Machine Learning Classification of Schizophrenia Based on Brain Structural Imaging

Authors: Zheng Hong, Pu Chengcheng, Wang Yi, Chen Chuqiao, Wang Yi

Date: 2019-09-28T00:00:00+00:00

Abstract

The application of machine learning to clinical and basic research on psychiatric disorders has emerged as a prominent trend in recent years. Researchers have applied machine learning to brain imaging data, including T1-weighted imaging and diffusion tensor imaging, from patients with schizophrenia and high-risk populations, thereby contributing to the elucidation of disease pathophysiological mechanisms. A review of previous studies reveals that brain structural features of the frontal and temporal lobes demonstrate high discriminative capacity, and classification performance achieved through the integration of behavioral and neuroimaging data surpasses that of single-modality data. Current research is constrained by limitations including insufficient sample sizes and inadequate generalizability; future investigations should emphasize expanding sample sizes and establishing standardized classification methodologies to further explore the role of machine learning in psychiatric disorders.

Full Text

The Classification of Schizophrenia Based on Brain Structural Features: A Machine Learning Approach

Hong Zheng[1,2]; Cheng-cheng Pu[3]; Yi Wang[1,2]; Raymond C. K. Chan[1,2]

[1]Neuropsychology and Applied Cognitive Neuroscience Laboratory, CAS Key Laboratory of Mental Health, Institute of Psychology, Chinese Academy of Sciences, Beijing 100101, China
[2]Department of Psychology, University of Chinese Academy of Sciences, Beijing 100049, China
[3]Peking University Sixth Hospital, Peking University Institute of Mental Health, NHC Key Laboratory of Mental Health (Peking University), National Clinical

Research Center for Mental Disorders (Peking University Sixth Hospital), Beijing 100191, China

Abstract

Applying machine learning to clinical and basic research on mental disorders has become an emerging trend. Researchers have applied machine learning to T1-weighted imaging and diffusion tensor imaging data from schizophrenia patients and high-risk populations, providing insights into the pathophysiological mechanisms of the disease. A review of previous studies reveals that structural features of the frontal and temporal lobes demonstrate high discriminative power, and classification performance is superior when combining behavioral data with brain imaging data compared to single-modality data alone. Current research is limited by insufficient sample sizes and inadequate generalization ability. Future studies should focus on expanding sample sizes and developing standardized classification methods to further explore the role of machine learning in mental disorders.

Keywords: structural Magnetic Resonance Imaging; Diffusion Tensor Imaging; machine learning; schizophrenia; high-risk population

Schizophrenia is a severe mental disorder with a prevalence of approximately 1%. Patients exhibit bizarre thought patterns and behaviors, distorted perception, cognitive dysfunction, and often present with features such as emotional blunting and social withdrawal (Carpenter & Buchanan, 1994). In recent years, neuroimaging studies have revealed structural brain abnormalities in schizophrenia patients, including alterations in gray matter volume and density, white matter abnormalities, and enlarged ventricles, providing support for understanding the pathological mechanisms of the disorder (Bakhshi & Chance, 2015; Palaniyappan, 2017). Current schizophrenia neuroimaging research predominantly employs group mean difference tests between clinical patients and healthy controls to examine abnormal features. However, these results cannot be applied at the individual level, a limitation that machine learning technology can address. The concept of machine learning was first proposed by Arthur Samuel of IBM in 1959. Tom Mitchell later provided a broader and more formal definition: "A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P if its performance at tasks in T, as measured by P, improves with experience E" (Mitchell, 1997). In the field of mental illness, researchers utilize machine learning to identify patterns from neuroimaging data, construct models, and make predictions at the individual level while identifying the most discriminative features.

Researchers have already employed machine learning classification methods to identify schizophrenia patients from the general population (Ardekani et al., 2011; Nieuwenhuis et al., 2012; Winterburn et al., 2017), predict which individuals at high risk will convert to schizophrenia (Zarogianni, Storkey, Johnstone, Owens, & Lawrie, 2017), and differentiate patients at various disease

stages (Guo, Palaniyappan, Liddle, & Feng, 2016). During model training, selected features are typically those that can significantly distinguish between categories. Therefore, machine learning classification studies of schizophrenia patients, high-risk populations, and healthy controls can facilitate the identification of biomarkers relevant to disease diagnosis, prognosis, and treatment efficacy.

This article systematically reviews the literature from four perspectives: first, we briefly outline the main steps of machine learning; second, we systematically review previous machine learning studies using structural brain imaging in schizophrenia patients and high-risk populations; third, we discuss current limitations in schizophrenia machine learning research; and finally, we propose future research directions.

1. Main Steps of Machine Learning

When applying machine learning to structural Magnetic Resonance Imaging (sMRI) data in schizophrenia, the primary steps include feature extraction from brain images, feature dimensionality reduction and selection, model training, model evaluation, and prediction of new samples using the trained model (see Figure 1 [Figure 1: see original paper]).

Figure 1. Machine learning analysis steps for brain imaging data

1.1 Data Extraction

In schizophrenia sMRI studies, high signal-to-noise ratio brain images are fundamental to analysis. Researchers must perform image quality control to ensure the reliability and validity of data used for machine learning. Following data extraction from brain images, the dataset is divided into training and test sets, or training, validation, and test sets. Researchers use the training set to identify patterns and generate models, the validation set to adjust model parameters and select optimal models, and the final test set to evaluate model performance.

1.2 Feature Dimensionality Reduction and Selection

Feature dimensionality reduction and selection effectively improve model generalization ability and constitute crucial steps in data processing, particularly for high-dimensional data. Dimensionality reduction transforms high-dimensional data into low-dimensional representations, with Principal Components Analysis (PCA) being a commonly used method. Feature selection involves choosing a subset of features with optimal classification performance from original or reduced data, where the class distribution of the selected subset should resemble that of the original data. Guyon, Weston, Barnhill, and Vapnik (2002) categorized feature selection methods into filter, wrapper, and embedded approaches based on evaluation criteria, all of which have been applied in schizophrenia neuroimaging classification studies. Filter methods score features based on sta-

tistical metrics or correlations with class labels, selecting features according to predetermined thresholds or numbers, which helps reduce inter-feature correlations while enhancing feature-class correlations. Wrapper methods add or remove features based on algorithmic classification performance, such as using Support Vector Machine-Recursive Feature Elimination (SVM-RFE) for feature selection. Embedded methods integrate feature selection with classifier training, automatically selecting optimal feature subsets during training; for example, decision tree algorithms prioritize features with stronger classification abilities during training. Additionally, researchers can select or remove features based on previous literature, such as excluding imaging data from the striatum, which is heavily affected by medication in schizophrenia patients (Nieuwenhuis et al., 2012).

1.3 Model Training

Researchers employ machine learning algorithms to identify patterns in neuroimaging data and generate models. Machine learning algorithms are generally categorized into supervised and unsupervised learning, depending on whether they utilize class labels or target values from sample data. In schizophrenia research, supervised learning algorithms such as Support Vector Machine (SVM) and Linear Regression are commonly used to investigate classification between schizophrenia patients or high-risk populations and healthy controls.

1.4 Model Evaluation and Prediction of New Data

Generalization ability refers to a model' s applicability to new data, which researchers measure using performance metrics to evaluate model quality (Zhou, 2016). In schizophrenia research, the most intuitive performance metrics include Accuracy, Sensitivity (Recall), Precision, and Specificity. Accuracy represents the percentage of correctly classified samples among the total sample. Sensitivity refers to the proportion of true positive samples among all positive samples. Precision indicates the percentage of true positive samples among those classified as positive. Specificity represents the proportion of true negative samples among all negative samples. Based on performance metrics, researchers may modify algorithms or adjust model parameters.

1.5 Cross-Validation

Cross-validation is an effective method for evaluating model performance during feature selection, parameter tuning, and final model testing. Data processing operations (such as dimensionality reduction and feature selection) must be embedded within cross-validation. K-fold cross-validation (K-fold CV) is the most common approach, dividing samples into K folds, using K-1 folds as training data and the remaining fold as test data, with average accuracy serving as the final prediction result. Most studies employ 5-fold or 10-fold cross-validation, though K values can also be determined based on sample size and statistical criteria (Dwyer, Falkai, & Koutsouleris, 2018). Due to sample size limitations,

neuroimaging machine learning studies often use a special form of K-fold cross-validation—Leave-One-Out Cross-Validation (LOOCV)—where K equals the total number of samples N. While simple, LOOCV suffers from high bias and computational expense.

Additionally, some researchers recommend using Nested Cross-Validation (Madsen, Krohne, Cai, Wang, & Chan, 2018). Nested cross-validation includes an outer loop for evaluating model performance and an inner loop for selecting optimal features and parameters, enabling small datasets to use different data for model evaluation and feature/parameter selection. This approach minimizes bias, alleviates issues with small sample availability, and improves result credibility. Multiple neuroimaging machine learning studies have already employed nested cross-validation methods (Borgwardt et al., 2013; Dyrba, Grothe, Kirste, & Teipel, 2015; Peng et al., 2017; Zarogianni et al., 2017).

2. Classification Studies of Schizophrenia Patients and Healthy Controls Based on Brain Structural Features

We conducted a literature search for machine learning studies in schizophrenia since 2010 using structural imaging data (including T1-weighted imaging and Diffusion Tensor Imaging [DTI]). Databases searched included Web of Science, PubMed, and Elsevier, using keyword combinations such as "structural MRI," "Machine Learning," "Support Vector Machine," SVM, schizophrenia, "high risk," schizotype, "voxel based morphometry," VBM, "grey matter," "gray matter," DTI, and "diffusion tensor imaging." Table 1 summarizes the main methods and results from 34 studies identified as of January 17, 2019.

Based on whether classification features originate from the same modality, neuroimaging data are divided into single-modality and multi-modality data. To investigate the impact of these two data types and different brain structural features on machine learning classification of schizophrenia neuroimaging data, the following review is organized according to: single-modality machine learning studies based on structural MRI, single-modality machine learning studies based on DTI, multi-modality machine learning studies, and machine learning studies combining brain imaging with behavioral data.

2.1 Single-Modality Machine Learning Studies Based on Structural MRI

In machine learning studies based on structural MRI, the most commonly used features include gray matter volume, white matter volume, and cortical thickness. Studies classifying chronic schizophrenia patients, first-episode schizophrenia patients, and individuals at high risk for psychosis have all been reported. In terms of classification performance, reported accuracies range from 36.84% to 98%. Regarding features, structural characteristics of the frontal and temporal lobes can achieve high classification accuracy.

2.1.1 Schizophrenia In classification studies of schizophrenia, total sample sizes range from 39 to 606, with most studies exceeding 100 participants. Classification accuracies range from 44.74% to 98%. Results indicate that structural features of the superior frontal gyrus, middle frontal gyrus, inferior frontal gyrus, middle temporal gyrus, superior temporal gyrus, fusiform gyrus, hippocampus, precuneus, thalamus, and occipital lobe yield high classification accuracy. The impact of algorithms on classification accuracy is not significant, though feature selection methods may affect performance, and the importance of the same features may differ across patients with varying disease courses.

Brain structure in schizophrenia patients changes with disease progression. One study analyzed cortical thickness data from 98 schizophrenia patients with different disease courses and 83 healthy controls, removing age and gender effects on cortical thickness. Using SVM with an RBF (Radial Basis Function) kernel for classification and LOOCV for validation, they achieved an average classification accuracy of 81.77%, average sensitivity of 77.55%, and average specificity of 86.75%. Further analysis revealed that classification accuracy between patients with less than two years of disease course and healthy controls was highest (96.3%, with 98.8% specificity and 88% sensitivity). Patients, particularly those with long disease courses, showed reduced cortical thickness in the temporal-limbic and frontal-parietal systems, but increased cortical thickness in the occipital lobe (Guo et al., 2016).

During classification, some researchers have proposed that using feature selection methods to select gray matter density from partial brain regions yields higher accuracy than using whole-brain gray matter density data. Chin, You, Meng, Zhou, and Sim (2018) used a sequential region-of-interest (ROI) selection method to select optimal features from gray matter density data across 64 brain regions. Results showed this approach increased classification accuracy from 86.6% to 92.0% for 84 schizophrenia patients and 43 healthy controls (training set), and from 83.5% to 89.4% for 57 schizophrenia patients and 28 healthy controls (test set). The method also identified an optimal classification subset containing 7 ROIs: fusiform gyrus, middle frontal gyrus, inferior frontal gyrus, superior temporal gyrus, superior frontal gyrus, left thalamus, and left ventricle. Other researchers have compared the effects of different feature selection methods on classification accuracy. Nieuwenhuis et al. (2012) used Voxel-Based Morphometry (VBM) to extract gray matter density data for each voxel from structural images, removing age, gender, and handedness effects using linear regression, and using residuals as original input data. They employed both embedded methods and prior knowledge-based feature selection. The embedded method used SVM to train models, selecting features in the top 10% of absolute weight values as input features. The prior knowledge method removed gray matter density from the striatum (affected by medication) based on previous research, using remaining brain regions as input features. Using linear SVM to build models in 239 subjects (128 patients and 111 healthy controls) and LOOCV, results showed the embedded feature selection method achieved better classification performance: 71.4% average accuracy, 73.4% average sen-

sitivity, and 69.4% average specificity. Validation in 277 subjects (155 patients and 122 healthy controls) yielded 70.4% classification accuracy, 67.1% sensitivity, and 73.8% specificity. The study found reduced gray matter density in the frontal lobe, superior temporal gyrus, and hippocampus, but increased gray matter density in the basal ganglia and left occipital lobe in patients.

In addition to using different feature selection methods, researchers have also investigated the impact of different algorithms on classification accuracy. Salvador et al. (2017) used gray matter volume data from 127 healthy controls and 128 schizophrenia patients to compare classification performance across different algorithms, including Ridge, Lasso, Elastic Net, L0 Norm Regularized Logistic Regression, linear SVM, Regularized Discriminant Analysis, Random Forests, and Gaussian Process Classifier. Results showed no substantial differences in classification outcomes across these eight algorithms, with accuracies ranging from 74.1% to 76%.

Existing research has applied machine learning not only to distinguish schizophrenia patients from healthy controls but also to differentiate subtypes of schizophrenia patients. One study used machine learning to distinguish schizophrenia patients with and without obsessive-compulsive symptoms, finding that gray matter density in the paracentral lobule (including the supplementary motor area), middle cingulate gyrus, and fronto-subcortical regions could effectively differentiate the two groups with 78.26% accuracy in a sample of 23 patients without obsessive-compulsive symptoms and 23 with obsessive-compulsive symptoms (Tas et al., 2018).

2.1.2 First-Episode Schizophrenia Due to potential brain structural changes from long-term disease course and medication, chronic schizophrenia patients may show altered classification performance. First-episode schizophrenia patients are less affected by medication and disease progression, making studies of this population valuable for investigating the neuropathological mechanisms of schizophrenia. In classification studies of first-episode schizophrenia, total sample sizes range from 42 to 480, with classification accuracies between 44.74% and 96.7%. Regarding features, structural characteristics of the frontal lobe (such as superior frontal gyrus, middle frontal gyrus, and frontal pole), middle temporal gyrus, fusiform gyrus, and lingual gyrus can improve classification accuracy.

Classification accuracy is a primary concern for researchers. In a study of 127 first-episode patients and 127 healthy controls, Squarcina et al. (2017) extracted cortical thickness data from 68 ROIs, applying SVM to each ROI individually and selecting the 5 optimal features based on classification results. Using these optimal features to build models and training/predicting with SVM or Multiple Kernel Learning (MKL) algorithms, they achieved a maximum accuracy of 85% in the superior frontal gyrus, rostral middle frontal gyrus, and middle temporal gyrus.

Some researchers have investigated classification performance of different algorithms applied to gray matter density and cortical thickness data in first-episode schizophrenia patients. Winterburn et al. (2017) used Logistic Regression (LR), SVM (including linear SVM and SVM-RBF), and Linear Discriminant Analysis (LDA) for classification. Brain structural data included gray matter density and cortical thickness from 50 first-episode schizophrenia patients and 50 healthy controls. Researchers first divided samples into training (2/3) and test (1/3) sets. When applying linear SVM to gray matter density data, maximum classification accuracy reached 69.6% (sensitivity 77.8%, specificity 57.9%). Using SVM-RBF algorithm with cortical thickness data achieved a maximum accuracy of 73.5% (sensitivity 58.8%, specificity 88.2%).

2.1.3 High-Risk Populations Compared to first-episode schizophrenia patients, identifying high-risk populations and investigating their brain structure helps explain the mechanisms underlying schizophrenia onset and predict illness development. Currently, few classification studies of high-risk populations using structural MRI exist, with total sample sizes not exceeding 50 and classification accuracies ranging from 36.84% to 72%. One study used gray matter volume data and SVM to classify 25 high-risk individuals and 25 healthy controls, applying LOOCV and achieving 72% classification accuracy (sensitivity 68%, specificity 76%). Brain regions with high feature weights included bilateral hippocampus, parahippocampal gyrus, putamen, superior frontal gyrus, middle frontal gyrus, fusiform gyrus, and inferior parietal lobule, as well as left inferior temporal gyrus, right superior temporal gyrus, left precuneus, and left cerebellum (Valli et al., 2016).

2.2 Single-Modality Machine Learning Studies Based on Diffusion Tensor Imaging

Diffusion Tensor Imaging is another commonly used structural brain imaging modality. Water molecules in the brain are constrained by white matter fibers (such as myelin and axonal membranes), showing different diffusion patterns across directions and exhibiting anisotropy. DTI utilizes this anisotropic diffusion to investigate the location and direction of white matter fiber tracts (Meoded, Poretti, Mori, & Zhang, 2017). Existing machine learning classification studies based on DTI are limited, with sample sizes ranging from 42 to 154. Commonly used features include Fractional Anisotropy (FA), Axial Diffusivity (AD), Radial Diffusivity (RD), and Mean Diffusivity (MD). Classification accuracies in current studies range from 60% to 98%.

One study divided 50 schizophrenia patients and 50 healthy controls into training and test sets, calculating FA and MD values from DTI data. After PCA dimensionality reduction and independent samples t-tests, features with P-values greater than 0.5 were removed, and models were trained using Fisher Linear Discriminant Analysis. Results showed that models built using FA values achieved 94% accuracy in the test set (sensitivity 96%, specificity 92%), with discrimi-

native features concentrated in deep white matter regions such as the bilateral external capsule. Models using MD values achieved 98% accuracy in the test set (sensitivity 96%, specificity 100%), with features affecting accuracy concentrated in the cortex and ventricles. Combining FA and MD data did not significantly change classification accuracy (accuracy, sensitivity, and specificity all at 96%) (Ardekani et al., 2011).

Another study calculated FA, MD, RD, and AD values for 18 major white matter fiber tracts from DTI data, using Random Forest-Recursive Feature Elimination (RF-RFE) for feature selection and Random Forest algorithm for training and testing in 52 first-episode patients and 48 healthy controls, achieving 71% classification accuracy (sensitivity 67.3%, specificity 75%). The final model distinguished 13 first-episode patients and 12 healthy controls with 76% accuracy (sensitivity 76.9%, specificity 75%). Discriminative features were located in commissural fibers, cerebello-thalamo-cortical circuits, and long fibers (Deng et al., 2019).

Pettersson-Yeo et al. (2013) used FA values to classify 19 high-risk individuals versus 23 healthy controls, and 19 first-episode patients versus 23 healthy controls, using SVM algorithm and achieving 65.79% classification accuracy for both comparisons, with sensitivity at 68.42% and specificity at 63.16% for both.

2.3 Multi-Modality Neuroimaging Data Machine Learning Studies

In addition to single-modality data, researchers have attempted to combine multiple neuroimaging modalities to classify schizophrenia patients or high-risk populations from healthy controls. Currently, there are 8 multi-modality neuroimaging machine learning classification studies, with total sample sizes ranging from 46 to 144 and classification accuracies between 54.9% and 100%. Brain imaging data commonly combined with T1-weighted imaging include resting-state functional Magnetic Resonance Imaging (rs-fMRI) and DTI data.

Peruzzo et al. (2015) combined T1-weighted imaging and DTI data, using non-parametric statistical tests and SVM for feature selection, and MKL algorithm to classify 23 first-episode schizophrenia patients and 23 healthy controls, achieving maximum accuracy of 93.5%. Features affecting accuracy included gray matter volume of the amygdala, superior frontal gyrus, inferior frontal gyrus, globus pallidus, and parahippocampal gyrus, as well as lateral ventricle volume; cortical thickness of the superior frontal gyrus, middle frontal gyrus, inferior frontal gyrus, fusiform gyrus, parahippocampal gyrus, middle temporal gyrus, orbitofrontal cortex, and insula; and diffusion tensor values of the uncinate fasciculus, cingulum, inferior cerebellar peduncle, and splenium of the corpus callosum.

Additionally, Sui et al. (2013) proposed a multi-modality fusion method—multi-set Canonical Correlation Analysis (mCCA) and joint Independent Component Analysis (jICA)—combining rs-fMRI, structural imaging, and DTI data pairwise to classify 35 schizophrenia patients and 28 healthy controls. They found single-

modality classification accuracies between 50%~70%, while multi-modality accuracies ranged from 60%~80%. The same research team also combined rs-fMRI, T1-weighted imaging, and Electroencephalogram (EEG) data, using independent samples t-tests, mCCA, and SVM-RFE for feature selection, applying SVM algorithm and 10-fold cross-validation to 43 schizophrenia patients and 48 healthy controls, achieving 91% classification accuracy and successfully distinguishing 5 schizophrenia patients and 5 healthy controls with 100% accuracy (Sui et al., 2014).

Although most current multi-modality neuroimaging machine learning studies report high classification accuracy, such studies typically have small sample sizes. Therefore, the impact of multi-modality neuroimaging data on classification results requires further investigation.

2.4 Machine Learning Studies Combining Brain Imaging and Behavioral Data

In addition to combining different neuroimaging modalities, some studies have combined brain imaging data with behavioral data as classification features. Current research suggests that combining brain imaging and behavioral data can improve classification accuracy.

One study combined gray matter, white matter, and cerebrospinal fluid volume with error rates from the Wisconsin Card Sorting Test (WCST), processed data using Independent Component Analysis, and applied SVM algorithm to data from 19 schizophrenia patients and 16 healthy controls, achieving 91.58% classification accuracy. This demonstrated that combining WCST error rates with neuroimaging data could effectively distinguish patients from healthy controls (Chu, Huang, Jian, Hsu, & Cheng, 2016). Zarogianni et al. (2017) predicted which 34 high-risk individuals with family history would convert to schizophrenia, applying SVM-RFE algorithm for feature selection among gray matter density data across brain regions, Rust Inventory of Schizotypal Cognitions (RISC) questionnaire scores, and Rey Auditory Verbal Learning Test (RAVLT) scores, using SVM algorithm for prediction and achieving 94% classification accuracy —higher than the 88% accuracy obtained using only gray matter density data. Ebdrup et al. (2018) classified 46 first-episode schizophrenia patients and 58 healthy controls, combining behavioral data (including Wechsler Adult Intelligence Scale III and Danish Adult Reading Test) with T1-weighted imaging, DTI, and EEG data, achieving only 51%~68% classification accuracy—lower than results using cognitive data alone (60%~69%) but slightly higher than single-modality data alone (49%~56%). Therefore, current research suggests that combining behavioral and neuroimaging data may yield better classification results than single-modality data, though whether this combination can achieve high accuracy remains to be further explored.

Figure 2 [Figure 2: see original paper]. Total sample size and classification accuracy in schizophrenia patients and high-risk populations based on structural

brain imaging. SCZ: schizophrenia; HC: healthy control; FEP: first-episode psychosis; HR: high-risk population. Solid dots represent maximum classification accuracy; hollow dots represent minimum classification accuracy; dashed lines connect maximum and minimum classification accuracy for the same study with identical sample sizes.

3. Limitations of Machine Learning Studies

Applying machine learning to schizophrenia research using brain structural features to classify schizophrenia patients and high-risk populations from healthy controls helps understand disease pathophysiological mechanisms. Although the number of studies has increased annually, limitations remain, particularly regarding sample size and model generalization ability, which require urgent resolution.

3.1 Sample Size

Figure 2 summarizes total sample sizes and maximum/minimum classification accuracies from existing machine learning studies in schizophrenia and high-risk populations. As shown in Figure 2, when sample sizes range from 50 to 150, maximum classification accuracy is relatively high; when sample size exceeds 400, maximum accuracy is only 60%~72%. Several reasons may explain why small samples yield high accuracy while large samples yield lower accuracy: (1) Small sample sizes lead to model instability and inaccurate testing. Insufficient training samples can cause inadequate training, resulting in model overfitting or instability; small test samples are prone to artificially high or low classification accuracy, making it difficult to validate model performance even when high accuracy is achieved in small-sample testing. Research indicates that validating model effectiveness requires at least 75 to 100 test samples (Beleites, Neugebauer, Bocklitz, Krafft, & Popp, 2013). (2) In current large-sample studies, classification features far outnumber samples without feature selection, easily causing model overfitting and reducing classification accuracy. (3) Large-sample studies often involve multi-center data with potentially higher patient heterogeneity, thereby reducing classification accuracy.

Insufficient sample size is a major problem facing current schizophrenia machine learning research (Madsen et al., 2018). Primary reasons include high MRI scanning costs, the special nature of schizophrenia, and difficulties in data sharing. Expensive MRI scanning fees and long scan times make it difficult for researchers to collect large amounts of neuroimaging data quickly. Schizophrenia patients, affected by clinical symptoms and medication side effects, require more human and material resources during neuroimaging data collection, making acquisition difficult and resulting in typically small sample sizes in schizophrenia imaging studies. Additionally, multi-center data sharing difficulties arise from varying scanner models and parameters across centers, as well as data management, ethical, and privacy protection issues.

In the future, multiple research institutions/centers can increase sample sizes by establishing detailed protocols to protect subject privacy and data security, unifying scanning parameters and inclusion criteria, and enabling multi-center data sharing without ethical violations. When incorporating multi-center data, researchers need to ensure consistent inclusion criteria as much as possible. When processing large-sample data, data cleaning and dimensionality reduction/feature selection methods should be used to retain useful features, ensure data quality, and obtain more reliable results, thereby enabling investigation of brain structural feature validity using large-sample data. Some researchers have proposed addressing sample size issues by merging multi-center models. Researchers collected structural brain data from first-episode schizophrenia patients and healthy controls across 4 centers, trained models separately using each center' s data, and used these 4 models to generate a final meta-model. The study found this meta-model showed high similarity to models generated from all data combined (Dluhoš et al., 2017). Therefore, merging multi-center models provides a new approach to solving data sharing limitations.

3.2 Generalization Ability

The goal of machine learning is to train models that generalize well to new samples. In addition to insufficient sample size, information leakage during training also affects generalization ability. A common form of information leakage involves using the same samples for parameter tuning and result testing, or using test set labels prematurely during feature selection, such as conducting independent samples t-tests across all samples to select significant brain regions before splitting training, test, and validation sets. While this premature use of test set information can yield high accuracy, it does not reflect true model generalization ability. Researchers should avoid information leakage and use nested cross-validation in future studies (Madsen et al., 2018; Dwyer et al., 2018) to make evaluation results more precise and improve model generalization. To test generalization ability, Dwyer et al. (2018) recommend establishing third-party testing agencies to collect data from various centers or institutions. Once new models are developed, researchers can submit them to third-party agencies for testing using their data, thereby ensuring better generalization ability.

4. Future Research Directions

Future schizophrenia neuroimaging machine learning studies could consider incorporating medication status, clinical symptoms, and sociodemographic information into features. Kambeitz et al. (2015) used sensitivity and specificity as effect sizes in a meta-analysis investigating the impact of age, positive-to-negative symptom ratio, and antipsychotic medication dosage on classification results in schizophrenia patients versus healthy controls. The meta-analysis found that older age was associated with higher sensitivity; higher positive-to-negative symptom ratio was associated with higher specificity; and higher antipsychotic medication dosage was associated with higher specificity. There-

fore, when distinguishing patients with different disease courses or investigating prognosis, clinical and sociodemographic information can be incorporated to improve classification performance.

In recent years, with rapid advances in computer technology, deep learning has regained attention. Deep learning is an algorithm that uses artificial neural network architectures for representation learning of data and has demonstrated excellent performance in image, video, speech, and other domains (Bengio, Courville, & Vincent, 2013; Lecun, Bengio, & Hinton, 2015; Schmidhuber, 2015). Pinaya et al. (2016) used Deep Belief Networks to achieve 73.6% classification accuracy (sensitivity 76.37%, specificity 70.74%) in 143 schizophrenia patients and 83 healthy controls, slightly higher than the 68.1% accuracy obtained using SVM. Dwyer et al. (2018) note that deep learning requires large-sample data to avoid the curse of dimensionality and overfitting. Therefore, researchers can attempt to use deep learning methods for classification and investigate the underlying mechanisms of schizophrenia once large-sample data are obtained.

Current studies employ different data processing pipelines and report results differently. Due to lack of specific information and methodological differences, comparisons across studies are difficult. Riley (2019) therefore suggests establishing unified data analysis and results reporting standards when applying machine learning across different disciplines.

Machine learning technology extracts useful information from data, assisting researchers in making predictions at the individual level, investigating schizophrenia characteristics and underlying mechanisms, and helping identify biomarkers for early recognition of mental illness. This review primarily focuses on structural brain imaging, summarizing single-modality, multi-modality, and brain-behavior combined machine learning studies. Single-modality neuroimaging machine learning classification accuracies range from 36.84% to 98%, mostly based on T1-weighted imaging, with results indicating that frontal and temporal lobe structural features can effectively distinguish schizophrenia patients or high-risk populations from healthy controls. DTI-based machine learning studies predominantly use diffusion tensor values as features, but due to small sample sizes and limited studies, brain regions or fiber features proposed to improve classification accuracy vary across researchers without consensus. Multi-modality neuroimaging classification accuracies range from 54.9% to 100%, with most studies reporting high accuracy but small sample sizes. Whether multi-modality data significantly outperforms single-modality data in improving classification accuracy requires further investigation. Classification combining behavioral and neuroimaging data performs better than single-modidity data, though whether this significantly improves accuracy remains to be verified. Current schizophrenia machine learning research is limited by insufficient sample size and weak generalization ability. Future research needs to train and validate models in large multi-center samples to obtain more reliable models and more effective features for greater clinical utility.

Table 1 Literature Summary

Study	Sample	Algorithm	Cross-Validation	Accuracy
Wang, L., Shen, H., Li, B., & Hu, D., 2011	SCZ(32) HC(32)	Linear SVM	10-fold CV	59.4%~82.8%
Greenstein, Malley, Weisinger, Clasen, & Gogtay, 2012	SCZ(98) HC(99)	Linear SVM	LOOCV	73.7%
Nieuwenhuis et al., 2012	SCZ(128) HC(111) SCZ(155) HC(122)	Linear SVM + SVM feature selection	LOOCV	67.5%~86.8% (embedded) 69.1%~70.6% (prior knowledge)
Iwabuchi, Liddle, & Palaniyappan, 2013	SCZ(19) HC(20)	Linear SVM	Leave-4-out, Leave-2-out	63.9%~77%
Schnack et al., 2014	SCZ(66) BD(66) HC(66) SCZ(46) BD(47) HC(43)	Linear SVM	Leave-2-out	SCZ vs HC: 90.15% SCZ vs BD: 87.9% HC vs BD: 59.85% HC vs SCZ: 76% BD vs SCZ: 66% HC vs BD: 61%
Gould et al., 2014	CD(74) CS(126) HC(134)	Linear SVM	Leave-2-out	HC vs CD: 64%~72% HC vs CS: 59%~67% CD vs CS: 54%~59%
Guo et al., 2016	SCZ(98) HC(83)	SVM-RBF	LOOCV	81.77%~96.3%
Lu et al., 2016	SCZ(41) HC(42)	SVM-RFE	Nested CV (outer: 10-fold, inner: 10-fold)	79.5%~88.4%

Study	Sample	Algorithm	Cross-Validation	Accuracy
Salvador et al., 2017	SCZ(128) BD(128) HC(127)	Ridge LR; Lasso LR; Elastic Net Regularization; L0-norm Regularization; SVM; RDA; GPC; Random Forests	10-fold CV	SCZ vs HC: 74.1%~76% SCZ vs BD: 49.2%~66% BD vs HC: 49.1%~65.5%
Chin et al., 2018	SCZ(84) HC(43) SCZ(57) HC(28)	Sequential ROI selection algorithm	10-fold CV	Training: 86.6%~92% Test: 83.5%~89.4%
de Pierrefeu et al., 2018	SCZ(276) HC(330)	linear SVM, Elastic Net, Enet-TV	2-fold CV	61%~72% (training) 61%~73% (test)
Takayanagi, Y. et.al., 2011	FEP(43) HC(90)	Linear SVM	LOOCV	81.2%~96.7%
Zanetti, M. V., et. al., 2013	FEP(62) HC(62)	Linear SVM + SVM feature selection	LOOCV	64.3%~73.4%
Dluhoš et al., 2017	FEP(258) HC(222)	Linear SVM	2-fold CV	52%~66%
Squarcina et al., 2017	FEP(127) HC(127)	SVM-RBF, MKL + SVM feature selection	10-fold CV	68%~85%
Xiao et al., 2017	FEP(163) HC(163)	Linear SVM	Two-sample t-test ($p<0.05$) + 10-fold CV	81.8%~85.0%
Winterburn et al., 2017	FEP(50) HC(220) SCZ(179)	LR, SVM(RBF/Linear), LDA	10-fold CV	SCZ vs HC: 50.0%~71.9% FEP vs HC: 50.0%~71.2% SCZ vs HC: 51.0%~70.8% FEP vs HC: 50.6%~73.5%

Study	Sample	Algorithm	Cross-Validation	Accuracy
de Moura et al., 2018	FEP(32) chronic SCZ(143) HC(32)	SVM	10-fold CV	chronic SCZ vs HC: 73.0% FEP vs HC: 59.4%
Ardekani et al., 2011	SCZ(25) HC(25) SCZ(25) HC(25)	Fisher Linear Discriminant	Two-sample t-test ($p<0.5$) + 100×2-fold CV	Training: 94%~98% Test: 96%
Rathi Y. et. al., 2010	FEP(21) HC(20)	Parzen window classifier, KNN	Leave-2-out	69%~86% sensitivity 68%~85% specificity
Mikolas, P. et. al., 2018	FEP(77) HC(77)	Linear SVM	Leave-2-out	62.34%
Deng et al., 2019	FEP(52) HC(48) FEP(13) HC(12)	RF-RFE	Training: 71.0% Test: 76.0%	
Sui et al., 2013	SCZ(35) HC(28)	SVM(RBF/Linear), KNN, GNB mCCA+jICA	10-fold CV	Single-modality: (50%~70%) Multi-modality: (60%~80%)
Sui et al., 2014	SCZ(43) HC(48) SCZ(5) HC(5)	Linear SVM mCCA, two-sample t-test, SVM-RFE	10-fold CV	rs-fMRI Multi-modality: (80%)~91% Single-modality: 50%~90% Multi-modality: 80%~100%
Qureshi, Oh, Cho, Jo, & Lee, 2017	SCZ(72) HC(72)	SVM(RBF/Linear), ELM, LDA, RF	10×10-fold CV	Single-modality: 42.4%~93.0% Multi-modality: 54.9%~99.3%

Study	Sample	Algorithm	Cross-Validation	Accuracy
Tas et al., 2018	SCZ(23) schizo-obsessive (23)	SVM-RBF	Nested CV (outer: 6-fold, inner: 5-fold)	78.26%
Peruzzo et al., 2015	FEP(23) HC(23)	SVM, MKL	Non-parametric test ($p<0.05$) + SVM feature selection	78.26%~93.48%
Lu et al., 2018	FEP(44) chronic SCZ(44) HC(56)	SVM-RFE	10-fold CV	Multi-modality: 85% Single-modality: 70.7%~72%
Karageorgiou, E., et al., 2011	SCZ(28) HC(47)	Neuropsychological testing + stepwise LDA	PCA-LDA	Wilk' s Λ ($p=0.05$ retain, $p=0.10$ remove)
Chu et al., 2016	SCZ(19) HC(16)	the error rate of the WCST	naïve Bayes, LR, SVM, decision tree, RF, auto-sklearn	76%~92%
Ebdrup et al., 2018	FEP(46) HC(58)	Cognition data (DART, WAIS III)	Backwards elimination feature selection	Neuropsychological + neu-roimaging: 82.794%~91.575% Single-modality: 49%~56% Cognitive + imaging: 51%~68%
Pettersson-Yeo et al., 2013	UHR(19) FEP(19) HC(23)	Molecular genetics Neu-ropsychology	Linear SVM	UHR vs HC: 36.84%~68.42% UHR vs FEP: 33.33%~76.67% HC vs FEP: 44.74%~68.42%
Zarogianni et al., 2017	HRill HRsymp	SVM-RFE	Nested CV (outer: LOOCV, inner: LOOCV)	Single-modality: 88% Behavioral + imaging: 94.1%

Abbreviations: SCZ: schizophrenia patients; HC: healthy controls; FEP: first-episode psychosis patients; UHR: ultra-high-risk population; HR: high-risk population, where HR[ill] are high-risk individuals who converted to schizophrenia and HR[symp] are those who did not; BD: bipolar disorder; DTI: Diffusion Tensor Imaging; T1: T1-weighted imaging; DWI: diffusion weighted imaging; fMRI: functional magnetic resonance imaging, rs-fMRI: resting-state functional magnetic resonance imaging; EEG: electroencephalogram; LDA: Linear Discriminant Analysis; SVM: Support Vector Machine, where SVM-RBF is SVM with Gaussian kernel and Linear SVM is linear support vector machine; ELM: extreme learning machine; naïve Bayes: naive Bayes; decision tree: decision tree; kNN: k-Nearest Neighbor; GNB: Gaussian Naive Bayes; MKL: Multiple Kernel Learning; Elastic Net Regularization: elastic net regularization, Enet-TV: ElasticNet-Total Variation; L0-norm Regularization: L0-norm regularization; RDA: Regularized Discriminant function Analysis; GPC: Gaussian Process Classifier; RF: Random Forests; LR: Logistic Regression, where Lasso LR is Lasso regression and Ridge LR is ridge regression; SVM-RFE: Support Vector Machine-Recursive Feature Elimination; DART: Danish Adult Reading Test; WAIS III: Wechsler Adult Intelligence Scale III; MLDA: Maximum Uncertainty Linear Discriminant Analysis; PCA: Principal Component Analysis; ACO: ant colony optimization; schizo-obsessive: schizophrenia with obsessive-compulsive symptoms; cognitive deficit subtype: cognitive deficit subtype of schizophrenia; cognitively spared subtype: cognitively normal subtype of schizophrenia
a: Articles did not explicitly report accuracy values

Zhou, Z. (2016). *Machine Learning*. Beijing: Tsinghua University Press.

References

Ardekani, B. A., Tabesh, A., Sevy, S., Robinson, D. G., Bilder, R. M., & Szeszko, P. R. (2011). Diffusion tensor imaging reliably differentiates patients with schizophrenia from healthy volunteers. *Human Brain Mapping, 32*(1), 1-9. https://doi.org/10.1002/hbm.20995

Bakhshi, K., & Chance, S. A. (2015). The neuropathology of schizophrenia: A selective review of past studies and emerging themes in brain structure and cytoarchitecture. *Neuroscience*. https://doi.org/10.1016/j.neuroscience.2015.06.028

Beleites, C., Neugebauer, U., Bocklitz, T., Krafft, C., & Popp, J. (2013). Sample size planning for classification models. *Analytica Chimica Acta, 760*, 25-33. doi:10.1016/j.aca.2012.11.007

Bengio, Y., Courville, A., & Vincent, P. (2013). Representation learning: A review and new perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. https://doi.org/10.1109/TPAMI.2013.50

Borgwardt, S., Koutsouleris, N., Aston, J., Studerus, E., Smieskova, R., Riecher-Rössler, A., & Meisenzahl, E. M. (2013). Distinguishing prodromal from first-episode psychosis using neuroanatomical single-

subject pattern recognition. *Schizophrenia Bulletin, 39*(5), 1105–1114. https://doi.org/10.1093/schbul/sbs095

Carpenter, W. T., & Buchanan, R. W. (1994). Schizophrenia. *New England Journal of Medicine, 330*(10), 681–690. https://doi.org/10.1056/NEJM199403103301006

Chin, R., You, A. X., Meng, F., Zhou, J., & Sim, K. (2018). Recognition of Schizophrenia with Regularized Support Vector Machine and Sequential Region of Interest Selection using Structural Magnetic Resonance Imaging. *Scientific Reports, 8*(1), 13858. https://doi.org/10.1038/s41598-018-32290-9

Chu, W.-L., Huang, M.-W., Jian, B.-L., Hsu, C.-Y., & Cheng, K.-S. (2016). A Correlative Classification Study of Schizophrenic Patients with Results of Clinical Evaluation and Structural Magnetic Resonance Images. *Behavioural Neurology, 2016*, 1–11. https://doi.org/10.1155/2016/7849526

de Moura, A. M., Pinaya, W. H. L., Gadelha, A., Zugman, A., Noto, C., Cordeiro, Q., ···Sato, J. R. (2018). Investigating brain structural patterns in first episode psychosis and schizophrenia using MRI and a machine learning approach. *Psychiatry Research - Neuroimaging, 275*(March), 14–20. https://doi.org/10.1016/j.pscychresns.2018.03.003

de Pierrefeu, A., Löfstedt, T., Laidi, C., Hadj-Selem, F., Bourgin, J., Hajek, T., ···Duchesnay, E. (2018). Identifying a neuroanatomical signature of schizophrenia, reproducible across sites and stages, using machine learning with structured sparsity. *Acta Psychiatrica Scandinavica*. https://doi.org/10.1111/acps.12964

Deng, Y., Hung, K. S. Y., Lui, S. S. Y., Chui, W. W. H., Lee, J. C. W., Wang, Y., ···Cheung, E. F. C. (2019). Tractography-based classification in distinguishing patients with first-episode schizophrenia from healthy individuals. *Progress in Neuro-Psychopharmacology and Biological Psychiatry, 88*, 66–73. https://doi.org/10.1016/j.pnpbp.2018.06.010

Dluhoš, P., Schwarz, D., Cahn, W., van Haren, N., Kahn, R., Španiel, F., ···Schnack, H. (2017). Multi-center machine learning in imaging psychiatry: A meta-model approach. *NeuroImage, 155*(April), 10–24. https://doi.org/10.1016/j.neuroimage.2017.03.027

Dwyer, D. B., Falkai, P., & Koutsouleris, N. (2018). Machine Learning Approaches for Clinical Psychology and Psychiatry. *Annual Review of Clinical Psychology, 14*(1), annurev-clinpsy-032816-045037. https://doi.org/10.1146/annurev-clinpsy-032816-045037

Dyrba, M., Grothe, M., Kirste, T., & Teipel, S. J. (2015). Multimodal analysis of functional and structural disconnection in Alzheimer' s disease using multiple kernel SVM. *Human Brain Mapping, 36*(6), 2118–2131. https://doi.org/10.1002/hbm.22759

Ebdrup, B. H., Axelsen, M. C., Bak, N., Fagerlund, B., Oranje, B., Raghava, J. M., ···Glenthøj, B. Y. (2018). Accuracy of diagnostic classification algorithms

neuroanatomical data in antipsychotic-naïve schizophrenia patients. *Psychological Medicine*. https://doi.org/10.1017/S0033291718003781

Gould, I. C., Shepherd, A. M., Laurens, K. R., Cairns, M. J., Carr, V. J., & Green, M. J. (2014). Multivariate neuroanatomical classification of cognitive subtypes in schizophrenia: a support vector machine learning approach. *NeuroImage: Clinical, 6*, 229-236.

Greenstein, D., Malley, J. D., Weisinger, B., Clasen, L., & Gogtay, N. (2012). Using Multivariate Machine Learning Methods and Structural MRI to Classify Childhood Onset Schizophrenia and Healthy Controls. *Frontiers in Psychiatry, 3*(June), 1–12. https://doi.org/10.3389/fpsyt.2012.00053

Guo, S., Palaniyappan, L., Liddle, P. F., & Feng, J. (2016). Dynamic cerebral reorganization in the pathophysiology of schizophrenia: A MRI-derived cortical thickness study. *Psychological Medicine, 46*(10), 2201–2214. https://doi.org/10.1017/S0033291716000994

Guyon, I., Weston, J., Barnhill, S., & Vapnik, V. (2002). Gene selection for cancer classification using support vector machines. *Machine Learning*. https://doi.org/10.1023/A:1012487302797

Iwabuchi, S. J., Liddle, P. F., & Palaniyappan, L. (2013). Clinical Utility of Machine-Learning Approaches in Schizophrenia: Improving Diagnostic Confidence for Translational Neuroimaging. *Frontiers in Psychiatry, 4*(August), 1–9. https://doi.org/10.3389/fpsyt.2013.00095

Kambeitz, J., Kambeitz-Ilankovic, L., Leucht, S., Wood, S., Davatzikos, C., Malchow, B., ···Koutsouleris, N. (2015). Detecting neuroimaging biomarkers for schizophrenia: A meta-analysis of multivariate pattern recognition studies. *Neuropsychopharmacology, 40*(7), 1742–1751. https://doi.org/10.1038/npp.2015.22

Karageorgiou, E., Schulz, S. C., Gollub, R. L., Andreasen, N. C., Ho, B. C., Lauriello, J., ···& Georgopoulos, A. P. (2011). Neuropsychological testing and structural magnetic resonance imaging as diagnostic biomarkers early in the course of schizophrenia and related psychoses. *Neuroinformatics, 9*(4), 321-333.

Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*. https://doi.org/10.1038/nature14539

Lu, X. B., Zhang, Y., Yang, D. Y., Yang, Y. Z., Wu, F. C., Ning, Y. P., & Wu, K. (2018). Analysis of first-episode and chronic schizophrenia using multi-modal magnetic resonance imaging. *European Review for Medical and Pharmacological Sciences, 22*(19), 6422–6435. https://doi.org/10.26355/eurrev_201810_16055

Lu, X., Yang, Y., Wu, F., Gao, M., Xu, Y., Zhang, Y., ···Wu, K. (2016). Discriminative analysis of schizophrenia using support vector machine and recursive feature elimination on structural MRI images. *Medicine (United States), 95*(30), e3973. https://doi.org/10.1097/MD.0000000000003973

Madsen, K. H., Krohne, L. G., Cai, X. L., Wang, Y., & Chan, R. C.

(2018). Perspectives on machine learning for classification of Schizotypy using fMRI data. *Schizophrenia bulletin, 44*(suppl_2), S480-S490. https://doi.org/10.1093/schbul/sby026

Meoded, A., Poretti, A., Mori, S., & Zhang, J. (2017). Diffusion Tensor Imaging (DTI) . *Reference Module in Neuroscience and Biobehavioral Psychology,* (March 2016), 1–11. https://doi.org/10.1016/B978-0-12-809324-5.02472-X

Mitchell, T. (1997). *Machine Learning*.

Nieuwenhuis, M., van Haren, N. E. M., Hulshoff Pol, H. E., Cahn, W., Kahn, R. S. R. S., & Schnack, H. G. (2012). Classification of schizophrenia patients and healthy controls from structural MRI scans in two large independent samples. *NeuroImage, 61*(3), 606–612. https://doi.org/10.1016/j.neuroimage.2012.03.079

Mikolas, P., Hlinka, J., Skoch, A., Pitra, Z., Frodl, T., Spaniel, F., & Hajek, T. (2018). Machine learning classification of first-episode schizophrenia spectrum disorders and controls using whole brain white matter fractional anisotropy. *BMC Psychiatry, 18*(1), 97. https://doi.org/http://dx.doi.org/10.1186/s12888-018-1678-y

Palaniyappan, L. (2017). Progressive cortical reorganisation: A framework for investigating structural changes in schizophrenia. *Neuroscience and Biobehavioral Reviews*. https://doi.org/10.1016/j.neubiorev.2017.04.028

Peng, B., Lu, J., Saxena, A., Zhou, Z., Zhang, T., Wang, S., & Dai, Y. (2017). Examining brain morphometry associated with self-esteem in young adults using Multilevel-ROI-Features-Based classification method. *Frontiers in Computational Neuroscience, 11*(May), 1–10. https://doi.org/10.3389/fncom.2017.00037

Peruzzo, D., Castellani, U., Perlini, C., Bellani, M., Marinelli, V., Rambaldelli, G., ···Brambilla, P. (2015). Classification of first-episode psychosis: A multimodal multi-feature approach integrating structural and diffusion imaging. *Journal of Neural Transmission (Vienna, Austria : 1996), 122*(6), 897–905. https://doi.org/10.1007/s00702-014-1324-x

Pettersson-Yeo, W., Benetti, S., Marquand, A. F., Dell' Acqua, F., Williams, S. C. R., Allen, P., ···Mechelli, A. (2013). Using genetic, cognitive and multi-modal neuroimaging data to identify ultra-high-risk and first-episode psychosis at the individual level. *Psychological Medicine, 43*(12), 2547–2562. https://doi.org/10.1017/S003329171300024X

Pinaya, W. H. L., Gadelha, A., Doyle, O. M., Noto, C., Zugman, A., Cordeiro, Q., ···Sato, J. R. (2016). Using deep belief network modelling to characterize differences in brain morphometry in schizophrenia. *Scientific Reports, 6*(1), 38897. https://doi.org/10.1038/srep38897

Qureshi, M. N. I., Oh, J., Cho, D., Jo, H. J., & Lee, B. (2017). Multimodal Discrimination of Schizophrenia Using Hybrid Weighted Feature Concatenation of Brain Functional Connectivity and Anatomical Features with an Ex-

treme Learning Machine. *Frontiers in Neuroinformatics, 11*(September), 1–14. https://doi.org/10.3389/fninf.2017.00059

Rathi Y, Malcolm J, Michailovich O, Goldstein J, Seidman L, McCarley RW et al (2010). Biomarkers for identifying first-episode schizophrenia patients using diffusion weighted imaging. *Med Image Comput Comput-Assist Interv 13*: 657–665.

Riley, P. (2019). Three pitfalls to avoid in machine learning. *Nature, 572*(7767), 27–29.

Salvador, R., Radua, J., Canales-Rodrõâguez, E. J., Solanes, A., Sarroa, S., Goikolea, J. M., ···Pomarol-Clote, E. (2017). Evaluation of machine learning algorithms and structural features for optimal MRI-based diagnostic prediction in psychosis. *PLoS ONE, 12*(4), 1–24. https://doi.org/10.1371/journal.pone.0175683

Samuel, A. L. (1959). Some studies in machine learning using the game of checkers. *IBM Journal of research and development, 3*(3), 210-229. https://doi.org/10.1147/rd.33.0210

Schmidhuber, J. (2015). Deep Learning in neural networks: An overview. *Neural Networks*. https://doi.org/10.1016/j.neunet.2014.09.003

Schnack, H. G., Nieuwenhuis, M., van Haren, N. E. M., Abramovic, L., Scheewe, T. W., Brouwer, R. M., ···Kahn, R. S. (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.

Squarcina, L., Castellani, U., Bellani, M., Perlini, C., Lasalvia, A., Dusi, N., ···Brambilla, P. (2017). Classification of first-episode psychosis in a large cohort of patients using support vector machine and multiple kernel learning techniques. *NeuroImage, 145*(Pt B), 238–245. https://doi.org/10.1016/j.neuroimage.2015.12.007

Sui, J., Castro, E., He, H., Bridwell, D., Du, Y., Pearlson, G. D., ···Calhoun, V. D. (2014). Combination of FMRI-SMRI-EEG data improves discrimination of schizophrenia patients by ensemble feature selection. *Conference Proceedings : ···Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference, 2014*(2), 3889–3892. https://doi.org/10.1109/EMBC.2014.6944473

Sui, J., He, H., Yu, Q., Chen, J., Rogers, J., Pearlson, G. D., ···Calhoun, V. D. (2013). Combination of Resting State fMRI, DTI, and sMRI Data to Discriminate Schizophrenia by N-way MCCA + jICA. *Frontiers in Human Neuroscience, 7*(May), 1–14. https://doi.org/10.3389/fnhum.2013.00235

Takayanagi, Y., Takahashi, T., Orikabe, L., Mozue, Y., Kawasaki, Y., Nakamura, K., ···& Kurachi, M. (2011). Classification of first-episode schizophrenia

patients and healthy subjects by automated MRI measures of regional brain volume and cortical thickness. *PloS one, 6*(6), e21047.

Tas, C., Mogulkoc, H., Eryilmaz, G., Gogcegoz-Gul, I., Erguzel, T. T., Metin, B., & Tarhan, N. K. (2018). Discriminating schizophrenia and schizo-obsessive disorder: a structural MRI study combining VBM and machine learning methods. *Neural Computing and Applications, 29*(2), 377–387. https://doi.org/10.1007/s00521-016-2451-0

Valli, I., Marquand, A. F., Mechelli, A., Raffin, M., Allen, P., Seal, M. L., & McGuire, P. (2016). Identifying Individuals at High Risk of Psychosis: Predictive Utility of Support Vector Machine using Structural and Functional MRI Data. *Frontiers in Psychiatry, 7*, 52. https://doi.org/10.3389/fpsyt.2016.00052

Wang, L., Shen, H., Li, B., & Hu, D. (2011). Classification of schizophrenic patients and healthy controls using multiple spatially independent components of structural MRI data. *Frontiers of Electrical and Electronic Engineering in China, 6*(2), 353–362. https://doi.org/10.1007/s11460-011-0142-2

Winterburn, J. L., Voineskos, A. N., Devenyi, G. A., Plitman, E., de la Fuente-Sandoval, C., Bhagwat, N., ···Chakravarty, M. M. (2017). Can we accurately classify schizophrenia patients from healthy controls using magnetic resonance imaging and machine learning? A multi-method and multi-dataset study. *Schizophrenia Research*. https://doi.org/10.1016/j.schres.2017.11.038

Xiao, Y., Yan, Z., Zhao, Y., Tao, B., Sun, H., Li, F., ···Lui, S. (2017). Support vector machine-based classification of first episode drug-naïve schizophrenia patients and healthy controls using structural MRI. *Schizophrenia Research*. https://doi.org/10.1016/j.schres.2017.11.037

Zarogianni, E., Storkey, A. J., Johnstone, E. C., Owens, D. G. C., & Lawrie, S. M. (2017). Improved individualized prediction of schizophrenia in subjects at familial high risk, based on neuroanatomical data, schizotypal and neurocognitive features. *Schizophrenia Research, 181*, 6–12. https://doi.org/10.1016/j.schres.2016.08.027

Zanetti, M. V., Schaufelberger, M. S., Doshi, J., Ou, Y., Ferreira, L. K., Menezes, P. R., ···& Busatto, G. F. (2013). Neuroanatomical pattern classification in a population-based sample of first-episode schizophrenia. *Progress in Neuro-Psychopharmacology and Biological Psychiatry, 43*, 116-125.

Note: Figure translations are in progress. See original paper for figures.