthymia” (without any symptoms of major mood episode) [3]. This mood disorder affects around 1% of the world’s adult population [4], regardless of continent, socioeconomic status, or ethnicity. The course of BD is lifelong, but is heterogeneous in terms of number of episodes, relapses, polarity (i.e., higher number of manic or depressive episodes), and response to treatment. The impact of the disease on cognitive function and quality of life can be major [4]. Diagnosis, treatment, health, and social care are major goals in the management of BD.

Manic episodes are defined by a period lasting at least 1 week, during which patients exhibit elevated mood and increased motor activity. The intensity of these symptoms defines the manic or hypomanic nature of the episode. During a manic episode, patients may experience psychotic symptoms such as hallucinations, delusions, disorganized thinking, and sleep disturbances. The delusions may be consistent with the manic mood, with individuals displaying

grandiosity, megalomania, or messianic ideas. Impaired judgment and risk of endangering the patient often lead to hospitalization. Hypomanic episodes are characterized by lower symptom intensity (abnormally high, expansive, or irritable mood, as well as abnormal increase in activity or energy, most of the day) and must last at least 4 consecutive days. Although there are no pathognomonic features of bipolar or unipolar depression, some clinical features are useful in distinguishing them: bipolar depression usually occurs at an earlier age, and the episodes are also more frequent and shorter, show an abrupt onset and termination, and are more frequently associated with substance abuse. Patients with bipolar depression may also present atypical symptoms, such as hypersomnia and weight instability. Psychosis (delusions and hallucinations) and catatonia are also more frequent in bipolar depression, whereas somatic complaints are more common in unipolar depression. The presence of a family history of mania is also a relevant indicator of bipolar depression. The establishment of the diagnosis of BD is a major challenge and has several consequences: stabilizing the disease, allowing good social reintegration, avoiding relapses and side effects, and, finally, limiting the long-term effects of the disease, particularly on the cognitive level. Treatment strategies usually combine pharmacotherapy (mostly mood stabilizers) and psychosocial care, tailored to each patient. Mood stabilizers aim at decreasing the frequency of major mood episodes. Lithium, some anticonvulsants (such as valproate and carbamazepine), and some antipsychotics (such as aripiprazole, quetiapine, or olanzapine) are the three classes of available mood stabilizers. Psychosocial care includes cognitive rehabilitation strategies, psychoeducation, and interpersonal social and rhythm therapies.

1.3 Schizophrenia

The annual incidence of schizophrenia is 0.2–0.4 per 1000, with a lifetime prevalence of about 0.8% [5], which can slightly vary between countries and cultural groups [6]. These differences are reduced when stricter diagnostic criteria are used for schizophrenia, such as the ones of the DSM-5. Research conducted by the WHO has further confirmed this observation by showing that schizophrenic disorder prevalence is similar across a wide range of cultures and countries, including developed and developing countries [6]. Its sex ratio is around 1:1.

Schizophrenia is characterized by three main types of symptoms, namely, positive symptoms, negative symptoms, and cognitive impairment [7]. Positive symptoms involve a loss of contact with reality; the patient has false beliefs (delusions) and perceptual experiences not shared with others (hallucinations) and may exhibit behavioral oddities. People with schizophrenia can experience different kinds of hallucinations: auditory, visual, olfactory, gustatory, or tactile. About delusions, patients with schizophrenia may have persecutory delusions, control delusions (e.g., belief in telepathy), grandiose delusions (e.g., belief in being a god), and somatic

delusions (e.g., belief that one's body is rotting from the inside) [8]. Negative symptoms are characterized by a deficit state during which basic emotional and behavioral processes are diminished or absent. The most common negative symptoms are blunted affect, anhedonia, avolition, apathy, and alogia (i.e., reduction in the amount or content of speech). Negative symptoms are more frequent and less fluctuating over time than positive symptoms [9]. They are also strongly associated with poor psychosocial functioning [10]. Cognitive impairments in schizophrenia include deficits with attention and concentration, psychomotor speed, learning and memory, and executive function. A decline in cognitive abilities from premorbid functioning is present in most of the patients, with cognitive functioning after the onset of the illness being relatively stable over time [10]. Despite this decline, cognitive functioning in some patients could be in the normal range. As for the negative symptoms, cognitive impairment is strongly associated with poor psychosocial functioning, particularly with regard to social and professional lives.

The etiology of schizophrenia is complex and multifactorial. Genetic and environmental factors seem to play a major role. The risk of developing schizophrenia is higher in patients' relatives than in the general population [11, 12]. Adoption and twin studies have shown that this increased risk is genetic, with the risk being increased by the presence of an affected first-degree relative [12]. There are two main approaches to the treatment of schizophrenia: pharmacological and psychosocial treatments [13]. Antipsychotics constitute the main medication, with major effects on reducing positive symptoms and preventing relapses. First-generation antipsychotics include molecules such as chlorpromazine or haloperidol. Second-generation antipsychotics were developed to decrease the neurological and cognitive side effects. They are the most used molecules nowadays (quetiapine, aripiprazole, risperidone, clozapine, etc.). In contrast, their effects on negative symptoms and cognitive impairment are much more moderate [14]. Psychosocial interventions improve the management of schizophrenia, e.g., through symptom management or relapse prevention. Other specific interventions that can improve the outcome of schizophrenia include family psychoeducation, supported employment, social skills training, psychoeducation, cognitive behavioral therapy, and integrated treatment of comorbid substance abuse [8].

The remainder of this chapter is organized as follows: We first describe the challenges in psychiatry that can potentially be addressed with machine learning. We then provide a non-exhaustive state of the art of machine learning with magnetic resonance imaging in psychiatry. We finally highlight the limitations of current approaches and propose perspectives for the field. Studies reviewed in this chapter are summarized in Table 1.

2 Challenges for Machine Learning in Psychiatry

Diagnosis and treatment are based on clinical diagnostic criteria developed from the subjective human experience, rather than on objective markers of illness. These criteria have been developed based on experts' opinion and are included in the DSM-5 and ICD-10 manuals. This approach has some limitations. Diagnosis can vary across interview methodologies [50], and clinically identical symptoms can be caused by different underlying conditions. Therefore, the common diagnostic criteria, which are based on symptom manifestation alone, are not always reliable in the clinical context [51]. They are indeed often unstable over time and unspecific [52] and provide little guidance to select the appropriate treatment. These misdiagnoses and misclassifications could lead to a poor therapeutic response and suboptimal management of the illness. Based on these observations, it appears necessary to develop objective markers and a better characterization of these illnesses.

In this section, we will discuss how machine learning could be used to improve diagnosis, to help characterize the different mental illnesses, and to improve treatment response and prognostic approach.

2.1 Improving the Diagnosis of Psychiatric Disorders

In the early stages of research on machine learning and psychiatric disorders, researchers wanted to explore whether different diagnoses could be predicted using machine learning algorithms applied to neuroimaging features. They mainly applied machine learning on structural MRI (sMRI) and functional MRI (fMRI) data (during tasks or at rest) [53]. Recent efforts have been made to apply machine learning on diffusion MRI [15], mostly in combination with other modalities [53, 54], and to explore whether adding modalities improves the diagnosis. Classification using machine learning in neuroimaging initially focused on major psychiatric disorders, such as MDD [55], schizophrenia [56], and bipolar disorder [54]. In a second phase, research has broadened the spectrum of psychiatric disorders such as anxiety disorders [23], anorexia [20], substance abuse [57], specific phobia [19], and autism spectrum disorders [58]. Machine learning using EEG has also been investigated for schizophrenia classification [59] as it is an affordable method for functional imaging and since it has a better temporal resolution than fMRI. While lots of machine learning studies in psychiatry focused on neuroimaging data, other fields of research were increasingly interested in using other modalities, such as proteomic, metabolomic [22], and genetic [24] data.

Machine learning also opens perspectives for the identification of relevant features (e.g., the measured variables) for the diagnosis. Using interpretable models such as support vector machines (SVM) or decision trees lets researchers investigate features that are used in

the decision. Deep learning could also be used to find useful features without a priori preprocessing of the images when it is used in combination with interpretation techniques [59]. Another way to identify relevant features for the classification is to compare the prediction performances of different machine learning models with different input features. It then allows us to evaluate if the information present in the different features helps the classification. For example, this was shown in the study of Lin et al. [16], where the authors established that the G72 protein alone yielded almost as much information for the diagnosis of schizophrenia than combined with other G72 single nucleotide polymorphisms. While this approach could be fruitful to build more resilient and interpretable algorithms, we should be careful when interpreting their results. We must keep in mind that statistical algorithms such as the machine learning ones are designed to predict (classes), while inference tests (i.e., univariate statistics) usually rely on association studies, which are more reliable to infer correlation and causal relations [60]. Moreover, when interpreting SVM weights, for example, one must keep in mind that some features are only including noise but are still important when considered in combination with other features [61]. For all these reasons, even though finding important features is necessary to better understand the models, their interpretation to infer pathophysiology or biomarkers must be cautious.

2.2 Refining the Classification of Psychiatric Disorders

Since there is a significant overlap in the clinical symptoms of different psychiatric disorders, many patients suffer from an important delay in the diagnostic establishment, after a potentially harmful diagnosis wavering. For instance, patients with BD wait on average 10 years before receiving an accurate diagnosis [62] and are often misdiagnosed with unipolar depression for years. As for MDD, it is often underdiagnosed even though fast and accurate diagnosis could avoid long-term cognitive impairment in under-treated patients [63]. For all these reasons, making the right diagnosis as early as possible is a major public health challenge.

Machine learning may be a useful tool to discriminate between different diagnoses. Indeed, the interest in machine learning is not only to distinguish a patient with a psychiatric disorder from a healthy subject – which is not the most difficult task for the clinician – but it could be used to help the clinician when the diagnosis becomes more difficult, e.g., to distinguish bipolar depression from unipolar depression [18] or to identify a patient at risk of psychosis [24].

As studies investigate new biomarkers to differentiate between different conditions, our current classification of psychiatric disorders appears to be limited. There are numerous different classification criteria to describe psychopathology, and theoretical frameworks are evolving rapidly [52], which contributes to our

limited understanding of these disorders. The classification of psychiatric disorders is also a complex issue at the biological level, since biological boundaries between conditions are not binary and are blurred by the imprecision of the current genetic and imaging tools (e.g., between BD and schizophrenia [17]). Moreover, the heterogeneity in the clinical presentation of the patients limits the efficiency of a binary classification task. A simple classification algorithm as SVM will only find the largest and shared biomarkers, leading to a suboptimal classification.

The question we might ask is whether changing our perspective and the way we approach psychiatric disorders' heterogeneity will improve our understanding and management of the patients. To consider this heterogeneity, unsupervised machine learning seems to be an appropriate method, as it allows to find new homogeneous subgroup within the population without preconceptions. Current research is using unsupervised machine learning to automatically detect new subgroups (i.e., clusters) of patients based on similar cognitive [25], genetic [64], and/or cerebral [64] profiles. After subgrouping, supervised machine learning can be used to automatically classify the patients into one group or another. For instance, Wu et al. [25] identified two phenotypic groups of patients with BD using a cognitive task battery. Then, they used classifiers to detect white matter tracts' microstructural differences between the two groups. Newly developed algorithms combining supervised learning and clustering show promising results [65], as they can disentangle the heterogeneity of some disorders and improve diagnostic prediction at the same time. The HYDRA model is one of those promising algorithms that has already been used to find some subtypes of Alzheimer disease and to reveal meaningful biomarkers of this disease at the same time [66]. These semi-supervised clustering algorithms [67] are also starting to be used in psychiatry [68] as they could help to reveal biomarkers while discriminating between two different homogeneous classes. Finally, these algorithms are of special interest as they are also handling common source of variation in the groups to be classified (i.e., the age, the sex, or other clinical or biological variables) [69].

Other approaches aim to identify differences between the patients (the cases) and a reference population [70] (the controls). These so-called normative models drop the hypothesis that the patients do not belong to a homogeneous group, which is a step toward a finer analysis. Indeed, recent studies showed important clinical and biological heterogeneity between the patients, especially regarding brain structural abnormalities. Therefore, the hypothesis of an average patient, as it is in classical "case-control" studies, could limit our understanding of the diseases in the long term. Normative modeling could overpass this limitation as it allows to situate a given patient among the "norm" while considering the strong heterogeneity within the patients' population. For

instance, Wolfer et al. [70] showed that deviations from the normative model of gray matter volume were frequent in both SZ and BD but highly heterogeneous. However, these models also induce an asymmetry as they hypothesize that the controls are homogeneous, which is debatable in practice. Nevertheless, it appears that subtyping leads to increased predictive accuracy in identifying individuals with mental illnesses compared with healthy controls, even though results are mixed [71]. This approach could gain attention with the development of new tools such as longitudinal normative brain charts that cover the whole lifespan [72].

2.3 Predicting Evolution and Treatment Response

Predicting the evolution of psychiatric disorders is an important challenge. As previously mentioned, clinicians' choices are guided by recommendations based on broad symptom classifications, such as depression, anxiety, or psychosis criteria, and become personalized over time through an empirical process of trials and errors. Being able to predict the prognosis of the mental illnesses would allow a better organization of care and more adapted psychoeducation consultations, would let clinicians set up strategies to prevent relapses, and would finally greatly improve the quality of life of the patients. Some studies tried to predict psychotic transition using neuroimaging [29] or using EEG [32] and clinical measures [35]. Schmaal et al. [31] used Gaussian process classifiers based on structural and functional MRI (emotional task) to characterize trajectories of depression (chronic, improvement, and rapid remission). They successfully classified the chronic group vs. the rapid remission group with an accuracy of 73%. Regarding other studies on depression, Kessler et al. [73] used self-reported clinical questionnaires of 1057 patients and machine learning algorithm to predict the course of MDD. They predicted the risk of suicide attempt with an AUC of 0.76 and whether the patient would experience a depressive episode lasting more than 2 weeks with an AUC of 0.71. Tran et al. [34] used electronic record's information such as medication, diagnosis, occurrence of interactions with health services, etc. with the aim of stratifying individuals according to their suicide risk. Interestingly, according to their results, their algorithms predicted the suicide risk better than clinicians, with an AUC of 0.73 vs. 0.57 for the prediction of high suicide risk patient vs. the rest of the population. It could also be possible to predict future substance abuse using neuroimaging data [33] and using combinations of demographic, clinical, cognitive, neuroimaging, and genetic data [30]. For schizophrenia, EEG-based machine learning could also be used to determine at-risk patients [59]. Machine learning could also be useful to predict the outcome of a first episode of psychosis [42] and to adapt the treatment. These studies highlight the possibility to stratify and classify individuals to optimize prognostic assessments, thanks to machine

learning. That would help the clinician to propose personalized care, such as primary care facilities for patients at high suicidal risk.

Regarding the treatment outcome, the major challenge is to determine whether machine learning could be used to predict treatment response. This knowledge would be extremely useful, as for now therapeutic choices are made through a trial-error process, which increases the time interval between the apparition of the symptoms and the administration of the adequate treatment. This leads to a serious socioeconomic burden and can have debilitating consequences. In depression, the interest of the machine learning approach was tested on pharmacological decision, for instance, to predict the response to serotonin reuptake inhibitor medications [27]. The authors were able to predict the treatment response using EEG-derived features with an accuracy of 87.9%. In another study, EEG features were also used to predict antipsychotics response in schizophrenia [74]. More recently, studies focused on anatomical and functional MRI. For instance, Whitfield-Gabrieli et al. [28] used resting-state fMRI combined with FA maps as well as initial severity assessment to predict the response to cognitive behavioral therapy in patients with social anxiety. They were able to classify good and poor responders with an accuracy of 81% in a sample of 38 patients. Predicting treatment response is particularly interesting when the treatment is more invasive, such as for the use of electroconvulsive therapy (ECT). Indeed, one team showed (with a sample of 122 depressed patients) that the brain structure can predict the ECT response with an accuracy of 78% [75]. Finally, choosing the right treatment is not just about measuring its effectiveness; it is always about balancing the cost and the acceptable benefit for the patients. In summary, all these features could be integrated in machine learning algorithms and used by the clinicians as tools to improve the accuracy of the therapeutic decisions.

3 MRI and Machine Learning in Psychiatry: State of the Art (Table 1)

To this day, unlike in some medical specialties such as neurology, MRI is rarely used for psychiatric clinical practice. However, it is extensively used in research as it provides a large variety of information about the brain structure and function. Currently, sMRI is the easiest method to implement and the most used in the MRI studies. It is preferentially used to measure the cortex thickness and the cortical surface and to estimate the gray and white matter density and/or volume. Diffusion-weighted imaging (DWI) is less used but provides useful information on the white matter microstructure, thanks to different markers such as fractional anisotropy (the most used), mean diffusivity, and radial diffusivity. fMRI is of particular interest to investigate the neural correlates of cognition and emotion processes and their alteration in patients with

psychiatric conditions. Predictive models are thus useful tools when analyzing MRI data, because they allow to handle high-dimensional inputs and fit more unknown variables than available observations. In neuroimaging, machine learning allows to model sets of effects rather than single effects and thus to build models that describe more than one isolated dimension of cognition.

3.1 Classification Versus Healthy Controls

Classification of patients with psychiatric disorder vs. healthy controls is a widely studied area of research. Even though most studies fail to obtain the 80% of accuracy needed for clinical relevance, they yield promising results and give important methodological insights.

Regarding MDD, using sMRI, machine learning studies [55] found accuracies ranging from 67.6% to 90.3%. These results should be taken with great caution since they are usually obtained from small samples. For example, Mwangi et al. [39] obtained an accuracy of 90.3% using relevance vector machines and a sample of 60 subjects. They also showed that the brain regions identified during the features selection process were consistent with those of previous studies that reported gray matter reductions in patients with MDD, which were mostly located in the frontal lobe, the orbitofrontal and cingulate cortex, the middle frontal gyrus, and the inferior and superior gyri [76]. As for fMRI studies, Gao et al. [55] found an accuracy ranging from 56% to 99%; Ramasubbu et al. [36] found a significant accuracy of 66% for very severe depression using resting-state fMRI in 19 control subjects vs. 45 patients with different intensities of depression; and Fu et al. [21] obtained an accuracy of 86% in a sample of 19 patients with MDD and 19 HC who were processing sad faces during fMRI scanning.

Regarding bipolar disorder (BD), a recent literature review counted 25 studies using machine learning with different MRI modalities to classify BD vs. HC [54]. They found a median accuracy of 66% for BD vs. HC classification. Even though most studies used samples of less than 100 subjects, a study stood out by the number of samples. Using 3040 subjects, sMRI, and a linear SVM, Nunes et al. [43] obtained an accuracy of 65.23% using aggregate subject-level analyses and an accuracy of 58.67% when testing on left out sites. Their results, which highlighted the importance of regions such as the hippocampus, the amygdala, and the inferior frontal gyrus for the classification, were in good accordance with previous MRI studies in BD [76–78]. Regarding fMRI, the review of Claude et al. [54] highlighted that machine learning studies performed with an accuracy range between 37.5% and 83.5%. The minimum accuracy was 37.5% for the classification of bipolar depression vs. HC, during angry face processing using a Gaussian process classifier (GPC) [37]. DWI was not investigated much. In the review of Claude et al. [54], only two DWI studies were referenced. Achalia et al. [15] used DWI and machine learning on 60 subjects and obtained an accuracy of 74% for DWI alone. Even

though DWI gave lower classification scores than sMRI (77.8%) and fMRI (80.3%), combining it with other modalities significantly enhanced the accuracy (87.6%). Mwangi et al. [40] also used DWI in combination with sMRI on 30 pediatric patients with BD and obtained a classification accuracy of 78.12%.

Regarding schizophrenia (SZ), Filippis et al. [56] conducted a systematic review focusing on sMRI and fMRI studies that attempt to classify SZ vs. HC. Notably, the study of Salvador et al. [38] focused on a sample of 128 patients with SZ and 127 HC and aimed to compare the classification score of different neuroimaging features such as voxel-based and wavelet-based (a transformation like Fourier transform) morphometry of gray and white matter, vertex-based cortical thickness and volume defined as regions of interest, as well as volumetric measures. They also compared different methods, such as random forest, regressions with different regularization methods and levels, and SVM. The best results were obtained using the voxel-based and wavelet-based morphometry in combination with a SVM, with respective accuracy of 77.2% and 71%. The authors stress on the fact that no algorithm clearly outperforms the others, but that the selection of features has a real influence on the classification accuracy. Another notable study focused on cortical thickness and surface area measurement to differentiate first-episode psychosis from healthy subjects [42]. This study witnessed that regions contributing to the classification accuracy included the default mode network (DMN), the central executive network, the salience network, and the visual network. They observed a classification accuracy of 85.0% for the surface area and 81.8% for the cortical thickness. Pinaya et al. [79] used a deep belief network, which is a deep neural network that extrapolated and interpreted features, on sMRI data from 83 HC and 143 patients with SZ. The deep belief network highlighted an accuracy of 73.6% vs. 68.1% for a classical SVM. It also detected large differences between classes among specific regions, particularly frontal, temporal, parietal, and insular cortices, the corpus callosum, the putamen, and the cerebellum. Finally, as already mentioned in Subheading 2.1, normative models constructed with MRI data could be a useful tool to handle the inter-subject variability in machine learning models [71].

3.2 Inter-Illness Classification and Clustering

One major challenge of machine learning studies using MRI is to be able to correctly distinguish or classify patients suffering from different disorders. Several studies focused on the classification between BD and SZ. In their review, Claude et al. [54] found that three studies used sMRI in combination with machine learning algorithms to discriminate between BD and SZ with an accuracy ranging between 58% and 66%. Precisely, Schnack et al. [44] showed good classification performance on an independent dataset, with an average classification accuracy of 66%. Mothi et al. [45]

used K-mean clustering after a non-linear PCA to separate patients with BD, SZ, or schizoaffective disorder. They found out that the separation in three clusters was optimal, comprising a cluster including a major proportion of patients with BD, a second with mostly patients with SZ, and a third with a balanced proportion of the three types of illnesses. To build their clusters, they used clinical and cognitive data and validated the robustness of their results with sMRI data. The cluster including more patients with SZ was the one to have a significantly reduced cortical thickness in the frontal lobe. In addition, the BD and the SZ clusters presented significant cortical thickness reductions in occipital and temporal regions.

Several studies attempted to predict the diagnosis of BD in a population of unipolar, bipolar depression, and healthy controls with a median accuracy of 79% and an accuracy ranging from 50% to 90.69% [54]. Burger et al. [37] focused on the classification of unipolar vs. bipolar depression using different regions of interest. They did not find any significant results using the whole brain but found an accuracy of 63.89% for the classification of BD vs. unipolar depression using a GPC based on a happy face processing paradigm and the amygdala activity. Their best accuracy was of 72.2% for the classification of bipolar vs. unipolar depression, using a fear processing paradigm and GPC on the anterior cingulate gyrus. Overall, the best performance was obtained by Grotegerd et al. [18] In a pilot study, they obtained an accuracy of 90% using fMRI with a happy vs. neutral contrast image and an SVM on 10 BD, 10 HC, and 10 MDD. Using sMRI and DWI with a multiple kernel learning and a sample of 74 MDD and 74 BD, Vai et al. [46] obtained an accuracy of 74.32%, with a positive predictive value of 73.33% (probability that subjects with a positive BD prediction suffer from BD). The accuracy for MDD was 72.97%, indicating the ability to correctly identify people with MDD, with a predictive value of 73.97%. Their models are particularly interesting as they included relevant covariates in their models, such as age, gender, number of previous episodes, and drug load, which can confound and bias the accuracy estimates. Taking into account all these factors helps to increase the performance of the algorithm, as they impact the brain structural measures. It is necessary since these effects were witnessed by the ENIGMA-BD Working Group that used a large cohort of 2447 BD and 4056 HC and found [80] that several commonly prescribed drugs for BD treatment, including lithium, anti-epileptic, and antipsychotic treatments, showed significant associations with cortical thickness and surface area, even after accounting for patients receiving multiple drugs.

3.3 Treatment Response and Illness Prediction

Another perspective is the use of MRI and machine learning algorithms to predict treatment response. This was done by the team of Liu et al. [47] who tested the sensitivity to antidepressants in patients with MDD. Precisely, the study included 17 subjects that

Table 1
Summary of reviewed studies

Study	Modalities	Diseases and goal	Number of subjects	Methods and validation procedure	Results
Achalia et al. [15]	sMRI, dMRI, rs-fMRI, cognitive test	Classification of BD vs. HC	30 HC, 30 BD	SVM and feature selection	Accuracy of 87.60%, sensitivity of 82.3% and specificity of 92.7%
Lin et al. [16]	G72 genotypes, proteins levels	Classification of SZ vs. HC	89 SZ, 60 HC	Naive Bayes, logistic regression, decision tree	Naive Bayes model: sensitivity = 0.7969, specificity = 0.9372, (AUC) = 0.9356
Kliemann et al. [17]	Emotion attribution task, localizer task and behavioral task in fMRI	Classification of Autism spectrum disorder (ASD) vs. HC	Study 1: ASD: 16, HC: 21 Study 2: ASD: 22, HC: 30	ROI-based MVPA (Multivariate pattern analysis) with a SVM	Only statistics on the MVPA results
Grotegerd et al. [18]	Facial emotion processing, fMRI	Classification of BD vs. HC vs. Unipolar depression (UD)	BD: 10, HC: 10, UD: 10	SVM, GPC	Best results: happy versus neutral with the SVM for UD vs. BD: 90% accuracy
Visser et al. [19]	Generalization task, fMRI	Identifying area that influence the fear of spiders in spider-phobia stricken patients	High spider fear patients (HSF): 18, Low spider fear patient (LSF): 20	SVM-based MVPA	Several above chance results of correct classification of the stimuli depending on the ROIs
Lavagnino et al. [20]	sMRI	Classification of Anorexia Nervosa (AN) patient versus HC	AN: 15, HC: 15	LASSO linear regression	83.3% accuracy (sensitivity 86.7%, specificity 80.0%)
Fu et al. [21]	Sad face processing, fMRI	Classification of MDD versus HC	MDD: 19, HC: 19	SVM	Accuracy of 86% (84% of sensitivity and 89% of specificity)

(continued)

Table 1
(continued)

Study	Modalities	Diseases and goal	Number of subjects	Methods and validation procedure	Results
Setoyama et al. [22]	Plasma metabolites concentration	Prediction of suicide ideation (SI) in MDD and BD patients	Dataset 1: 51 drug-free MDD patients Dataset 2: 23 medicated MDD patients Dataset 3: 27 MDD, 14 BD	Logistic regression, SVM and random forest. Results reported for each fold of a 10-fold cross validation	Logistic regression accuracy of SI vs. no SI: 57.14% (mean of the 10-fold) SVM accuracy of SI vs. no SI: 52.08% (mean of the 10-fold) Random forest: overfit
Lueken et al. [23]	Fear conditioning task, fMRI	Panic disorder, MDD	Panic disorder: 33, Panic disorder with depression: 26	AdaBoost	Comorbidity status was correctly predicted in 79% of patients (sensitivity: 73%; specificity: 85%) based on brain activation during fear conditioning (corrected for potential confounders: accuracy: 73%; sensitivity: 77%; specificity: 70%)
Petterson-Yeo et al. [24]	Genotype, sMRI, dMRI, cognitive data, fMRI	Classification of ultra high risk (UHR) patients vs. first episode psychosis patient (FEP) vs. HC	HC: 23, UHR: 19, FEP: 19	SVM	FEP vs. HC (genotype, 67.86%; dMRI, 65.79%; fMRI, 65.79% and 68.42%; cognitive data, 73.69%), UHR versus HC (sMRI, 68.42%; dMRI, 65.79%), and FEP

						versus UHR (sMRI, 76.67%; fMRI, 73.33%; cognitive data, 66.67%).
Wu et al. [25]	Neurocognitive data, dMRI (FA, MD)	Unsupervised clustering of BD patients confirmed by supervised classification	BD: 70	K-means, ElasticNet	Identification of two phenotypes by K-means	
Drysdale et al. [26]	rsfMRI	Unsupervised clustering of depressive disorder	Dataset 1: 333 DD, 378 HC Dataset 2: 352 HC, 109 DD, 16 BD II	K-means, SVM, Logistic regression, LDA with cross validation	Identification of two phenotypes for depressive disorder	
Khodayari-Rostamabad et al. [27]	EEG	Prediction of the response to SSRI antidepressant therapy in treatment-resistant MDD	21 MDD subjects	Mixture of factor analysis (MFA)	Specificity: 80.9%, sensitivity: 94.9%, accuracy: 87.9%.	
Whitfield-Gabrieli et al. [28]	Rs-fMRI, dMRI	Prediction of cognitive behavioral therapy results on social anxiety disorder inpatients	38 social anxiety disorder patients	Logistic regression and MVPA, leave-one out cross validation	Better responders vs. worst responders: 81% accuracy, 84% sensitivity and 78% specificity.	
Koutsouleris et al. [29]	sMRI	Classification of SZ vs. MDD	Database 1: 104 MDD, 158 SZ, 437 HC Database 2: 35 BD, 23 FEP, 89 At risk patients	SVM	Classification of MD vs. SZ: 76% of balanced accuracy.	

(continued)

Table 1
(continued)

Study	Modalities	Diseases and goal	Number of subjects	Methods and validation procedure	Results
The IMAGEN Consortium et al. [30]	Stop-signal task, fMRI, personality measure, cognitive test	Prediction of future alcohol-misuse	692 teenagers	ElasticNet logistic regression	94% of correct classification of binge-drinkers and 34% of non binge-drinkers at the optimum of the Receiver Operating Characteristic curve.
Schmaal et al. [31]	Face task and London tower task, fMRI, clinical datas, sMRI	Prediction of MDD prognosis	156 MDD, 145 HC	GPC	fMRI: 73% of correct classification for chronic patients vs. more favorable trajectories patients. 69% of accuracy for chronic vs. remitted patients with clinical data.
Ramyeed et al. [32]	EEG	Prediction of psychosis transition	53 at risk of psychosis inpatients (18 made a transition to psychosis)	LASSO regression	Prediction of psychotic transition: AUC of 0.77
Bertocci et al. [33]	reward-task, fMRI	Prediction of substance abuse in emotionally dysregulated teenagers	30 emotionally dysregulated female teenagers	LASSO regression	Classification of substance abuse vs. no substance abuse: 83.6%

Tran et al. [34]	Electronic medical record (EMR)	Risk stratification of suicidal behaviour	7399 subjects with risk assessment	L1 continuation-ratio model for ordinal outcomes	Classification of high-risk patients versus the rest using EMR data: AUC of 0.79
Mechelli et al. [35]	Clinical data	Psychotic disorders	416 subjects at risk for psychosis	SVM with leave-one-out cross validation	64.6% of accuracy for the prognostic prediction of transition to psychosis
Ramasubbu et al. [36]	Resting-state fMRI, Emotional-faces task fMRI	Classification of different MDD subgroups vs. HC	45 MDD, 19 HC	ICA for feature selection SVM for classification	Very severe group vs. HC with rs-fMRI: 66% accuracy Rest of the classifications at chance levels
Bürger et al. [37]	Emotional faces task fMRI. Features from the anterior cingulate gyrus or the amygdala	Classification of BD vs. HC vs. UD	36 HC, 36 BD, 36 UD	SVM or GPC for the MVPA	Significant results BD vs. UC with a SVM and fearful faces: 69.44%, with GPC: 72%. UC vs. HC with GPC and happy faces: 66.67%.
Salvador et al. [38]	sMRI features: vertex-based cortical thickness, grey and white matter voxels/wavelet-based morphometry, regional volumes and interactions.	Classification of HC vs. SZ vs. BD	127 HC, 128 BD, 128 SZ	Logistic regression, Lasso, Ridge and ElasticNet, SVC, RDA, GPC, Random forest	Best results with VBM. SZ vs. HC: 75%, BD vs. HC: 63% and SZ vs. BD: 62%.
Mwangi et al. [39]	sMRI features	Classification of MDD	30 MDD, 32 HC	SVM and Relevance Vector Machine, train/test split	Classification of 90.3% for MDD vs. HC, correlation of Relevance Vector Machine prediction and illness severity

(continued)

Table 1
(continued)

Study	Modalities	Diseases and goal	Number of subjects	Methods and validation procedure	Results
Mwangi et al. [40]	dMRI features	Prediction of diagnosis in pediatric BD	16 pediatric BD patients, 16 matched HC	SVM with leave-one-out cross-validation and grid search	The SVM algorithm trained with a linear kernel function and FA values performed best with accuracy = 78.12%, sensitivity = 68.75%, specificity = 87.5%
Costafreda et al. [41]	sMRI	Classification of MDD vs. HC and prognosis prediction of CBT and fluoxetine treatment	37 MDD, 37 HC	SVM and leave-one-out cross-validation	The whole brain structural neuroanatomy predicted 88.9% of the clinical response. Accuracy of the structural neuroanatomy as a diagnostic marker was 67.6%.
Xiao et al. [42]	sMRI features	Classification of first episode psychosis patients vs. HC	163 drug naive first episode psychosis patients, 163 HC	SVM and 10-fold cross-validation	Surface Area. Accuracy: 85.0% (specificity = 87.0%, sensitivity = 83.0%) Cortical thickness. Accuracy: 81.8% (specificity = 85.0%, sensitivity = 76.9%)

Abraham et al. for the ENIGMA Bipolar Disorders Working Group et al. [43]	sMRI features	Classification of BD vs. HC	853 BD, 2167 HC	SVM and 30-fold cross-validation or leave-one-site-out	Aggregate subject analysis: accuracy = 65.23% Leave-one-site-out cross-validation: accuracy = 58.67%
Schnack et al. [44]	sMRI features (GMD)	Classification of BD vs. SZ vs. HC	112 SZ, 113 BD and 109 HC	3 SVM models (HC vs. BD, HC vs. SZ and BD vs. SZ) tested on a validation sample	On the validation sample: ROC AUC = 66% for BD vs. SZ, 55.5% for BD vs. HC, 79.2% for SZ vs. HC
Mothi et al. [45]	Clinical measures, eye movement tracks, cognitive assessment, sMRI	Unsupervised clustering of BD, SZ and schizoaffective patients	251 SZ, 164 SZA, 195 BD	K-means with 10 first components of a PCA ran on clinical measures, eye movement, cognitive assessment	3 clusters with one with a majority of SZ, one with a majority of BD, and one distributed across the 3 disorders
Vai et al. [46]	sMRI and dMRI features (FA)	Classification of BD vs. MDD	74 BD, 74 MDD	MVPA with 10-fold cross-validation	Balanced accuracy of 73.65% with a sensitivity for BD of 74.32% and specificity for MDD of 72.97%
Liu et al. [47]	sMRI (VBM of GM or WM)	Classification of treatment resistant depression vs. treatment sensitive depression	17 TRD, 17 TSD, 17 HC	MVPA with leave-one-out cross-validation	GM results: TRD vs. TSD: 82.9%, TRD vs. HC 85.7%, TSD vs. HC 82.4%. WM results: TRD vs. TSD 82.9, TRD vs. HC 85.7%, TSD vs. HC 91.2%
Hajek et al. [48]	sMRI (GM and WM separately)	Classification of unaffected high-risk offspring (UHRO) versus controls (C)	45 UHRO, 45 C	SVM, GPC, leave-2-out cross-validation	SVM with WM for UHRO vs. C: 68.89%, SVM with GM for UHRO vs. C: 56.67%

(continued)

Table 1
(continued)

Study	Modalities	Diseases and goal	Number of subjects	Methods and validation procedure	Results
Fleck et al. [49]	fMRI (continuous performance task with distractions), proton magnetic resonance spectroscopy (PMRS)	Predicting treatment response	20 BD that received lithium treatment	Cascading genetic fuzzy tree	80% accuracy

Acronyms. *AN* anorexia nervosa, *ASD* autism spectrum disorder, *BD* bipolar disorder, *C* controls, *DD* depressive disorder, *dMRI* diffusion MRI, *EA* fractional anisotropy, *FEP* first episode psychosis, *fMRI* functional MRI, *GM* gray matter, *GMD* gray matter density, *GPC* Gaussian process classifier, *HC* healthy controls, *ICA* independent component analysis, *MD* mean diffusivity, *MDD* major depressive disorder, *MRI* magnetic resonance imaging, *MVPA* multivariate pattern analysis, *PCA* principal component analysis, *PMRS* proton magnetic resonance spectroscopy, *rs-fMRI* resting-state functional MRI, *sMRI* structural MRI, *SI* suicide ideation, *SVM* support vector machine, *SZ* schizophrenia, *SZA* schizoaffective, *TRD* treatment-resistant depression, *TSD* treatment-sensitive depression, *UD* unipolar depression, *UHRO* unaffected high-risk offspring, *VBM* voxel-based morphometry, *WM* white matter

were treatment resistant, 17 that were treatment sensitive, and 17 controls. The accuracy of the MVPA models that correctly distinguished resistant and sensitive patients from HC ranged from 85.7% to 91.2% depending on the features used. The authors highlighted differences in structural alterations between responders and non-responders suggesting that structural differences may be related to different responses to antidepressants. Furthermore, they found that the structural abnormalities were larger between responders and HC than between non-responders and HC. These results are somewhat counterintuitive as one would expect resistant patients to show more structural differences from HC than responders. However, this lack of specificity is probably related to a high degree of clinical heterogeneity and the small sample size that does not allow sufficient precision to distinguish more specific abnormalities.

Hajek et al. [48] used machine learning applied to white matter sMRI to distinguish 45 unaffected participants at high genetic risk of BD from 45 low-risk healthy controls with an accuracy of 68.9%. Similarly, Lin et al. [81] successfully classified HR individuals for BD with vs. without (sub)syndromic risk with an accuracy of 83.21% based on the gray matter volume. Finally, a pilot study was conducted using a novel machine learning system based on a “multi-cascade fuzzy genetic tree” with sMRI capable of accurately classifying subjects with BD in a first manic episode into groups that responded or did not respond to lithium treatment [49].

4 Limitations and Perspectives

As illustrated in this chapter, numerous studies have been conducted to classify psychiatric disorders and refine the definition of psychiatric subgroups using machine learning. However, methods and results are heterogeneous. In fact, many authors point to a major limitation of most studies, that is, the limited number of samples [53–55]. Claude et al. [54] also pointed out a negative correlation between the accuracy and the number of subjects, leading to think that the results obtained from small samples are artificially high. Another effect of this limited number of samples resides on the fact that models need to be trained on a population that is representative of the population on which we will use them. Indeed, models trained on a young population will be biased when used on an older one, and similar bias could be raised when using a model trained with a population from a specific country on subjects from another country.

As it is difficult to recruit enough patients to obtain a sufficient statistical power, this limitation may persist in the long term, unless

collective efforts for data sharing are undertaken. This issue deepens when looking at more specific subsets of patients. The field therefore needs more and larger datasets to work on. These datasets start to be collected, with, e.g., the UK BioBank dataset (~40,000 subjects). Even though they are not focused on psychiatric disorders, they are interesting because they are multimodal datasets, with genetic, clinical, and MRI data, and some participants will develop psychiatric syndromes throughout the follow-up. Recent efforts have been specifically made for psychiatric disorders, e.g., by the ENIGMA Consortium, a multisite and multimodal project including several working groups focused on different diseases, such as bipolar disorder, schizophrenia, autism, ADHD, etc.

Larger datasets are often multisite ones, and they bring their own challenges. Since the MRI devices that are used for different studies have different magnetic field strengths, different vendors, coils, etc., there are large site effects that need to be considered. These site effects are particularly important for DWI and fMRI, but they even appear for sMRI [82], the most robust method of imaging. A second source of site effects lies in the preprocessing of the data, which may vary between different sites and protocols. The preprocessing steps are of major importance and need to be homogenized since different softwares can lead to different results [83]. The remaining “site effects” can be partially corrected, thanks to different methods. Statistics-based methods include adjusted residualizations or ComBat [84, 85], a method originally proposed to remove batch effects in genomics [86] and then adapted for DWI and then for sMRI [87]. Other methods are more specific to MRI, such as RAVEL [88], which aims at capturing the sites’ variability using the signal from the CSF, with mixed results for now. Since the extent of the efficiency of these corrections is still under discussion [89], we must consider the site effect in our models and use validation methods such as leave-one-site-out validation to evaluate the reproducibility of our approaches.

The site effect highlights a deeper and more fundamental limitation of our studies, the signal-to-noise ratio. That issue, which is faced by all imaging studies, is particularly present in neuroimaging for psychiatric diseases as the changes that we are looking for are subtle and probably not the main causes of variation in our datasets (e.g., one important cause of variance is the age, which produces consequent variations in the gray and white matter density [72]). We therefore need to be vigilant and make specific efforts when interpreting the results of machine learning algorithms as they can learn some information that are irrelevant for psychiatric disorders. Nevertheless, it is possible to improve this signal-to-noise ratio. One way to do so is to improve the signal; the second is to diminish the noise. Larger datasets improve the statistical power of the algorithms but may induce noise (such as the multisite noise). In addition to the fact that methodological modifications can change

and improve the performance of machine learning, technological improvements seem to bring better performance as shown by the team of Iwabuchi et al. [90], who showed that 7 T MRI compared to 3 T MRI gave higher classification accuracy when distinguishing patients with schizophrenia vs. controls (77% versus 66%). Moreover, the use of multimodal datasets has shown promising results in increasing the signal-to-noise ratio in current studies [91]. While trying to determine to what extent machine learning using MRI can still improve its results, Schulz et al. [91] highlighted two interesting perspectives: first, that there is still room for improvement of the classification accuracy by getting larger datasets and second, that multimodal MRI and more specifically fMRI could improve the classification.

Other ways to collect data could also be thought about, with, for example, the use of tools such as smartphones. Data can be provided through active monitoring (self-reporting), passive monitoring of various activities, mobility, or statistics on phone calls [92]. Promising results show that voice data from daily phone calls could be a valid marker of mood states and hold promise for monitoring BD [93]. Taken together, the development of our knowledge of machine learning and the growing data resources could provide new tools for the management of psychiatric disorders soon. However, their development can only be done by considering the challenges they raise, such as personal data protection, but also by considering all the ethical issues that these new tools will raise.

Finally, machine learning in psychiatry is a promising field of research, with still a lot to do to characterize the different biomarkers and psychiatric disorders properly and accurately. The use of MRI and other clinical and biological features could in a near future bring new tools for diagnosis, risk assessment, and treatment selection that could be used by the clinician. However, due to the actual social stigma around psychiatric disorders and the apparent arbitrary character of classification algorithms, their use would need an important ethical discussion beforehand, notably when people would like to use them to identify at-risk healthy subjects or when using them to determine the treatment of already symptomatic patients.

Acknowledgments

We are grateful to the reviewer, Anton Iftimovici, for his very helpful comments and suggestions.

References

1. Kupfer DJ, Frank E, Phillips ML (2012) Major depressive disorder: new clinical, neurobiological, and treatment perspectives. *Lancet* 379(9820):1045–1055. [https://doi.org/10.1016/S0140-6736\(11\)60602-8](https://doi.org/10.1016/S0140-6736(11)60602-8)
2. Kessler RC, Angermeyer M, Anthony JC et al (2007) Lifetime prevalence and age-of-onset distributions of mental disorders in the World Health Organization's world mental health survey initiative. *World Psychiatry* 6(3): 168–176
3. Diagnostic and statistical manual of mental disorders: DSM-5™, 5th ed. American Psychiatric Publishing, Inc. 2013. p. xlv, 947. <https://doi.org/10.1176/appi.books.9780890425596>
4. Grande I, Berk M, Birmaher B, Vieta E (2016) Bipolar disorder. *Lancet* 387(10027): 1561–1572. [https://doi.org/10.1016/S0140-6736\(15\)00241-X](https://doi.org/10.1016/S0140-6736(15)00241-X)
5. Jablensky A (1997) The 100-year epidemiology of schizophrenia. *Schizophr Res* 28(2–3): 111–125. [https://doi.org/10.1016/S0920-9964\(97\)85354-6](https://doi.org/10.1016/S0920-9964(97)85354-6)
6. Institute of Medicine (US) Committee on Nervous System Disorders in Developing Countries (2001) Neurological, psychiatric, and developmental disorders: meeting the challenge in the developing world. National Academies Press (US). Accessed September 10, 2021. <http://www.ncbi.nlm.nih.gov/books/NBK223475/>
7. Mueser KT, McGurk SR (2004) Schizophrenia. *Lancet* 363(9426):2063–2072. [https://doi.org/10.1016/S0140-6736\(04\)16458-1](https://doi.org/10.1016/S0140-6736(04)16458-1)
8. Fenton WS (1991) Natural history of schizophrenia subtypes: II. Positive and negative symptoms and long-term course. *Arch Gen Psychiatry* 48(11):978. <https://doi.org/10.1001/archpsyc.1991.01810350018003>
9. Sayers SL, Curran PJ, Mueser KT (1996) Factor structure and construct validity of the scale for the assessment of negative symptoms. *Psychol Assess* 8(3):269–280. <https://doi.org/10.1037/1040-3590.8.3.269>
10. Green MF, Kern RS, Braff DL, Mintz J (2000) Neurocognitive deficits and functional outcome in schizophrenia: are we measuring the “right stuff”? *Schizophr Bull* 26(1):119–136. <https://doi.org/10.1093/oxfordjournals.schbul.a033430>
11. McGuffin P, Owen Michael J, Farmer Anne E (1995) Genetic basis of schizophrenia. *Lancet* 346(8976):678–682. [https://doi.org/10.1016/S0140-6736\(95\)92285-7](https://doi.org/10.1016/S0140-6736(95)92285-7)
12. Cardno AG, Marshall EJ, Coid B et al (1999) Heritability estimates for psychotic disorders: the Maudsley twin psychosis series. *Arch Gen Psychiatry* 56(2):162. <https://doi.org/10.1001/archpsyc.56.2.162>
13. Davis JM, Chen N, Glick ID (2003) A meta-analysis of the efficacy of second-generation antipsychotics. *Arch Gen Psychiatry* 60(6): 553. <https://doi.org/10.1001/archpsyc.60.6.553>
14. Insel TR, Cuthbert BN (2015) Brain disorders? Precisely. *Science* 348(6234):499–500. <https://doi.org/10.1126/science.aab2358>
15. Achalia R, Sinha A, Jacob A et al (2020) A proof of concept machine learning analysis using multimodal neuroimaging and neurocognitive measures as predictive biomarker in bipolar disorder. *Asian J Psychiatr* 50:101984. <https://doi.org/10.1016/j.ajp.2020.101984>
16. Lin E, Lin CH, Lai YL, Huang CH, Huang YJ, Lane HY (2018) Combination of G72 genetic variation and G72 protein level to detect schizophrenia: machine learning approaches. *Front Psych* 9:566. <https://doi.org/10.3389/fpsy.2018.00566>
17. Yamada Y, Matsumoto M, Iijima K, Sumiyoshi T (2020) Specificity and continuity of schizophrenia and bipolar disorder: relation to biomarkers. *Curr Pharm Des* 26(2):191–200. <https://doi.org/10.2174/1381612825666191216153508>
18. Grotegerd D, Suslow T, Bauer J et al (2013) Discriminating unipolar and bipolar depression by means of fMRI and pattern classification: a pilot study. *Eur Arch Psychiatry Clin Neurosci* 263(2):119–131. <https://doi.org/10.1007/s00406-012-0329-4>
19. Visser RM, Haver P, Zwitser RJ, Scholte HS, Kindt M (2016) First steps in using multi-voxel pattern analysis to disentangle neural processes underlying generalization of spider fear. *Front Hum Neurosci* 10. <https://doi.org/10.3389/fnhum.2016.00222>
20. Lavagnino L, Amianto F, Mwangi B et al (2015) Identifying neuroanatomical signatures of anorexia nervosa: a multivariate machine learning approach. *Psychol Med* 45(13): 2805–2812. <https://doi.org/10.1017/S0033291715000768>
21. Fu CHY, Mourao-Miranda J, Costafreda SG et al (2008) Pattern classification of sad facial processing: toward the development of neurobiological markers in depression. *Biol Psychiatry* 63(7):656–662. <https://doi.org/10.1016/j.biopsych.2007.08.020>

22. Setoyama D, Kato TA, Hashimoto R et al (2016) Plasma metabolites predict severity of depression and suicidal ideation in psychiatric patients – a multicenter pilot analysis. Hashimoto K, ed. *PLoS One* 11(12): e0165267. <https://doi.org/10.1371/journal.pone.0165267>
23. Lueken U, Straube B, Yang Y et al (2015) Separating depressive comorbidity from panic disorder: a combined functional magnetic resonance imaging and machine learning approach. *J Affect Disord* 184:182–192. <https://doi.org/10.1016/j.jad.2015.05.052>
24. Pettersson-Yeo W, Benetti S, Marquand AF et al (2013) Using genetic, cognitive and multi-modal neuroimaging data to identify ultra-high-risk and first-episode psychosis at the individual level. *Psychol Med* 43(12): 2547–2562. <https://doi.org/10.1017/S003329171300024X>
25. Wu MJ, Mwangi B, Bauer IE et al (2017) Identification and individualized prediction of clinical phenotypes in bipolar disorders using neurocognitive data, neuroimaging scans and machine learning. *NeuroImage* 145:254–264. <https://doi.org/10.1016/j.neuroimage.2016.02.016>
26. Drysdale AT, Grosenick L, Downar J et al (2017) Resting-state connectivity biomarkers define neurophysiological subtypes of depression. *Nat Med* 23(1):28–38. <https://doi.org/10.1038/nm.4246>
27. Khodayari-Rostamabad A, Reilly JP, Hasey GM, de Bruin H, MacCrimmon DJ (2013) A machine learning approach using EEG data to predict response to SSRI treatment for major depressive disorder. *Clin Neurophysiol* 124(10):1975–1985. <https://doi.org/10.1016/j.clinph.2013.04.010>
28. Whitfield-Gabrieli S, Ghosh SS, Nieto-Castanon A et al (2016) Brain connectomics predict response to treatment in social anxiety disorder. *Mol Psychiatry* 21(5):680–685. <https://doi.org/10.1038/mp.2015.109>
29. Koutsouleris N, Meisenzahl EM, Borgwardt S et al (2015) Individualized differential diagnosis of schizophrenia and mood disorders using neuroanatomical biomarkers. *Brain* 138(7): 2059–2073. <https://doi.org/10.1093/brain/awv111>
30. The IMAGEN Consortium, Whelan R, Watts R et al (2014) Neuropsychosocial profiles of current and future adolescent alcohol misusers. *Nature* 512(7513):185–189. <https://doi.org/10.1038/nature13402>
31. Schmaal L, Marquand AF, Rhebergen D et al (2015) Predicting the naturalistic course of major depressive disorder using clinical and multimodal neuroimaging information: a multivariate pattern recognition study. *Biol Psychiatry* 78(4):278–286. <https://doi.org/10.1016/j.biopsych.2014.11.018>
32. Ramyeed A, Studerus E, Kometer M et al (2016) Prediction of psychosis using neural oscillations and machine learning in neuroleptic-naïve at-risk patients. *World J Biol Psychiatry* 17(4):285–295. <https://doi.org/10.3109/15622975.2015.1083614>
33. Bertocci MA, Bebko G, Versace A et al (2017) Reward-related neural activity and structure predict future substance use in dysregulated youth. *Psychol Med* 47(8):1357–1369. <https://doi.org/10.1017/S0033291716003147>
34. Tran T, Luo W, Phung D et al (2014) Risk stratification using data from electronic medical records better predicts suicide risks than clinician assessments. *BMC Psychiatry* 14(1):76. <https://doi.org/10.1186/1471-244X-14-76>
35. Mechelli A, Lin A, Wood S et al (2017) Using clinical information to make individualized prognostic predictions in people at ultra high risk for psychosis. *Schizophr Res* 184:32–38. <https://doi.org/10.1016/j.schres.2016.11.047>
36. Ramasubbu R, Brown MRG, Cortese F et al (2016) Accuracy of automated classification of major depressive disorder as a function of symptom severity. *Neuroimage Clin* 12:320–331. <https://doi.org/10.1016/j.nicl.2016.07.012>
37. Bürger C, Redlich R, Grotegerd D et al (2017) Differential abnormal pattern of anterior cingulate gyrus activation in unipolar and bipolar depression: an fMRI and pattern classification approach. *Neuropsychopharmacology* 42(7): 1399–1408. <https://doi.org/10.1038/npp.2017.36>
38. Salvador R, Radua J, Canales-Rodríguez EJ et al (2017) Evaluation of machine learning algorithms and structural features for optimal MRI-based diagnostic prediction in psychosis. Hu D, ed. *PLoS One* 12(4):e0175683. <https://doi.org/10.1371/journal.pone.0175683>
39. Mwangi B, Ebmeier KP, Matthews K, Douglas SJ (2012) Multi-centre diagnostic classification of individual structural neuroimaging scans from patients with major depressive disorder. *Brain* 135(5):1508–1521. <https://doi.org/10.1093/brain/aws084>
40. Mwangi B, Wu MJ, Bauer IE et al (2015) Predictive classification of pediatric bipolar disorder using atlas-based diffusion weighted imaging and support vector machines. *Psychiatry Res Neuroimaging* 234(2):265–271.

- <https://doi.org/10.1016/j.psychres.2015.10.002>
41. Costafreda SG, Chu C, Ashburner J, Fu CHY (2009) Prognostic and diagnostic potential of the structural neuroanatomy of depression. *PLoS One* 4(7):e6353. <https://doi.org/10.1371/journal.pone.0006353>
 42. Xiao Y, Yan Z, Zhao Y et al (2019) Support vector machine-based classification of first episode drug-naïve schizophrenia patients and healthy controls using structural MRI. *Schizophr Res* 214:11–17. <https://doi.org/10.1016/j.schres.2017.11.037>
 43. ENIGMA Bipolar Disorders Working Group, Nunes A, Schnack HG et al (2020) Using structural MRI to identify bipolar disorders – 13 site machine learning study in 3020 individuals from the ENIGMA bipolar disorders working group. *Mol Psychiatry* 25(9): 2130–2143. <https://doi.org/10.1038/s41380-018-0228-9>
 44. Schnack HG, Nieuwenhuis M, van Haren NEM et al (2014) Can structural MRI aid in clinical classification? A machine learning study in two independent samples of patients with schizophrenia, bipolar disorder and healthy subjects. *NeuroImage* 84:299–306. <https://doi.org/10.1016/j.neuroimage.2013.08.053>
 45. Mothi SS, Sudarshan M, Tandon N et al (2019) Machine learning improved classification of psychoses using clinical and biological stratification: update from the bipolar-schizophrenia network for intermediate phenotypes (B-SNIP). *Schizophr Res* 214:60–69. <https://doi.org/10.1016/j.schres.2018.04.037>
 46. Vai B, Parenti L, Bollettini I et al (2020) Predicting differential diagnosis between bipolar and unipolar depression with multiple kernel learning on multimodal structural neuroimaging. *Eur Neuropsychopharmacol* 34:28–38. <https://doi.org/10.1016/j.euroneuro.2020.03.008>
 47. Liu F, Guo W, Yu D et al (2012) Classification of different therapeutic responses of major depressive disorder with multivariate pattern analysis method based on structural MR scans. Fan Y, ed. *PLoS One* 7(7):e40968. <https://doi.org/10.1371/journal.pone.0040968>
 48. Hajek T, Cooke C, Kopecek M, Novak T, Hoschl C, Alda M (2015) Using structural MRI to identify individuals at genetic risk for bipolar disorders: a 2-cohort, machine learning study. *J Psychiatry Neurosci* 40(5):316–324. <https://doi.org/10.1503/jpn.140142>
 49. Fleck DE, Ernest N, Adler CM et al (2017) Prediction of lithium response in first-episode mania using the LITHium intelligent agent (LITHIA): pilot data and proof-of-concept. *Bipolar Disord* 19(4):259–272. <https://doi.org/10.1111/bdi.12507>
 50. Miller PR, Dasher R, Collins R, Griffiths P, Brown F (2001) Inpatient diagnostic assessments: 1. Accuracy of structured vs. unstructured interviews. *Psychiatry Res* 105:255
 51. ECNP consensus meeting. Bipolar depression. Nice, March 2007 <https://doi.org/10.1016/j.euroneuro.2008.03.003>. Published September 8, 2021
 52. Alda M (2021) The moving target of psychiatric diagnosis. *J Psychiatry Neurosci* 46(3): E415–E417. <https://doi.org/10.1503/jpn.210098>
 53. Arbabshirani MR, Plis S, Sui J, Calhoun VD (2017) Single subject prediction of brain disorders in neuroimaging: promises and pitfalls. *NeuroImage* 145:137–165. <https://doi.org/10.1016/j.neuroimage.2016.02.079>
 54. Claude L, Houenou J, Duchesnay E, Favre P (2020) Will machine learning applied to neuroimaging in bipolar disorder help the clinician? A critical review and methodological suggestions. *Bipolar Disord* 22(4):334–355. <https://doi.org/10.1111/bdi.12895>
 55. Gao S, Calhoun VD, Sui J (2018) Machine learning in major depression: from classification to treatment outcome prediction. *CNS Neurosci Ther* 24(11):1037–1052. <https://doi.org/10.1111/cns.13048>
 56. de Filippis R, Carbone EA, Gaetano R et al (2019) Machine learning techniques in a structural and functional MRI diagnostic approach in schizophrenia: a systematic review. *Neuropsychiatr Dis Treat* 15:1605–1627. <https://doi.org/10.2147/NDT.S202418>
 57. Kruschwitz JD, Ludwig VU, Waller L et al (2018) Regulating craving by anticipating positive and negative outcomes: a multivariate pattern analysis and network connectivity approach. *Front Behav Neurosci* 12:297. <https://doi.org/10.3389/fnbeh.2018.00297>
 58. Kliemann D, Richardson H, Anzellotti S et al (2018) Cortical responses to dynamic emotional facial expressions generalize across stimuli, and are sensitive to task-relevance, in adults with and without autism. *Cortex* 103: 24–43. <https://doi.org/10.1016/j.cortex.2018.02.006>
 59. Barros C, Silva CA, Pinheiro AP (2021) Advanced EEG-based learning approaches to predict schizophrenia: promises and pitfalls. *Artif Intell Med* 114:102039. <https://doi.org/10.1016/j.artmed.2021.102039>

60. Bzdok D, Ioannidis JPA (2019) Exploration, inference, and prediction in neuroscience and biomedicine. *Trends Neurosci* 42(4):251–262. <https://doi.org/10.1016/j.tins.2019.02.001>
61. Haufe S, Meinecke F, Görgen K et al (2014) On the interpretation of weight vectors of linear models in multivariate neuroimaging. *NeuroImage* 87:96–110. <https://doi.org/10.1016/j.neuroimage.2013.10.067>
62. Bowden CL (2001) Strategies to reduce misdiagnosis of bipolar depression. *Psychiatr Serv* 52(1):51–55. <https://doi.org/10.1176/appi.ps.52.1.51>
63. Galimberti C, Bosi MF, Volontè M, Giordano F, Dell’Osso B, Viganò CA (2020) Duration of untreated illness and depression severity are associated with cognitive impairment in mood disorders. *Int J Psychiatry Clin Pract* 24(3):227–235. <https://doi.org/10.1080/13651501.2020.1757116>
64. Arnedo J, Svrakic DM, del Val C et al (2015) Uncovering the hidden risk architecture of the schizophrenias: confirmation in three independent genome-wide association studies. *Am J Psychiatry* 172(2):139–153. <https://doi.org/10.1176/appi.ajp.2014.14040435>
65. Schnack HG (2019) Improving individual predictions: machine learning approaches for detecting and attacking heterogeneity in schizophrenia (and other psychiatric diseases). *Schizophr Res* 214:34–42. <https://doi.org/10.1016/j.schres.2017.10.023>
66. Varol E, Sotiras A, Davatzikos C (2017) HYDRA: revealing heterogeneity of imaging and genetic patterns through a multiple max-margin discriminative analysis framework. *NeuroImage* 145:346–364. <https://doi.org/10.1016/j.neuroimage.2016.02.041>
67. Schulz MA, Chapman-Rounds M, Verma M, Bzdok D, Georgatzis K (2020) Inferring disease subtypes from clusters in explanation space. *Sci Rep* 10(1):12900. <https://doi.org/10.1038/s41598-020-68858-7>
68. Yang T, Frangou S, Lam RW et al (2021) Probing the clinical and brain structural boundaries of bipolar and major depressive disorder. *Transl Psychiatry* 11:48. <https://doi.org/10.1038/s41398-020-01169-7>
69. Louiset R, Gori P, Dufumier B, Houenou J, Grigis A, Duchesnay E. UCSL: a machine learning expectation-maximization framework for unsupervised clustering driven by supervised learning. *ArXiv210701988 Cs Stat*. Published online July 5, 2021. Accessed March 1, 2022. <http://arxiv.org/abs/2107.01988>
70. Wolfers T, Doan NT, Kaufmann T et al (2018) Mapping the heterogeneous phenotype of schizophrenia and bipolar disorder using normative models. *JAMA Psychiatry* 75(11):1146. <https://doi.org/10.1001/jamapsychiatry.2018.2467>
71. Marquand AF, Rezek I, Buitelaar J, Beckmann CF (2016) Understanding heterogeneity in clinical cohorts using normative models: beyond case-control studies. *Biol Psychiatry* 80(7):552–561. <https://doi.org/10.1016/j.biopsych.2015.12.023>
72. Bethlehem RAI, Seidlitz J, White SR et al (2022) Brain charts for the human lifespan. *Nature* 604(7906):525–533. <https://doi.org/10.1038/s41586-022-04554-y>
73. Kessler RC, van Loo HM, Wardenaar KJ et al (2016) Testing a machine-learning algorithm to predict the persistence and severity of major depressive disorder from baseline self-reports. *Mol Psychiatry* 21(10):1366–1371. <https://doi.org/10.1038/mp.2015.198>
74. Suvisaari J, Mantere O, Keinänen J et al (2018) Is it possible to predict the future in first-episode psychosis? *Front Psych* 9:580. <https://doi.org/10.3389/fpsyg.2018.00580>
75. Sun H, Jiang R, Qi S et al (2019) Preliminary prediction of individual response to electroconvulsive therapy using whole-brain functional magnetic resonance imaging data. *Neuroimage Clin* 26:102080. <https://doi.org/10.1016/j.nicl.2019.102080>
76. Hajek T, Cullis J, Novak T et al (2013) Brain structural signature of familial predisposition for bipolar disorder: replicable evidence for involvement of the right inferior frontal gyrus. *Biol Psychiatry* 73(2):144–152. <https://doi.org/10.1016/j.biopsych.2012.06.015>
77. Hajek T, Kopecek M, Kozeny J, Gunde E, Alda M, Höschl C (2009) Amygdala volumes in mood disorders — meta-analysis of magnetic resonance volumetry studies. *J Affect Disord* 115(3):395–410. <https://doi.org/10.1016/j.jad.2008.10.007>
78. Ganzola R, Duchesne S (2017) Voxel-based morphometry meta-analysis of gray and white matter finds significant areas of differences in bipolar patients from healthy controls. *Bipolar Disord* 19(2):74–83. <https://doi.org/10.1111/bdi.12488>
79. Pinaya WHL, Gadelha A, Doyle OM et al (2016) Using deep belief network modelling to characterize differences in brain morphometry in schizophrenia. *Sci Rep* 6(1):38897. <https://doi.org/10.1038/srep38897>
80. Hibar DP, Westlye LT, Doan NT et al (2018) Cortical abnormalities in bipolar disorder: an MRI analysis of 6503 individuals from the ENIGMA bipolar disorder working group.

- Mol Psychiatry 23(4):932–942. <https://doi.org/10.1038/mp.2017.73>
81. Lin K et al (2018) Illness, at-risk and resilience neural markers of early-stage bipolar disorder. *J Affect Disord* 238:16–23
 82. Keenan KE, Gimbutas Z, Dienstfrey A et al (2021) Multi-site, multi-platform comparison of MRI T1 measurement using the system phantom. Lundberg P, ed. *PLoS One* 16(6): e0252966. <https://doi.org/10.1371/journal.pone.0252966>
 83. Yamada H, Abe O, Shizukuishi T et al (2014) Efficacy of distortion correction on diffusion imaging: comparison of FSL Eddy and Eddy_Correct using 30 and 60 directions diffusion encoding. Najbauer J, ed. *PLoS One* 9(11): e112411. <https://doi.org/10.1371/journal.pone.0112411>
 84. Fortin JP, Cullen N, Sheline YI et al (2018) Harmonization of cortical thickness measurements across scanners and sites. *NeuroImage* 167:104–120. <https://doi.org/10.1016/j.neuroimage.2017.11.024>
 85. Fortin JP, Parker D, Tunc B et al (2017) Harmonization of multi-site diffusion tensor imaging data. *Neuroimage* 161:149–170
 86. Johnson WE, Li C, Rabinovic A (2007) Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* 8(1):118–127. <https://doi.org/10.1093/biostatistics/kxj037>
 87. Pomponio R, Erus G, Habes M et al (2020) Harmonization of large MRI datasets for the analysis of brain imaging patterns throughout the lifespan. *NeuroImage* 208:116450. <https://doi.org/10.1016/j.neuroimage.2019.116450>
 88. Eshaghzadeh Torbati M, Minhas DS, Ahmad G et al (2021) A multi-scanner neuroimaging data harmonization using RAVEL and ComBat. *NeuroImage* 245:118703. <https://doi.org/10.1016/j.neuroimage.2021.118703>
 89. Cetin-Karayumak S, Stegmayer K, Walther S et al (2020) Exploring the limits of ComBat method for multi-site diffusion MRI harmonization. *Neuroscience*. <https://doi.org/10.1101/2020.11.20.390120>
 90. Iwabuchi SJ, Liddle PF, Palaniyappan L (2013) Clinical utility of machine-learning approaches in schizophrenia: improving diagnostic confidence for translational neuroimaging. *Front Psych* 4:95. <https://doi.org/10.3389/fpsy.2013.00095>
 91. Schulz MA, Bzdok D, Haufe S, Haynes JD, Ritter K (2022) Performance reserves in brain-imaging-based phenotype prediction. *Neuroscience*. <https://doi.org/10.1101/2022.02.23.481601>
 92. Antosik-Wójcińska AZ, Dominiak M, Chojnacka M et al (2020) Smartphone as a monitoring tool for bipolar disorder: a systematic review including data analysis, machine learning algorithms and predictive modelling. *Int J Med Inform* 138:104131. <https://doi.org/10.1016/j.ijmedinf.2020.104131>
 93. Faurholt-Jepsen M, Bauer M, Kessing LV (2018) Smartphone-based objective monitoring in bipolar disorder: status and considerations. *Int J Bipolar Disord* 6(1):6. <https://doi.org/10.1186/s40345-017-0110-8>

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

