



## OPEN ACCESS

## EDITED BY

Valerio Zerbi,  
Swiss Federal Institute of Technology  
Lausanne, Switzerland

## REVIEWED BY

Francois Maurice Lalonde,  
National Institute of Mental Health  
(NIH), United States  
Renato de Filippis,  
Magna Graecia University, Italy

## \*CORRESPONDENCE

Jacob Levman  
jlevman@stfx.ca

## SPECIALTY SECTION

This article was submitted to  
Neurodevelopment,  
a section of the journal  
Frontiers in Neuroscience

RECEIVED 22 April 2022

ACCEPTED 05 July 2022

PUBLISHED 15 August 2022

## CITATION

Levman J, Jennings M, Rouse E,  
Berger D, Kabaria P, Nangaku M,  
Gondra I and Takahashi E (2022) A  
morphological study of schizophrenia  
with magnetic resonance imaging,  
advanced analytics, and machine  
learning.  
*Front. Neurosci.* 16:926426.  
doi: 10.3389/fnins.2022.926426

## COPYRIGHT

© 2022 Levman, Jennings, Rouse,  
Berger, Kabaria, Nangaku, Gondra and  
Takahashi. This is an open-access  
article distributed under the terms of  
the [Creative Commons Attribution  
License \(CC BY\)](#). The use, distribution  
or reproduction in other forums is  
permitted, provided the original  
author(s) and the copyright owner(s)  
are credited and that the original  
publication in this journal is cited, in  
accordance with accepted academic  
practice. No use, distribution or  
reproduction is permitted which does  
not comply with these terms.

# A morphological study of schizophrenia with magnetic resonance imaging, advanced analytics, and machine learning

Jacob Levman<sup>1,2,3\*</sup>, Maxwell Jennings<sup>1,4</sup>, Ethan Rouse<sup>1</sup>,  
Derek Berger<sup>1</sup>, Priya Kabaria<sup>5</sup>, Masahito Nangaku<sup>5</sup>,  
Iker Gondra<sup>1</sup> and Emi Takahashi<sup>3,5</sup>

<sup>1</sup>Department of Computer Science, St. Francis Xavier University, Antigonish, NS, Canada, <sup>2</sup>Center for Clinical Research, Nova Scotia Health Authority - Research, Innovation and Discovery, Halifax, NS, Canada, <sup>3</sup>Department of Radiology, Massachusetts General Hospital, Athinoula A. Martinos Center for Biomedical Imaging, Harvard Medical School, Massachusetts Institute of Technology, Boston, MA, United States, <sup>4</sup>Department of Mathematics and Statistics, St. Francis Xavier University, Antigonish, NS, Canada, <sup>5</sup>Department of Medicine, Boston Children's Hospital, Harvard Medical School, Boston, MA, United States

We have performed a morphological analysis of patients with schizophrenia and compared them with healthy controls. Our analysis includes the use of publicly available automated extraction tools to assess regional cortical thickness (inclusive of within region cortical thickness variability) from structural magnetic resonance imaging (MRI), to characterize group-wise abnormalities associated with schizophrenia based on a publicly available dataset. We have also performed a correlation analysis between the automatically extracted biomarkers and a variety of patient clinical variables available. Finally, we also present the results of a machine learning analysis. Results demonstrate regional cortical thickness abnormalities in schizophrenia. We observed a correlation ( $\rho = 0.474$ ) between patients' depression and the average cortical thickness of the right medial orbitofrontal cortex. Our leading machine learning technology evaluated was the support vector machine with stepwise feature selection, yielding a sensitivity of 92% and a specificity of 74%, based on regional brain measurements, including from the insula, superior frontal, caudate, calcarine sulcus, gyrus rectus, and rostral middle frontal regions. These results imply that advanced analytic techniques combining MRI with automated biomarker extraction can be helpful in characterizing patients with schizophrenia.

## KEYWORDS

schizophrenia, magnetic resonance imaging, machine learning, depression, correlation analysis

## Introduction

Schizophrenia is a mental disorder characterized by hallucinations, delusions, thought disorders and a lack of motivation. Although a number of post-mortem brain studies in the early twentieth century failed to find abnormalities in schizophrenia (e.g., Southard, 1910, 1915; Jacobi and Winkler, 1928; Haug, 1962), the first computer assisted tomography (CT) study succeeded in observing enlarged lateral ventricles (Johnstone et al., 1976). More recently, magnetic resonance imaging (MRI) was introduced in schizophrenia research. MRI-based studies, as well as CT-based studies, confirmed ventricular enlargement, and it has become a consensus that ventricular enlargement is present in first episode patients who are not yet treated by medication (Nordström and Williamson, 2003). The use of MRI has also led to successes in evaluating abnormalities regionally, including investigations of the medial temporal lobe, the prefrontal, orbitofrontal, and parietal lobe in addition to the ventricles (see reviews in Shenton et al., 2001; Kubicki et al., 2005; Haukvik et al., 2013).

The analysis of patients with schizophrenia by MRI examinations has been the subject of many studies. However, existing research has been limited in the populations studied, particularly in terms of the ages of the patients in the study, and in terms of the regions analyzed (Kim et al., 2012; Zalesky et al., 2012; Wu et al., 2015). The use of MRI in schizophrenia research has been further supported by the use of automated measurement tools, such as FreeSurfer (Fischl, 2012). FreeSurfer is a brain imaging analysis technology that aids in identifying and extracting measurements from cortical and non-cortical regions of the brain. It performs volume-based and surface-based analyses, reconstructs accurate models of gray/white matter and pial surfaces, and provides a variety of measurements, such as cortical thicknesses, surface areas, and folding measurements. Studies of schizophrenia combining MRI technology and FreeSurfer have found interdependence between abnormal white matter connections and altered gray matter structure (Liu et al., 2014), reduced hippocampal volume (Arnold et al., 2015; Haukvik et al., 2015; Singh et al., 2018), reduced amygdalar volume (Rich et al., 2016), reduced cerebellar cortical volume (Laidi et al., 2015), thinner cortices (van Erp et al., 2018), and cortical (Yasuda et al., 2020) and subcortical brain volume abnormalities (van Erp et al., 2014) associated with the condition. These findings imply that a variety of abnormalities are present in patients with schizophrenia, and that FreeSurfer technology (Fischl, 2012) is capable of assisting in characterizing regional irregularities potentially associated with the condition.

The application of machine learning to MRI examinations of patients with schizophrenia has been the subject of multiple review articles (de Fillippis et al., 2019; Steardo et al., 2020). Briefly, previous research in the application of machine learning to schizophrenia diagnostics has focused on discriminating

schizophrenic patients from those with bipolar disorder (Schnack et al., 2014) and healthy controls (Nieuwenhuis et al., 2012; Iwabuchi et al., 2013; Schnack et al., 2014). Work has also focused on using machine learning to classify childhood-onset schizophrenia (Greenstein et al., 2012). Additional research has focused on differential diagnosis between patients with schizophrenia with and without auditory hallucinations using resting state functional MRI (Chyzyk et al., 2015), as well as classifying schizophrenic patients into cognitive subtypes (Gould et al., 2014). The combination of MRI and genetic data to improve classification of schizophrenia has also been investigated (Yang et al., 2010). Machine learning (ML) applied to MRI exams of patients with schizophrenia has been the subject of a review article (Veronese et al., 2013) and meta-analysis (Kambeitz et al., 2015). More recently, research has focused on the use of deep learning on structural MRI exams, achieving widely varying performance (area under the curve ranges from 0.71 to 0.90) on unseen datasets (Oh et al., 2020), ML based on whole brain white matter fractional anisotropy (accuracy: 62%) (Mikolas et al., 2018), ML applied to functional MRI (accuracy: 79%) (Gallos et al., 2021), and ML based on voxel-based morphometry on a small dataset (accuracy: 88%) (Lu et al., 2016), as well as a more recent study on a larger cohort (accuracy: 75%) (Salvador et al., 2017). Unfortunately, the differing datasets employed result in variation in the patient populations and image acquisition quality across studies, making direct comparison of widely accepted performance metrics, such as overall accuracy, extremely difficult. Additionally, the ability to explain what a ML model has learned, a critical component for medical applications of artificial intelligence (AI), is extremely limited in emerging methods, such as deep learning.

Since FreeSurfer produces such a large array of anatomical measurements (regional volumes, cortical thicknesses, cortical surface area, surface curvature measurements and more), and machine learning typically involves the multivariate analysis of many feature measurements in order to better inform predictions, we hypothesized that FreeSurfer may play a useful role in statistical machine learning-based diagnostic technology. The combination of FreeSurfer measurements with ML has produced papers presenting their leading ML models based on the random forest (Jo et al., 2019; Yassin et al., 2020), the decision tree (Liang et al., 2019) and the support vector machine (de Pierrefeu et al., 2018), including an approach developed on a small cohort focused on only the amygdaloid and hippocampal regions (Guo et al., 2020), as well as methods that combine FreeSurfer with voxel based morphometry (Schwartz et al., 2019). It should also be noted that ML combined with FreeSurfer has also been applied to the prediction of first episode psychosis (Vieira et al., 2020). FreeSurfer's validated pipeline produces a wide array of anatomically meaningful and specific measurements, thus ML models employing feature selection and FreeSurfer-based measurement extraction offer inherent

improvements in explaining what the model has learned, relative to many of the approaches employed in the literature.

In this study, we hypothesize that a thorough analysis of cortical thickness biomarkers, inclusive of a correlation analysis with clinical characteristics in a schizophrenia population, and a machine learning analysis applied to the many biomarkers available in FreeSurfer (Fischl, 2012), may assist in the characterization of schizophrenia.

## Materials and methods

### Participants, data acquisition and preprocessing

Following approval by BCH's Institutional Review Board (informed consent was waived due to the lack of risk to participants included in retrospective analyses), the MCIC medical imaging electronic database (Gollub et al., 2013)<sup>1</sup> was accessed, and all examinations with clinical data available from the MCIC dataset were downloaded for further analysis. Imaging was performed with MRI scanners. The MCIC project was supported by the Department of Energy under Award Number DE-FG02-08ER64581. MCIC is the result of efforts of investigators from the *University of Iowa, University of Minnesota, University of New Mexico, and Massachusetts General Hospital*. Patients were only included if they met the DSM-IV criteria for schizophrenia, schizoaffective disorder or schizophreniform disorder. A major effort was made to only include patients who were antipsychotic drug naïve and were recruited early in the course of their illness. The control participants were selected based on having no history of substance abuse or psychiatric illness, and were matched to the patient cohort for age, sex, and parental education (Gollub et al., 2013). Data used in the preparation of this work were obtained from the Mind Clinical Imaging Consortium database (see text footnote 1) through the Mind Research Network.<sup>2</sup> The detailed protocol descriptions available are in the literature (Gollub et al., 2013).

Each T1 structural examination was processed with FreeSurfer (Fischl, 2012), using the recon-all command which aligns the input examination to all available atlases. All atlases that have cortical thickness measurements were included for further analysis. These combined atlases include definitions of 331 cortical regions in the brain, divided into both the left and right hemispheres. Each FreeSurfer output T1 structural examination was displayed with label map overlays and visually inspected for regional segmentation quality. If FreeSurfer results were observed to substantially fail, they were excluded from

this analysis [i.e., FreeSurfer regions-of-interest (ROIs) that do not align to the MRI and examinations where major problems were observed with an ROI such as a cerebellar segmentation extending far beyond the extent of the cerebellum]. This resulted in a collection of 174 MRI examinations that passed quality control (from 213 examinations publicly accessed), including 99 examinations of patients with schizophrenia, and 75 examinations of healthy control participants.

### Statistical analysis

This study included the acquisition of 662 regionally distributed cortical thickness measurements per imaging examination (mean and standard deviation of within region thicknesses across both left and right hemispheres), as extracted by FreeSurfer's recon-all command, which processes the participant's examination with all available atlases (Fischl, 2012). This included extracting measurements of the average and the standard deviation of within-region cortical thicknesses for each supported gray matter region. Cohen's *d* statistic (positive/negative values indicate a higher/lower average value in the schizophrenic population relative to the neurotypical population) and a *p*-value based on the standard *t*-test (Student, 1908) for two groups of samples were computed to assist in the assessment of group-wise differences between our two populations. The *p*-value was selected as an established method to demonstrate that it is unlikely that our findings were the result of random chance, Cohen's *d* statistic was selected as it is the most established method to assess effect sizes. This yielded a total of  $m = 662$  group-wise comparisons, yielding a Bonferroni corrected threshold for achieving statistical significance of  $p < 0.05/m = 7.55e^{-5}$ .

In order to confirm that the findings reported are the result of group-wise differences between the schizophrenic and neurotypical participants, a statistical model was constructed based on multivariate regression (using MATLAB's mvregress function), adjusting each measurement in order to control for group-wise differences in age, gender and imaging site. This model was used to adjust each cortical thickness (mean and standard deviation) measurement, in order to evaluate whether group-wise differences between our pathological and typically developing populations are the result of age, gender or imaging site effects.

A correlation analysis was performed to assess possible relationships between cortical thickness measurements extracted from MRI examinations and patient clinical variables from the schizophrenia population. In the MCIC dataset, a variety of clinical variables are available for each pathological patient, including, age, gender, parental education levels, a variety of neurological test results, etc. A detailed description of all of the variables available is provided in the literature (Gollub et al., 2013). Each of 662 cortical thickness measurements

<sup>1</sup> <https://coins.trendscenter.org/>

<sup>2</sup> [www.mrn.org](http://www.mrn.org)

were compared with each of the 90 clinical variables available for our schizophrenic population with a correlation analysis, computing Pearson's correlation coefficient and an associated p-value. This resulted in  $m = 662 \times 90 = 59,580$  comparisons, yielding a Bonferroni corrected threshold for achieving statistical significance of  $p < 0.05/m = 8.39e^{-7}$ .

## Machine learning analysis

Five machine learning algorithms are compared in this analysis, the support vector machine (SVM), the decision tree (DT), the random forest (RF), bagged logistic regression (BL), and an artificial neural network (ANN). All examinations that passed quality control were included in the machine learning analysis. All 4,784 feature measurements produced by FreeSurfer's recon-all command were included as potential features to be selected by the machine learning algorithms. Hyperparameter tuning was employed with each algorithm in order to optimize classifier performance using Optuna Bayesian

optimization, which performed extremely well in a recent public competition (Turner et al., 2021). The parameters subjected to Optuna Bayesian optimization for each ML method are as follows: RF – the number of estimators/trees, the criterion (gini vs. entropy), the tree depth, and whether or not to bootstrap; SVM – C the regularization parameter; DT – criterion (gini vs. entropy), tree depth, whether splits are random vs. best; BL – number of estimators, feature down sampling, whether to bootstrap or not; ANN – number of layers, layer size, regularization, batch size, learning rate, learning rate decay. Each of these algorithms were combined with each of three feature selection techniques: ranking features based on Cohen's *d* statistic, Principal Components Analysis (PCA) and stepwise feature selection. The number of features included for each of the feature selection methods was set to each of 10, 50, and 100. Each combination of machine learning algorithm and feature selection technique was evaluated as part of a K-Fold cross validation, with  $K = 5$ , K-Fold cross validation with  $K = 10$ , and Monte-Carlo style bootstrapping, ensuring a large share of examinations were available for model training with held out samples used for testing/model validation. This validation was repeated 100 times. In each run of the validation, the overall accuracy (OA) is computed alongside the area under the receiver operating characteristic curve (AUC), the sensitivity and the specificity. Summary statistics (mean and standard deviation) are computed across validation runs. All machine learning, validation and statistical analyses were performed in python. Public domain software has been made available at <https://github.com/stfxcutables/df-analyze> to facilitate other researchers to (1) reproduce the findings of this study, and (2) to facilitate the application of the ML methods used herein to any given data frame (samples in rows, feature measurements in columns), helping make the addition of Optuna hypertuned ML easy in future studies.

TABLE 1 Effect sizes for regions exhibiting statistically significant group-wise differences in regional average thickness.

Regional measurement	Cohen's <i>d</i>
Left superior temporal average thickness	−0.728
Right temporal pole average thickness	−0.717
Left middle temporal average thickness	−0.712
Right fusiform average thickness	−0.678
Left superior segment of the circular sulcus of the insula average thickness	−0.672
Left inferior temporal average thickness	−0.654
Left hemisphere average thickness	−0.654
Left lateral superior temporal gyrus average thickness	−0.650
Left temporal pole average thickness	−0.645
Left middle occipital gyrus average thickness	−0.636
Right superior temporal average thickness	−0.634
Right anterior transverse collateral sulcus average thickness	−0.631
Left superior temporal sulcus average thickness	−0.630
Right anterior segment of the circular sulcus of the insula average thickness	−0.625
Left inferior segment of the circular sulcus of the insula average thickness	−0.618
Right inferior temporal average thickness	−0.618
Right pars orbitalis average thickness	−0.600
Right middle temporal average thickness	−0.593
Left fusiform average thickness	−0.592
Left Brodmann's area 45 average thickness	−0.592

Negative Cohen's *d* values indicate that the schizophrenic population have a lower average measured value as compared with the healthy patients.

## Results

### Cortical thickness analysis

Many regions demonstrated group-wise differences between participants with schizophrenia and healthy controls. Decreases in average cortical thicknesses with the largest effect sizes were observed in the superior temporal, temporal pole, middle temporal, and fusiform regions. These findings are summarized in Table 1, demonstrating consistently decreased cortical thicknesses in patients with schizophrenia. All findings provided in Table 1 represent raw (not multivariable regression adjusted) effect size calculations, all of which had statistically significant associated raw p-values, as well as statistically significant multivariable adjusted p-values computed with the results of the multivariable regression analysis that controls for the effect of age, gender, and imaging site.

Unlike other cortical thickness studies, which generally focus on average cortical thicknesses, our study has also included intra-regional cortical thickness variability measurements in the form of the standard deviation. Results indicate a common tendency for schizophrenic patients to exhibit increased cortical thickness variability across many sub-regions of the brain. However, the findings are based on small to medium effect sizes that do not reach our stringent standard for statistical significance. The largest Cohen's  $d$  statistics associated with the abnormalities of intra-regional cortical thickness variability was found in the left rostral anterior cingulate ( $d = 0.44$ ), the left pars orbitalis ( $d = 0.38$ ), the triangular part of the left inferior frontal gyrus ( $d = 0.34$ ) and the left rostral middle frontal region ( $d = 0.33$ ).

## Cortical thickness to clinical variable correlation analysis

The correlation analysis considered the schizophrenia population whose examinations passed quality control, thus this analysis helps assess possible relations between automated extractable brain biomarkers and clinical variables in a schizophrenic population. The correlation analysis yielded 156 biomarker measurement and clinical variable pairs that exceeded the Bonferroni correction, or just 0.26% of all pairings between automatically extracted biomarkers and clinical variables. The most common statistically significant correlations established were negative correlations between

patient age and average cortical thicknesses (78/156 statistically significant measurements) of a wide variety of regions, confirming known cortical thinning that progresses with age (e.g., Salat et al., 2004), as well as confirming this effect specifically in schizophrenia (Kubota et al., 2011). The second most common statistically significant correlations were found to be negative correlations between the amount of time a patient takes to complete a clinical assessment task and regional cortical thicknesses (57/156 statistically significant measurements), a finding that may also be age-dependent. The third most common statistically significant correlations were found to be negative correlations between illness duration and regional average cortical thicknesses (17/156 statistically significant measurements), another finding that might be a product of age-dependency, as older patients naturally have thinner cortices and older patients, on average, have had schizophrenia for longer periods of time. We observed a negative correlation between the thickness of the left precentral cortex and the number of errors patients make (Error Score clinical variable;  $\rho = -0.482$ ). Finally, a positive correlation was observed between the Calgary Depression Scale (Addington et al., 1993) total score and the right medial orbitofrontal average cortical thickness ( $\rho = 0.474$ ).

## Machine learning analysis

In Table 2, the mean and standard deviation of OA and AUC, as well as sensitivity and specificity, are provided for each of our top performing combination of machine

TABLE 2 Comparative results of 5 ML algorithms and three feature selection (FS) techniques in terms of overall accuracy (OA), area under the receiver operating characteristic curve (AUC), sensitivity, and specificity.

FS: stepwise	SVM	DT	RF	BL	ANN
OA mean (Std Dev)	0.84 (0.04)	0.75 (0.07)	0.84 (0.06)	0.83 (0.05)	0.74 (0.08)
AUC mean (Std Dev)	0.83 (0.04)	0.75 (0.08)	0.84 (0.05)	0.83 (0.06)	0.74 (0.09)
Sensitivity mean (Std Dev)	0.92 (0.08)	0.79 (0.05)	0.84 (0.11)	0.85 (0.04)	0.79 (0.13)
Specificity mean (Std Dev)	0.74 (0.09)	0.70 (0.12)	0.85 (0.07)	0.81 (0.13)	0.69 (0.22)
# of features	50	100	50	100	50
FS: PCA	SVM	DT	RF	BL	ANN
OA mean (Std Dev)	0.64 (0.11)	0.62 (0.05)	0.63 (0.05)	0.63 (0.07)	0.59 (0.08)
AUC mean (Std Dev)	0.63 (0.11)	0.62 (0.05)	0.62 (0.05)	0.61 (0.08)	0.57 (0.08)
Sensitivity mean (Std Dev)	0.70 (0.16)	0.66 (0.08)	0.68 (0.10)	0.77 (0.04)	0.71 (0.22)
Specificity mean (Std Dev)	0.56 (0.18)	0.58 (0.11)	0.57 (0.12)	0.44 (0.15)	0.44 (0.25)
# of features	10	10	10	50	100
FS: D	SVM	DT	RF	BL	ANN
OA mean (Std Dev)	0.64 (0.12)	0.62 (0.07)	0.64 (0.06)	0.67 (0.05)	0.60 (0.05)
AUC mean (Std Dev)	0.63 (0.12)	0.61 (0.08)	0.64 (0.06)	0.67 (0.06)	0.58 (0.07)
Sensitivity mean (Std Dev)	0.70 (0.10)	0.71 (0.08)	0.67 (0.11)	0.69 (0.07)	0.73 (0.19)
Specificity mean (Std Dev)	0.57 (0.15)	0.51 (0.15)	0.61 (0.16)	0.65 (0.11)	0.43 (0.29)
# of features	100	50	50	50	10

The feature selection (FS) method is listed in the top left. The symbol # denotes the word number.

learning methods and feature selection techniques. Results in [Table 2](#) demonstrate considerable variability in machine learning performance depending on the underlying technology as well as the feature selection technique chosen. The best performing model was the support vector machine (SVM) with stepwise feature selection (OA: mean 84%, standard deviation 4%) limited to 50 features. Similar results were obtained from the SVM with stepwise feature selection with only 10 features (OA: mean 83%, standard deviation 5%). These leading models both based their predictions on the variability of the thickness of the left calcarine sulcus, the volume of the left gyrus rectus, the number of vertices on the left superior frontal gyrus, the average thickness of the left rostral middle frontal region, the signal intensity of the right insula, right superior frontal gray matter, and right caudate, the left hemisphere surface integral, and the irregularity of the left hemisphere's surface reconstruction (in terms of defect holes). All results in [Table 1](#) report validation findings for K-Fold cross validation ( $K = 5$ ) repeated 100 times reporting on the average and standard deviation of each performance metric. Overall, stepwise feature selection was consistently the best performing feature selection method available.

## Discussion

We performed a retrospective analysis of a public schizophrenia dataset including a cortical thickness analysis and a correlation analysis. Our findings from the cortical thickness analysis confirm known average group-wise cortical thinning in schizophrenia ([van Erp et al., 2018](#)) in the superior temporal cortex ([Schultz et al., 2010](#); [van Haren et al., 2011](#)), middle temporal gyrus ([Cui et al., 2018](#)), right fusiform gyrus ([Goghari et al., 2015](#)), inferior temporal cortex ([Buchy et al., 2011](#)), and thinning of the right temporal pole, which was observed in patients at high risk for schizophrenia ([Li et al., 2016](#)). We observed cortical thinning of the left temporal pole, which has previously been reported to be correlated with processing speed ([Hartberg et al., 2011](#)). Additionally, we have observed cortical thinning in the left middle occipital gyrus, which has been reported to be linked with the verbal intelligence quotient (IQ) in schizophrenia ([Hartberg et al., 2010](#)). We were also able to confirm cortical thinning in Brodmann's area 45 ([Narayan et al., 2007](#)). Various studies have reported abnormalities of the insula ([Shepherd et al., 2012](#)), including abnormalities of insular cortical thickness ([Song et al., 2015](#); [Emami et al., 2016](#)), which this study has confirmed alongside reporting thinning of the insular sulcus.

In this study, the majority of the statistically significant correlations between clinical variables and structural cortical thickness measurements were observed between patient age and regional cortical thicknesses that were highly negative correlations. These findings are expected, as it is well

known that cortical thinning (reductions in average cortical thickness) occurs naturally with age ([Fjell et al., 2009](#); [Levman et al., 2017](#)), that schizophrenia populations exhibit increased cortical thinning relative to neurotypical controls ([Kubota et al., 2011](#); [Nenadic et al., 2015](#)), and, specifically, negative correlations have been reported between temporal pole thickness and age in schizophrenia ([van Erp et al., 2018](#)). Furthermore, we observed many (17/156) statistically significant negative correlations between illness duration and regional average cortical thicknesses, a finding in agreement with previously completed research ([van Erp et al., 2018](#)). Of potentially more interest, findings include a positive correlation between the Calgary Depression Scale For Schizophrenia Total Score ([Addington et al., 1993](#)) and the average cortical thickness of the right medial orbitofrontal region, which has been implicated in depression outside of a schizophrenic population ([Drevets, 2007](#); [Cheng et al., 2016](#)). These results are supportive of previous literature findings of observed orbitofrontal abnormalities in schizophrenia ([Waltz and Gold, 2007](#); [Kanahara et al., 2013](#)), and imply that the right medial orbitofrontal region might be directly implicated in the experience of depression among patients with schizophrenia. We observed a negative correlation between the thickness of the left precentral region and the number of errors patients make (Error Score clinical variable;  $\rho = -0.482$ ), which may be supportive of reported associations of the precentral gyrus with No-Go errors ([Stevens et al., 2009](#)) and reported functional MRI-based activation-behavior correlations in the precentral gyrus in Go trials ([Zhang et al., 2019](#)), both in control populations.

Multivariable regression models were included to assess the possible confounding effects of age, gender and imaging site. Results indicated that our primary findings remained statistically significant when controlling for these factors. We elected to present our raw (unadjusted) results rather than our multivariable regression model results for ease of comparison and reproducibility with future studies, and because it has been previously demonstrated that the use of multivariable regression in neuroscience is associated with the introduction of error, and this was specifically demonstrated on this same schizophrenia dataset ([Levman et al., 2021](#)).

We performed a comparative analysis of five machine learning methods and three feature selection methods combined with Optuna hyperparameter tuning toward the automated diagnosis of schizophrenia from T1 structural MRI examinations. The leading model was a support vector machine employing stepwise feature selection, yielding an accuracy of 84%. This leading model based its predictions on a variety of identified abnormalities implicating the right insula, right superior frontal gray matter, right caudate, left calcarine sulcus, left gyrus rectus, left superior frontal gyrus, the left rostral middle frontal region, and whole left hemisphere surface measurements in schizophrenia. Many of these brain

regions have been previously identified as being involved in schizophrenia, including the insula (Wylie and Tregellas, 2010), superior frontal gyrus (Gao et al., 2015; Vogel et al., 2016), caudate (Ebdrup et al., 2010; Williams, 2016), calcarine sulcus (Sulejmanpašić et al., 2016), gyrus rectus (Kim et al., 2017), and the rostral middle frontal region (Kikinis et al., 2010). The results of our ML analysis directly support these historical literature findings. Furthermore, these historical findings lend support to the potential reliability of the leading machine learning model developed as part of the methods from this analysis.

Deep learning is a popular approach to machine learning applications in medical imaging. Research has focused on the use of deep learning on structural MRI exams for the prediction of schizophrenia, achieving widely varying performance (area under the curve ranges from 0.71 to 0.90) on unseen datasets (Oh et al., 2020), implying that current deep learning technologies may not meet the high reliability standards expected in medical applications. This is further supported by recent findings applying deep learning to schizophrenia diagnostics, which achieved accuracy of 70% at best (Vieira et al., 2020).

Accuracy of reported ML models that rely on more traditional techniques (not deep learning) varies considerably, with techniques reporting accuracies of 62% for ML based on whole brain white matter fractional anisotropy (Mikolas et al., 2018), 79% for ML applied to functional MRI (Gallos et al., 2021), 88% for ML based on voxel-based morphometry on a small dataset (Lu et al., 2016), as well as 75% accuracy on a larger cohort of participants (Salvador et al., 2017). Automated feature extraction technologies, like FreeSurfer (Fischl, 2012), provide detailed feature measurements that are more informative relative to that obtainable from deep learning and voxel based morphometry (VBM) based approaches. The application of ML to FreeSurfer measurements has been the subject of existing work, which has produced papers presenting their leading ML models with accuracies of 76 and 69%, respectively, based on the random forest (Jo et al., 2019; Yassin et al., 2020), 72% based on the support vector machine (de Pierrefeu et al., 2018), 82% based only on the amygdaloid and hippocampal regions in a small cohort of study participants (Guo et al., 2020), and 76% accuracy from combining FreeSurfer with VBM (Schwartz et al., 2019). Unfortunately, the differing datasets employed result in variation in the patient populations and image acquisition quality across studies, making direct comparison of widely accepted performance metrics, such as overall accuracy, extremely difficult. Thus, directly comparing two OA values from differing studies should be done with great caution. However, it is noted that our approach, reporting OA of 84%, exhibits strong performance based on FreeSurfer (Fischl, 2012) extracted measurements, and the closest performing model (Guo et al., 2020), with 82% OA, was based on measurements extracted from just the amygdaloid and hippocampal regions

and was based on a small patient population. Our approach achieves the results outlined with biomarkers extracted from many regions, including the insula, superior frontal region, etc. FreeSurfer based methods have considerably improved potential to inform clinicians of the brain features that contribute to a patient being diagnosed with schizophrenia.

From a technical perspective, the best performing ML technique investigated was the support vector machine (SVM) combined with stepwise feature selection. The SVM is a classical ML approach that attempts to minimize error on unseen samples, which is a statistical sampling dependent approach that is capable of providing substantial reliability improvements over many ML techniques. Additionally, the SVM operates by defining a boundary between predictions based on training samples located near the boundary itself, providing potentially statistically robust solutions to the discrimination/prediction problem. Artificial neural networks (ANNs) perform feature weighting and are based on extensive random initialization, which potentially contributes to variability in performance across validation runs. The decision tree and the random forest do not weight their individual features, however, some features are prioritized over others based on their respective positions in the resultant decision trees associated with these methods. The support vector machine does not perform feature weighting, however, its robust sample down sampling strategy can sometimes result in a minimization of the effects of nuisance features. Employing feature selection with limits on the number of features helps to ensure that models do not rely on excessive numbers of spurious features to inform prediction, thus potentially contributing to improvements in ML reliability.

Stepwise feature selection was the best performing of the three feature selection methods considered in this analysis. Stepwise feature selection involves repeatedly adding feature measurements to the machine learning model based on which measurement is anticipated to provide the most improvement to the accuracy of the resultant technology. Unfortunately, stepwise feature selection is by far the slowest method of those evaluated. Stepwise feature selection does allow the user to select the number of features included in the ML model, which when set appropriately low, can assist in the creation of reliable artificial intelligence (AI) technologies. It is noteworthy that stepwise feature selection outperformed the other two feature selection methods (PCA and ranking with Cohen's *d* statistic) for all five machine learning algorithms assessed. PCA tries to rotate the high dimensional data space in order to identify salient projections that act as rotated features. It is a popular dimensionality reduction technique that can be used to reduce the number of features by keeping only those rotated measurements that allow us to retain as much variance in the dataset as possible, and thus hopefully retains as much discriminatory information as possible. Ideally, the technique will keep rotated features that result in large interclass distances and small intraclass variance in the feature space. It is also

desirable to assess correlations between features, as is done in PCA, so that any redundant information can be removed. The basic idea in PCA is to find the linearly transformed components that can explain the maximum possible amount of variance in the dataset. Unfortunately, the input dataspace is very large in terms of input features and sparsely populated, which helps explain why the technique is not functioning reliably. Cohen's  $d$  statistic represents an established method for assessing a feature measurement's effect size. Ranking our features with Cohen's  $d$  statistic and providing the leading set of features to the machine learning algorithm produced results of intermediate quality in this application.

The limitations of the study include that disease sub-status was not included as a variable in the analyses, thus no group-wise comparisons nor statistical models were constructed to control for issues such as whether a given patient had experienced chronic or first episode schizophrenia, respectively. Additional limitations include that the study was performed on a modest sized publicly available dataset. Strengths of the study include a correlation analysis, as well as a machine learning analysis, and open source machine learning code to facilitate the reproducibility of our study findings.

In conclusion, our analysis included the observation of correlations between cortical thickness of the right medial orbitofrontal cortex and patient depression in schizophrenia, as well as provided extensive sets of results from machine learning, indicating good performance for predicting the condition from computer code that we have made publicly available. Future work is needed to further optimize the performance of machine learning in this application. Future work will investigate the extent of improvements to machine learning diagnostics attainable by dramatically increasing the sample size and investigating the use of deep learning based convolutional neural networks. It is hoped that as new technologies in explainable deep learning become available, that methods such as those presented in this paper can assist with comparative validation between models to ensure sensible behavior from emergent deep learning technologies. All underlying features and trained models need to be subjected to rigorous validation on independently acquired datasets. We will also extend the analysis to include additional MRI modalities, such as tractography and functional MRI, as well as expanding our analyses to larger datasets. It is hoped that reliable ML models based on validated FreeSurfer measurement extraction technology can act as comparative models to help validate emerging explainable deep learning models as part of future work as well. Future work will also perform additional analyses to control for the effects of disease sub-status, which was not included as a variable in our analyses, in order to assess the effects associated with whether a given patient had experienced chronic or first episode schizophrenia, respectively. It is hoped that these research avenues will assist toward better understanding schizophrenia as well as improved characterization, diagnosis and classification of the disorder.

## Data availability statement

The data relied upon in this study is publicly available. Data used in the preparation of this work were obtained from the Mind Clinical Imaging Consortium database through the Mind Research Network: <https://coins.trendscenter.org/>.

## Ethics statement

The studies involving human participants were reviewed and approved by Boston Children's Hospital REB. Note this is a publicly available dataset so no REB permission is required. The patients/participants provided their written informed consent to participate in this study.

## Author contributions

JL: conceived the study. JL, ET, and IG: study supervision and feedback on and editing the manuscript. MJ, ER, and DB: software development. JL, PK, and MN: authoring the manuscript. All authors contributed to the article and approved the submitted version.

## Funding

This work was supported by the National Institutes of Health (grant numbers R01HD078561 and R03NS091587) to ET; Natural Science and Engineering Research Council of Canada's Canada Research Chair grant (grant number 231266) to JL, a Canada Foundation for Innovation and Nova Scotia Research and Innovation Trust infrastructure grant (R0176004) to JL, and a St. Francis Xavier University research startup grant to JL (grant number R0168020).

## Acknowledgments

We would like to thank Henry Feldman, Principal Biostatistician at Boston Children's Hospital for advice on conducting statistical analyses.

## Conflict of interest

JL was founder of Time Will Tell Technologies, Inc.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated

organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

- Addington, D., Addington, J., and Maticka-Tyndale, E. (1993). Assessing Depression in Schizophrenia: The Calgary Depression Scale. *Br. J. Psychiatry* 163, 39–44.
- Arnold, S. J., Ivleva, E. I., Gopal, T. A., Reddy, A. P., Jeon-Slaughter, H., Sacco, C. B., et al. (2015). Hippocampal Volume is Reduced in Schizophrenia and Schizoaffective Disorder But Not in Psychotic Bipolar I Disorder Demonstrated by Both Manual Tracing and Automated Parcellation (FreeSurfer). *Schizophr. Bull.* 41, 233–249. doi: 10.1093/schbul/sbu009
- Buchy, L., Ad-Dab'bagh, Y., Malla, A., Lepage, C., Bodnar, M., Joobar, R., et al. (2011). Cortical thickness is associated with poor insight in first-episode psychosis. *J. Psychiatr. Res.* 45, 781–787.
- Cheng, W., Rolls, E. T., Qiu, J., Liu, W., Tang, Y., Huang, C. C., et al. (2016). Medial reward and lateral non-reward orbitofrontal cortex circuits change in opposite directions in depression. *Brain* 139, 3296–3309. doi: 10.1093/brain/aww255
- Chyzyk, D., Graña, M., Öngür, D., and Shinn, A. K. (2015). Discrimination of Schizophrenia Auditory Hallucinators by Machine Learning of Resting-State Functional MRI. *Int. J. Neural Syst.* 25:1550007.
- Cui, Y., Liu, B., Song, M., Lipnicki, D. M., Li, J., Xie, S., et al. (2018). Auditory verbal hallucinations are related to cortical thinning in the left middle temporal gyrus of patients with schizophrenia. *Psychol. Med.* 48, 115–122.
- de Fillippis, R., Carbone, E. A., Gaetano, R., Bruni, A., Pugliese, V., Segura-Garcia, C., 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. doi: 10.2147/NDT.S202418
- de Pierrefeu, A., Löfstedt, T., Laidi, C., Hadj-Selem, F., Bourgin, J., Hajek, T., et al. (2018). Identifying a neuroanatomical signature of schizophrenia, reproducible across sites and stages, using machine learning with structured sparsity. *Acta Psychiatr. Scand.* 138, 571–580. doi: 10.1111/acps.12964
- Drevets, W. (2007). Orbitofrontal cortex function and structure in depression. *Ann. N.Y. Acad. Sci.* 1121, 499–527.
- Ebdrup, B. H., Glenthøj, B., Rasmussen, H., Aggernaes, B., Langkilde, A. R., Paulson, O. B., et al. (2010). Hippocampal and caudate volume reductions in antipsychotic-naïve first-episode schizophrenia. *J. Psychiatry Neurosci.* 35, 95–104. doi: 10.1503/jpn.090049
- Emami, S., Guimond, S., Chakravarty, M. M., and Lepage, M. (2016). Cortical thickness and low insight into symptoms in enduring schizophrenia. *Schizophr. Res.* 170, 66–72. doi: 10.1016/j.schres.2015.10.016
- Fischl, B. (2012). FreeSurfer. *NeuroImage* 62, 774–781.
- Fjell, A. M., Westlye, L. T., Amlien, I., Espeseth, T., Reinvang, I., Raz, N., et al. (2009). High Consistency of Regional Cortical Thinning in Aging across Multiple Samples. *Cereb. Cortex* 19, 2001–2012.
- Gallos, I. K., Gkiatis, K., Matsopoulos, G. K., Siettos, C. (2021). ISOMAP and machine learning algorithms for the construction of embedded functional connectivity networks of anatomically separated brain regions from resting state fMRI data of patient with Schizophrenia. *AIMS Neurosci.* 8, 295–321. doi: 10.3934/Neuroscience.2021016
- Gao, B., Wang, Y., Liu, W., Chen, Z., Zhou, H., Yang, J., et al. (2015). Spontaneous Activity Associated with Delusions of Schizophrenia in the Left Medial Superior Frontal Gyrus: A Resting-State fMRI Study. *PLoS One* 10:e0133766. doi: 10.1371/journal.pone.0133766
- Goghari, V. M., Truong, W., and Spilka, M. J. (2015). A magnetic resonance imaging family study of cortical thickness in schizophrenia. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* 168, 660–668.
- Gollub, R., Shoemaker, J. M., King, M. D., White, T., Ehrlich, S., Sponheim, S. R., et al. (2013). The MCIC collection: A shared repository of multi-modal, multi-site brain image data from a clinical investigation of schizophrenia. *Neuroinformatics* 11, 367–388. doi: 10.1007/s12021-013-9184-3
- Gould, I. C., Shepherd, A. M., Laurens, K. R., Cairns, M. J., Carr, V. J., Green, M. J., et al. (2014). Multivariate neuroanatomical classification of cognitive subtypes in schizophrenia: A support vector machine learning approach. *Neuroimage Clin.* 6, 229–236. doi: 10.1016/j.nicl.2014.09.009
- Greenstein, D., Malley, J. D., Weisinger, B., Clasen, L., and Gogtay, N. (2012). Using multivariate machine learning methods and structural MRI to classify childhood onset schizophrenia and healthy controls. *Front. Psychiatry* 3:53. doi: 10.3389/fpsy.2012.00053
- Guo, Y., Qiu, J., and Lu, W. (2020). Support Vector Machine-Based Schizophrenia Classification Using Morphological Information from Amygdaloid and Hippocampal Subregions. *Brain Sci.* 10:562. doi: 10.3390/brainsci10080562
- Hartberg, C. B., Lawyer, G., Nyman, H., Jönsson, E. G., Haukvik, U. K., Saetre, P., et al. (2010). Investigating relationships between cortical thickness and cognitive performance in patients with schizophrenia and healthy adults. *Psychiatry Res.* 182, 123–133.
- Hartberg, C. B., Sundet, K., Rimol, L. M., Haukvik, U. K., Lange, E. H., Nesvåg, R., et al. (2011). Brain Cortical Thickness and Surface Area Correlates of Neurocognitive Performance in Patients with Schizophrenia, Bipolar Disorder, and Healthy Adults. *J. Int. Neuropsychol. Soc.* 17, 1080–1093. doi: 10.1017/S1355617711001081
- Haug, J. O. (1962). Pneumoencephalographic studies in mental disease. *Acta Psychiatr. Scand. Suppl.* 38:104.
- Haukvik, U. K., Hartberg, C. B., and Agartz, I. (2013). Schizophrenia—what does structural MRI show?. *Tidsskr Nor Laegeforen* 133, 850–853. doi: 10.4045/tidsskr.12.1084
- Haukvik, U. K., Westlye, L. T., Murch-Johnsen, L., Jorgensen, K. N., Lange, E. H., Dale, A. M., et al. (2015). In vivo hippocampal subfield volumes in schizophrenia and bipolar disorder. *Biol. Psychiatry* 77, 581–588.
- Iwabuchi, S. J., Liddle, P. F., and Palaniyappan, L. (2013). Clinical utility of machine-learning approaches in schizophrenia: Improving diagnostic confidence for translational neuroimaging. *Front. Psychiatry* 4:95. doi: 10.3389/fpsy.2013.00095
- Jacobi, W., and Winkler, H. (1928). Encephalographische Studien an Schizophrenen. *Archiv. f. Psychiatrie* 84, 208–226. doi: 10.1007/BF01814508
- Jo, Y. T., Joo, S. W., Shon, S. H., Kim, H., Kim, Y., Lee, J., et al. (2019). Diagnosing schizophrenia with network analysis and a machine learning method. *Int. J. Methods Psychiatr. Res.* 29:e1818. doi: 10.1002/mpr.1818
- Johnstone, E. C., Crow, T. J., Frith, C. D., Husband, J., and Kreel, L. (1976). Cerebral ventricular size and cognitive impairment in chronic schizophrenia. *Lancet* 2, 924–926.
- Kambeitz, J., Kambeitz-Ilanovic, L., Leucht, S., Wood, S., Davatzikos, C., Malchow, B., et al. (2015). Detecting Neuroimaging Biomarkers for Schizophrenia: A Meta-Analysis of Multivariate Pattern Recognition Studies. *Neuropsychopharmacology* 40, 1742–1751. doi: 10.1038/npp.2015.22
- Kanahara, N., Sekine, Y., Haraguchi, T., Uchida, Y., Hashimoto, K., Shimizu, E., et al. (2013). Orbitofrontal cortex abnormality and deficit schizophrenia. *Schizophr. Res.* 143, 246–252.
- Kikinis, Z., Fallon, J. H., Niznikiewicz, M., Nestor, P., Davidson, C., Bobrow, L., et al. (2010). Gray matter volume reduction in rostral middle frontal gyrus in patients with chronic schizophrenia. *Schizophr. Res.* 123, 153–159. doi: 10.1016/j.schres.2010.07.027
- Kim, G.-W., Kim, Y. H., and Jeong, G. W. (2017). Whole brain volume changes and its correlation with clinical symptom severity in patients with schizophrenia: A DARTEL-based VBM study. *PLoS One* 12:e0177251. doi: 10.1371/journal.pone.0177251

- Kim, J. S., Chung, C. K., Jo, H. J., Lee, J. M., and Kwon, J. S. (2012). Regional thinning of cerebral cortical thickness in first-episode and chronic schizophrenia. *Int. J. Imaging Syst. Technol.* 22, 73–80. doi: 10.1002/ima.22002
- Kubicki, M., Park, H., Westin, C. F., Nestor, P. G., Mulkern, R. V., Maier, S. E., et al. (2005). DTI and MTR abnormalities in schizophrenia: Analysis of white matter integrity. *NeuroImage* 26, 1109–1118. doi: 10.1016/j.neuroimage.2005.03.026
- Kubota, M., Miyata, J., Yoshida, H., Hirao, K., Fujiwara, H., Kawada, R., et al. (2011). Age-related cortical thinning in schizophrenia. *Schizophr. Res.* 125, 21–29.
- Laidi, C., d'Albis, M. A., Wessa, M., Linke, J., Phillips, M. L., Delavest, M., et al. (2015). Cerebellar volume in schizophrenia and bipolar I disorder with and without psychotic features. *Acta Psychiatr. Scand.* 131, 223–233.
- Levman, J., Forgeron, C., Shiohama, T., MacDonald, P., Stewart, N., Lim, A., et al. (2021). Hold-out validation for the assessment of stability and reliability of multivariable regression demonstrated with magnetic resonance imaging of patients with schizophrenia. *Int. J. Dev. Neurosci.* 81, 655–662. doi: 10.1002/jdn.10144
- Levman, J., MacDonald, P., Lim, A. R., Forgeron, C., and Takahashi, E. (2017). A Pediatric Structural MRI Analysis of Healthy Brain Development From Newborns to Young Adults. *Hum. Brain Mapp.* 38, 5931–5942.
- Li, G., Wang, L., Shi, F., Lyall, A. E., Ahn, M., Peng, Z., et al. (2016). Cortical thickness and surface area in neonates at high risk for schizophrenia. *Brain Struct. Funct.* 221, 447–461. doi: 10.1007/s00429-014-0917-3
- Liu, X., Lai, Y., Wang, X., Hao, C., Chen, L., Zhou, Z., et al. (2014). A combined DTI and structural MRI study in medicated-naive chronic schizophrenia. *Magn. Reson. Imaging* 32, 1–8.
- Liang, S., Li, Y., Zhang, Z., Kong, X., Wang, Q., Deng, W., et al. (2019). Classification of First-Episode Schizophrenia Using Multimodal Brain Features: A Combined Structural and Diffusion Imaging Study. *Schizophr. Bull.* 45, 591–599.
- Lu, X., Yang, Y., Wu, F., Gao, M., Xu, Y., Zhang, Y., et al. (2016). Discriminative analysis of schizophrenia using support vector machine and recursive feature elimination on structural MRI images. *Medicine* 95:e3973. doi: 10.1097/MD.0000000000003973
- Mikolas, P., Hlinka, J., Skoch, A., Pitra, Z., Frodl, T., Spaniel, F., et al. (2018). Machine learning classification of first-episode schizophrenia spectrum disorders and controls using whole brain white matter fractional anisotropy. *BMC Psychiatry* 18:97. doi: 10.1186/s12888-018-1678-y
- Narayan, V. M., Narr, K. L., Kumari, V., Woods, R. P., Thompson, P. M., Toga, A. W., et al. (2007). Regional Cortical Thinning in Subjects With Violent Antisocial Personality Disorder or Schizophrenia. *Am. J. Psychiatry* 164, 1418–1427.
- Nenadic, I., Yotter, R. A., Sauer, H., and Gaser, C. (2015). Patterns of cortical thinning in different subgroups of schizophrenia. *Br. J. Psychiatry* 206, 479–483.
- Nieuwenhuis, M., van Haren, N. E., Hulshoff Pol, H. E., Cahn, W., Kahn, R. S., and Schnack, H. G. (2012). Classification of schizophrenia patients and healthy controls from structural MRI scans in two large independent samples. *NeuroImage* 61, 606–612. doi: 10.1016/j.neuroimage.2012.03.079
- Nordström, A.-L., and Williamson, P. (2003). Structural neuroimaging in schizophrenia. *Acta Psychiatr. Scand.* 108, 321–323. doi: 10.1034/j.1600-0447.2003.00211.x
- Oh, J., Oh, B. L., Lee, K. U., Chae, J. H., and Yun, K. (2020). Identifying Schizophrenia Using Structural MRI With a Deep Learning Algorithm. *Front. Psychiatry* 11:16. doi: 10.3389/fpsy.2020.00016
- Rich, A. M., Cho, Y. T., Tang, Y., Savic, A., Krystal, J. H., Wang, F., et al. (2016). Amygdala volume is reduced in early course schizophrenia. *Psychiatry Res. Neuroimaging* 250, 50–60. doi: 10.1016/j.psychres.2016.02.006
- Salat, D. H., Buckner, R. L., Snyder, A. Z., Greve, D. N., Desikan, R. S., Busa, E., et al. (2004). Thinning of the cerebral cortex in aging. *Cereb. Cortex* 14, 721–730.
- Salvador, R., Radua, J., Canales-Rodríguez, E. J., Solanes, A., Sarró, S., Goikolea, J. M., et al. (2017). Evaluation of machine learning algorithms and structural features for optimal MRI-based diagnostic prediction in psychosis. *PLoS One* 12:e0175683. doi: 10.1371/journal.pone.0175683
- Schnack, H. G., Nieuwenhuis, M., van Haren, N. E., Abramovic, L., Scheewe, T. W., Brouwer, R. M., 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. doi: 10.1016/j.neuroimage.2013.08.053
- Schultz, C. C., Koch, K., Wagner, G., Roebel, M., Schachtzabel, C., Gaser, C., et al. (2010). Reduced cortical thickness in first episode schizophrenia. *Schizophr. Res.* 116, 204–209.
- Schwartz, E., Doan, N. T., Pergola, G. L., and Westlye, T. (2019). Reproducible grey matter patterns index a multivariate, global alteration of brain structure in schizophrenia and bipolar disorder. *Transl. Psychiatry* 9:12. doi: 10.1038/s41398-018-0225-4
- Shenton, M. E., Dickey, C. C., Frumin, M., and McCarley, R. W. (2001). A review of MRI findings in schizophrenia. *Schizophr. Res.* 49, 1–52.
- Shepherd, A. M., Matheson, S. L., Laurens, K. R., Carr, V. J., and Green, M. J. (2012). Systematic Meta-Analysis of Insula Volume in Schizophrenia. *Biol. Psychiatry* 72, 775–784.
- Singh, S., Khushu, S., Kumar, P., Goyal, S., Bhatia, T., Deshpande, S. N., et al. (2018). Evidence for regional hippocampal damage in patients with schizophrenia. *Neuroradiology* 60, 199–205.
- Song, X., Quan, M., Lv, L., Li, X., Pang, L., Kennedy, D., et al. (2015). Decreased cortical thickness in drug naïve first episode schizophrenia: In relation to serum levels of BDNF. *J. Psychiatr. Res.* 60, 22–28. doi: 10.1016/j.jpsychires.2014.09.009
- Southard, E. E. (1910). A study of the dementia præcox group in the light of certain cases showing anomalies or sclerosis in particular brain-regions. *Am. J. Psychiatry* 67, 119–176. doi: 10.1176/ajp.67.1.119
- Southard, E. E. (1915). On the topographical distribution of cortex lesions and anomalies in dementia præcox, with some account of their functional significance. *Am. J. Psychiatry* 71, 603–671. doi: 10.1176/ajp.71.3.603
- Steardo, L., Carbone, E. A., de Filippis, R., Pisanu, C., Segura-García, C., Squassina, A., et al. (2020). Application of Support Vector Machine on fMRI Data as Biomarkers in Schizophrenia Diagnosis: A Systematic Review. *Front. Psychol.* 11:588. doi: 10.3389/fpsy.2020.00588
- Stevens, M. C., Kiehl, K. A., Pearlson, G. D., and Calhoun, V. D. (2009). Brain network dynamics during error commission. *Hum. Brain Mapp.* 30, 24–37.
- Sulejmanpašić, G., Suljić, E., and Šabanagić-Hajrić, S. (2016). Occipital sulci patterns in patients with schizophrenia and migraine headache using magnetic resonance imaging (MRI). *Med. Glas.* 13, 95–102.
- Turner, R., Eriksson, D., McCourt, M., Kiili, J., Laaksonen, E., Xu, Z., et al. (2021). Bayesian Optimization is Superior to Random Search for Machine Learning Hyperparameter Tuning: Analysis of the Black-Box Optimization Challenge 2020. *arXiv [Preprint]*. doi: 10.48550/arXiv.2104.10201
- van Erp, T. G., Greve, D. N., Rasmussen, J., Turner, J., Calhoun, V. D., Young, S., et al. (2014). A multi-scanner study of subcortical brain volume abnormalities in schizophrenia. *Psychiatry Res.* 222, 10–16. doi: 10.1016/j.psychres.2014.02.011
- van Erp, T. G. M., Walton, E., Hibar, D. P., Schmaal, L., Jiang, W., Glahn, D. C., et al. (2018). Cortical Brain Abnormalities in 4474 Individuals With Schizophrenia and 5098 Control Subjects via the Enhancing Neuro Imaging Genetics Through Meta Analysis (ENIGMA) Consortium. *Biol. Psychiatry* 84, 644–654. doi: 10.1016/j.biopsych.2018.04.023
- van Haren, N. E., Schnack, H. G., Cahn, W., van den Heuvel, M. P., Lepage, C., Collins, L., et al. (2011). Changes in Cortical Thickness During the Course of Illness in Schizophrenia. *JAMA Psychiatry* 68, 871–880.
- Veronese, E., Castellani, U., Peruzzo, D., Bellani, M., and Brambilla, P. (2013). Machine Learning Approaches: From Theory to Application in Schizophrenia. *Comput. Math. Methods Med.* 2013:867924
- Vieira, S., Gong, Q. Y., Pinaya, W. H. L., Scarpazza, C., Tognin, S., Crespo-Facorro, B., et al. (2020). Using Machine Learning and Structural Neuroimaging to Detect First Episode Psychosis: Reconsidering the Evidence. *Schizophr. Bull.* 46, 17–26. doi: 10.1093/schbul/sby189
- Vogel, T., Smieskova, R., Schmidt, A., Walter, A., Harrisberger, F., Eckert, A., et al. (2016). Increased superior frontal gyrus activation during working memory processing in psychosis: Significant relation to cumulative antipsychotic medication and to negative symptoms. *Schizophr. Res.* 175, 20–26. doi: 10.1016/j.schres.2016.03.033
- Waltz, J. A., and Gold, J. M. (2007). Probabilistic reversal learning impairments in schizophrenia: Further evidence of orbitofrontal dysfunction. *Schizophr. Res.* 93, 296–303. doi: 10.1016/j.schres.2007.03.010
- Williams, M. (2016). An introduction to the caudate in schizophrenia. *CNS J.* 2, 40–42.
- Wu, C.-H., Hwang, T.-J., Chen, Y.-J., Hsu, Y.-C., Lo, Y.-C., Liu, C.-M., et al. (2015). Altered integrity of the right arcuate fasciculus as a trait marker of schizophrenia: A sibling study using tractography-based analysis of the whole brain. *Hum. Brain Mapp.* 36, 1065–1076. doi: 10.1002/hbm.22686

- Wylie, K. P., and Tregellas, J. R. (2010). The Role of the Insula in Schizophrenia. *Schizophr. Res.* 123, 93–104. doi: 10.1016/j.jad.2017.11.026
- Yang, H., Liu, J., Sui, J., Pearson, G., and Calhoun, V. D. (2010). A hybrid machine learning method for fusing fMRI and genetic data: Combining both improves classification of schizophrenia. *Front. Hum. Neurosci.* 4:192. doi: 10.3389/fnhum.2010.00192
- Yassin, W., Nakatani, H., Zhu, Y., Kojima, M., Owada, K., Kuwabaras, H., et al. (2020). Machine-learning classification using neuroimaging data in schizophrenia, autism, ultra-high risk and first-episode psychosis. *Transl. Psychiatry* 10:278. doi: 10.1038/s41398-020-00965-5
- Yasuda, Y., Okada, N., Nemoto, K., Fukunaga, M., Yamamori, H., Ohi, K., et al. (2020). Brain morphological and functional features in cognitive subgroups of schizophrenia. *Psychiatry Clin. Neurosci.* 74, 191–203.
- Zalesky, A., Fornito, A., Egan, G. F., Pantelis, C., and Bullmore, E. T. (2012). The relationship between regional and inter-regional functional connectivity deficits in schizophrenia. *Hum. Brain Mapp.* 33, 2535–2549. doi: 10.1002/hbm.21379
- Zhang, H., Eppes, A., and Diaz, M. T. (2019). Task difficulty modulates age-related differences in the behavioral and neural bases of language production. *Neuropsychologia* 124, 254–273. doi: 10.1016/j.neuropsychologia.2018.11.017