Classification of First-Episode Schizophrenia Using Multimodal Brain Features: A Combined Structural and Diffusion Imaging Study

Sugai Liang[1,2,6], Yinfei Li[1,2,6], Zhong Zhang[3], Xiangzhen Kong[4], Qiang Wang[1], Wei Deng[1,2], Xiaojing Li[1], Liansheng Zhao[1], Mingli Li[1], Yajing Meng[1], Feng Huang[3], Xiaohong Ma[1], Xin-min Li[5], Andrew J. Greenshaw[5], Junming Shao[3], and Tao Li*[,1,2]

[1]Mental Health Centre and Psychiatric Laboratory, State Key Laboratory of Biotherapy, West China Hospital, Sichuan University, Chengdu, Sichuan, China; [2]West China Brain Research Centre, West China Hospital, Sichuan University, Chengdu, Sichuan, China; [3]Big Data Research Center, School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu, China; [4]Language and Genetics Department, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands; [5]Department of Psychiatry, Faculty of Medicine & Dentistry, University of Alberta, Edmonton, AB, Canada

[6]These authors contributed equally to this work.

*To whom correspondence should be addressed; West China Mental Health Centre, West China Hospital, Sichuan University, No. 28th Dianxin Nan Str., Chengdu, Sichuan 610041, China; tel.: 86-28-85423561, fax: 86-28-85422632, e-mail: xuntao26@hotmail.com

Recent neuroanatomical pattern recognition studies have shown some promises for developing an objective neuroimaging-based classification related to schizophrenia. This study explored the feasibility of reliably identifying schizophrenia using single and multimodal multivariate neuroimaging features. Multiple brain measures including regional gray matter (GM) volume, cortical thickness, gyrification, fractional anisotropy (FA), and mean diffusivity (MD) were extracted using fully automated procedures. We used Gradient Boosting Decision Tree to identify the most frequently selected features of each set of neuroanatomical metric and fused multimodal measures. The current classification model was trained and validated based on 98 patients with first-episode schizophrenia (FES) and 106 matched healthy controls (HCs). The classification model was trained and tested in an independent dataset of 54 patients with FES and 48 HCs using imaging data acquired on a different magnetic resonance imaging scanner. Using the most frequently selected features from fused structural and diffusion tensor imaging metrics, a classification accuracy of 75.05% was achieved, which was higher than accuracy derived from a single imaging metric. Most prominent discriminative features included cortical thickness of left transverse temporal gyrus and right parahippocampal gyrus, the FA of left corticospinal tract and right external capsule. In the independent cohort, average accuracy was 76.54%, derived from combined features selected from cortical thickness, gyrification, FA, and MD. These features characterized by GM abnormalities and white matter disruptions have discriminative power with respect to the underlying pathological changes in the brain of individuals having schizophrenia. Our results further highlight the potential advantage of multimodal data fusion for identifying schizophrenia.

Key words: schizophrenia/classification/diffusion tensor imaging/structural magnetic resonance imaging/gradient boosting

Introduction

Schizophrenia is a common and complex mental disorder with neuroimaging alterations. Recent neuroanatomical pattern recognition studies attempted to distinguish individuals with schizophrenia by structural magnetic resonance imaging (sMRI) and diffusion tensor imaging (DTI).[1,2] Applications of cutting-edge machine learning approaches in structural neuroimaging studies have revealed potential pathways to classification of schizophrenia based on regional gray matter volume (GMV) or density or cortical thickness.[3–5] Additionally, cortical folding may have high discriminatory value in correctly identifying symptom severity in schizophrenia.[6] Regional GMV and cortical thickness have also been combined in attempts to differentiate individuals with schizophrenia from healthy controls (HCs).[7] Applications of machine learning algorithms to diffusion imaging data analysis to predict individuals with first-episode schizophrenia (FES) have achieved encouraging accuracy.[8–10] White matter (WM) abnormalities in schizophrenia as estimated by DTI appear to be present in the early stage of the disorder, most likely reflecting the developmental stage of the sample of interest.[11,12] Microstructural

disruptions of WM could represent a hallmark of the neuroanatomical features of the brains of individuals having schizophrenia.[13] Despite these promising results, studies exploring the potential value of exploiting multivariate neuroanatomic measures to classify patients with FES relative to HC subjects are limited.

Different neuroimaging modalities can provide complementary information. The sMRI data provide an excellent window on the overall morphological properties of brain tissues, whereas DTI data reflect microstructural information on features associated with WM.[14] Multimodal fusion could shed further light on the neuronal mechanisms underlying pathophysiological features of the diseased brain.[15] In this study, we hypothesized that both structural and diffusion imaging would distinguish patients with FES from HCs, and that the combination of both modalities to predict schizophrenia would be more reliable than using either modality alone.

The primary aim of this study was to compare the classification performance for each single neuroanatomic metric and the combined multimodal imaging data in the classification of patients with FES and HCs. Second, the brain regions most frequently selected in the classification were assessed to enable interpretation in relation to the most significant discriminative features with respect to underlying neuroimaging biomarkers of schizophrenia. Third, the most highly discriminating regions, combined as a neuroimaging pattern, were confirmed in an independent dataset of patients with schizophrenia and HCs.

Materials and Methods

Participants

A total of 152 patients with schizophrenia and 154 HCs were recruited in this study. All participants were right-handed Chinese: 98 patients with FES (mean age: 24.75 years; 57 male) and 106 matched HCs (mean age: 25.97 years; 52 male) were acquired with 3.0 Tesla (T) Philip MRI machine (Dataset 1). Another 54 patients with FES (mean age: 28.40 years; 18 male) and 48 matched HCs (mean age: 25.25 years; 16 male) were acquired with 3.0 T general electric MRI machine (Dataset 2). Their demographic characteristics are displayed in supplementary tables S1 and S2. Symptom severity was assessed using the Positive and Negative Scale (PANSS). This study was approved by the ethics committee of the West China Hospital, Sichuan University, in accordance to the Declaration of Helsinki. In accord with the terms of ethics approval, written consent was obtained from each participant.

Image Data Acquisition

High resolution T1 and DTI data were acquired from each participant (see the Supplementary material for detailed scanning parameters).

MRI Data Preprocessing and Feature Extraction

sMRI data (T1-weighted image) were processed using FreeSurfer 5.3.0.[16] Reconstructed surfaces were visually inspected for topological defects and manually edited according to the standard FreeSurfer guidelines. Parcellation of the brain surfaces was based on the Desikan–Killiany atlas with 68 cortical regions.[17] Cortical GMV, cortical thickness, and cortical folding (local gyrification index) were computed for each of these respective 68 brain regions and for each neuroimaging metric. All metrics were used as cortical features in the classifier.

DTI data were processed using FSL software (FMRIB Software Library, FMRIB),[18] including motion and eddy current correction, brain extraction, tensor model fitting, and Monteal Neurological Institute normalization of tract-based spatial statistics. Fractional anisotropy (FA) and mean diffusivity (MD) were computed for 48 fiber tracts defined by the Johns Hopkins University DTI-based WM atlas.[19] These fiber tracts values were used as WM features in the classifier. Quality control of DTI data was carried out using DTIPrep (translation < 2 mm, rotation < 0.5 mm).[20]

A general linear model was applied in the case of each neuroimaging metric to control for potential confounding covariations of age and gender,[5,7] subsequently each brain measure was normalized by conversion to z scores.

Feature Selection

After obtaining the variables from each neuroimaging metric, the Gini importance of each feature was computed to identify which neuroanatomical features contributed to the discriminative ability of the classifier.[21,22] The Gini importance as well as Gini feature importance can be used as a general indicator of feature relevance. The formula was based on the work of Louppe et al.[23]

Feature Fusion

Three metrics, GMV, cortical thickness and cortical folding, were derived from gray matter of cortical regions. FA and MD of fiber tracts were calculated from WM. The features were concatenated horizontally, to build a matrix in the dimension of $N_{subjects} by (N_{sMRI} + N_{DTI})$ and then input to the feature selection.

Gradient Boosting Decision Tree

The Gradient Boosting Decision Tree (GBDT) is an improved boosting algorithm for regression and classification problems.[24] The basic theory of GBDT is to produce a prediction model constructed by an ensemble of base learners (decision trees), and each tree is built sequentially using the residuals of previous trees. GBDT includes a slight refinement to the basic gradient boosting

algorithm: after fitting a regression tree to the negative gradient, we reestimate the parameters at the leaves of the tree to minimize the loss as follows:[25]

$$r_{jm} = \arg\min_{r} \sum_{x_i \in R_{jm}} L(y_i, f_{m-1}(x_i) + r)$$

where R_{jm} is the region for leaf j in the mth tree, and r_{jm} is the corresponding parameter. Here, the mean response of y_i is for the most probable class label. The grid search with cross-validation was adopted for the hyper-parameters (see Supplementary material for details).

Application to Neuroimaging Metric

In this study, the most frequently selected features were identified in Dataset 1, and then applied to the independent cohort (Dataset 2). Analyses were computed based on scikit-learn (http://scikit-learn.org/stable/index.html) and SciPy (https://www.scipy.org/).

Training and Validating Classification Model in Dataset 1. To select the most frequent features, the whole dataset was randomly shuffled 100 times. After each shuffle, Gini importance was computed for selecting features, and then a 10-fold cross-validation was conducted to evaluate the performance of the GBDT classifier (see figure 1 for the data analysis schema). Average (standard deviation) of Gini importance for each feature was calculated. Observed ratios ≥ 0.95 were defined as the most frequently selected features in this study. The number of most frequently selected features was identified using different thresholds and final values were chosen as a function of their classification accuracy (see Supplementary figure S1).

The most frequently selected features were randomly shuffled 30 times. After each shuffle, 10-fold cross-validation was conducted. To represent model performance

objectively and conservatively, average classification measures were computed, including mean accuracy, sensitivity, specificity, positive and negative predictive values, and the area under the receiver operating characteristic curve (AUC).

Permutation tests were performed to compare accuracy between the observed group and the results of randomization. We created a distribution by assigning group labels randomly to each of the groups 1000 times and then estimated the difference between groups each time. If the observed difference between 2 groups was within 2.5% on either end of the distribution, we considered the difference to be significant at the 5% level.

Training and Testing Classification Model in Dataset 2. The most frequently selected features identified in Dataset 1 were extracted from Dataset 2. Then features extracted from Dataset 2 were randomly shuffled 30 times. After each shuffle, Dataset 2 was randomly partitioned into 10 equal size subsamples using the *KFold* in scikit-learn. Each fold was used once as a testing set whereas the 9 remaining folds formed the training set. As described earlier, to reflect model performance objectively and conservatively, average classification measures were computed for the independent cohort.

Results

In the automated prediction of FES from HC, using Gini importance and training with GBDT, the most frequently selected features were identified for each respective neuroimaging metric and for fused multimodal features. Table 1 lists the most frequently selected features for FA and gyrification. Higher feature importance, degree of robustness, and the relative importance of brain regions contributed to the resulting predictions. In Dataset 1, based on the most frequently selected 4 fiber tracts of
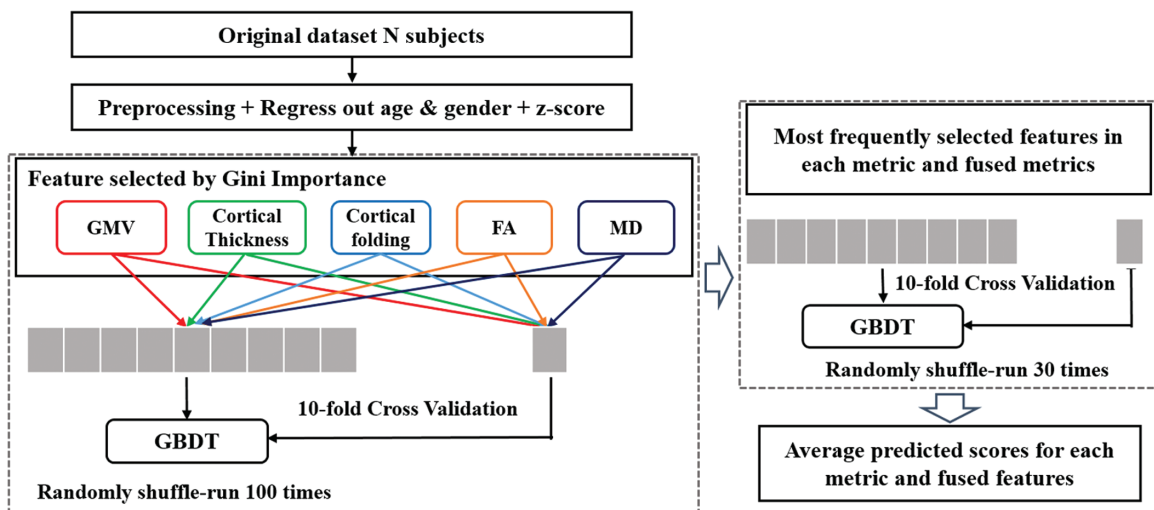

Fig. 1. Data analysis schema. GMV, gray matter volume; FA, fractional anisotropy; MD, mean diffusivity.

FA, cross-validated average accuracy was 70.27% in distinguishing individuals with schizophrenia from HCs. By contrast, in the independent cohort (Dataset 2), classification accuracy was 60.12% based on the 4 fiber tracts of FA. Using Dataset 1, based on the most frequently selected 6 brain regions of gyrification, an average accuracy of 67.93% was obtained. Similar accuracy (63.41%) was achieved in differentiating patients with schizophrenia and HCs in Dataset 2.

Supplementary table S3 presents the most frequently selected features for GMV, cortical thickness, and MD. Seven brain regions from GMV, 5 regions from cortical thickness, and 14 regions from MD were selected as individual metric predictors of schizophrenia. For Dataset 1, based on the most frequently selected features of each metric, cross-validated average accuracy was 63.50% for GMV, 66.47% for cortical thickness, and 66.00% for MD. In Dataset 2, average accuracy was 54.70% for GMV, 60.94% for cortical thickness, and 67.43% for MD.

On the basis of the most frequently selected features of cortical thickness, gyrification, FA, and MD, an accuracy of 75.05% and an AUC of 75.12% were achieved in Dataset 1. The results of this study yielded an accuracy of 76.54% and an AUC of 77.86% based on regions of the 4 fused metrics in Dataset 2. Table 2 presents the most frequently selected features from the fused metrics of cortical thickness, gyrification, FA, and MD. Specifically, the left transverse temporal gyrus, right parahippocampal gyrus, right middle temporal gyrus, right frontal pole, left entorhinal cortex, and right lingual gyrus were identified as potentially key brain areas for schizophrenia classification. For DTI measures, the bilateral corticospinal tract, right external capsule, left tapetum, part of fornix, left superior fronto-occipital fasciculus, left superior longitudinal fasciculus, and right posterior limb of internal capsule were apparently prominent predictive areas. Figure 2 illustrates these brain regions.

Supplementary table S4 shows the most frequently selected features for 3 concatenated gray matter metrics and 2 WM metrics. Based on the most frequently selected features of these 5 fused neuroimaging measures, the classification of FES and HCs was associated with a cross-validated accuracy of 74.14% with respect to Dataset 1. A similar average accuracy (75.30%) was achieved with respect to Dataset 2, based on the brain regions selected from the 5 fused metrics. The classification accuracy of the features selected from integrated metrics of cortical
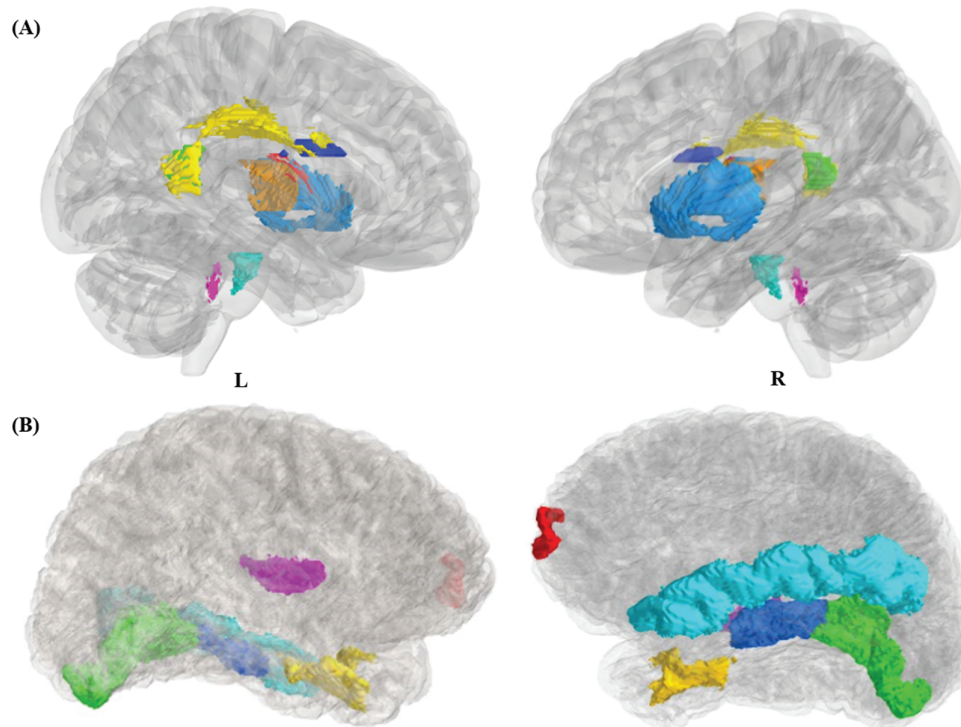
Table 1. Most Frequently Selected Features From Each Single Metric

Metric	Brain Region	Side	Feature Importance
FA	Corticospinal tract	L	0.278 (0.043)
	External capsule	R	0.223 (0.042)
	Superior fronto-occipital fasciculus	L	0.168 (0.023)
	Tapetum	L	0.124 (0.025)
Gyrification	Caudal anterior cingulate	R	0.182 (0.026)
	Cuneus cortex	R	0.178 (0.024)
	Frontal pole	R	0.157 (0.022)
	Isthmus cingulate	L	0.146 (0.016)
	Rostral middle frontal	L	0.116 (0.023)
	Parahippocampal	R	0.107 (0.016)

FA, fractional anisotropy.
Average (standard deviation) of Gini feature importance is listed for each selected feature in single metric. The higher feature importance, the more robustness and importance of the brain regions contribute to the prediction.

Table 2. Most Frequently Selected Features From 4 Fused Metrics

Brain Region	Metric	Side	Feature Importance
Corticospinal tract	FA	L	0.124 (0.016)
Transverse temporal cortex	Cortical thickness	L	0.090 (0.022)
Parahippocampal gyrus	Cortical thickness	R	0.086 (0.0127)
External capsule	FA	R	0.080 (0.012)
Middle temporal gyrus	Gyrification	R	0.075 (0.011)
Tapetum	MD	L	0.063 (0.016)
Corticospinal tract	FA	R	0.059 (0.008)
Fornix (column and body of fornix)	MD	—	0.058 (0.011)
Frontal pole	Cortical thickness	R	0.053 (0.013)
Superior fronto-occipital fasciculus	MD	L	0.054 (0.010)
Entorhinal cortex	Cortical thickness	L	0.051 (0.012)
Superior longitudinal fasciculus	FA	L	0.051 (0.012)
Medial lemniscus	MD	R	0.049 (0.009)
Posterior limb of internal capsule	FA	R	0.045 (0.008)
Lingual gyrus	Cortical thickness	R	0.028 (0.006)

FA, fractional anisotropy; MD, mean diffusivity.
Average (standard deviation) of feature importance is listed for each selected feature in multimodal metrics. The higher feature importance, the more robustness and importance of the brain regions contribute to the prediction. The 4 fused metrics included cortical thickness, gyrification, FA, and MD.

Fig. 2. Most frequently selected brain regions from the 4 fused metrics. (A) fiber tracts in the Monteal Neurological Institute space, corticospinal tract (cyan), tapetum L (green), superior longitudinal fasciculus L (yellow), superior frontal-occipital fasciculus L (blue), fornix (scarlet), external capsule R (light blue), medial lemniscus R (purple), posterior limb of internal capsule R (golden). (B) cortical regions in FreeSurfer space, transverse temporal cortex L (purple), entorhinal cortex L (yellow), parahippocampal gyrus R (blue), middle temporal gyrus R (cyan), frontal pole R (red), lingual gyrus R (green).

thickness, gyrification, FA, and MD was higher than that of the features selected from the 5 fused metrics. Figure 3 illustrates a comparison of the classification accuracy of single and fused neuroimaging metrics. The averages and standard deviations of classification performance measures are presented in Supplementary tables S5 and S6 for Dataset 1 and Dataset 2, respectively.

Based on single modality metrics, classification accuracy in Dataset 2 was lower than in Dataset 1, except the performance of MD. According to permutation tests in Dataset 1, average accuracy in the validation set remains significantly better than chance except for the result of the GMV (see supplementary figure S2). With respect to multimodal metrics, accuracies in Dataset 2 were slightly higher than Dataset 1. The 2-sample t test was conducted to compare the classification accuracies of unimodal and multimodal. The accuracies of multimodal were significantly higher than the those of unimodal ($P < .001$) both in Dataset 1 and Dataset 2.

Discussion

In this study, we applied a data-driven method based on single and multimodal neuroimaging metrics to identify discriminative features of FES. Classification performance based on the discriminative brain regions selected from the fused structural and DTI metrics was relatively

higher and more robust than performance based on features selected from single imaging metrics. Subsequent analysis using an independent cohort confirmed these results. Notably, the most distinctive brain regions from the fused measures of gray matter (cortical thickness and gyrification) and WM (fractional anisotropy and mean diffusivity) revealed a relatively high discriminatory ability in reliably classifying individuals with schizophrenia from HCs, with average accuracy of 76.54% based on the independent sample. In addition to the findings earlier mentioned, most important strength of our current study is that we performed the analysis in relatively 2 large and independent sample sets of patients with FES (majority of them were treatment-naïve) to rule out the confounding factors of chronicity of the illness and treatment effects.

Our study achieved a modest but robust classification accuracy by adopting the GBDT to multimodal brain characteristics. Compared to previous studies in this context,[26,27] although our datasets did not differ for age or gender at the group level in this study, the potentially confounding influence of these covariates were taken into account. Additionally, in this study, brain regions were automatically extracted from the whole-brain atlas and the discriminative brain areas were the most frequently features selected from the multiple repetition of the algorithm. Moreover, to evaluate the model objectively and conservatively,
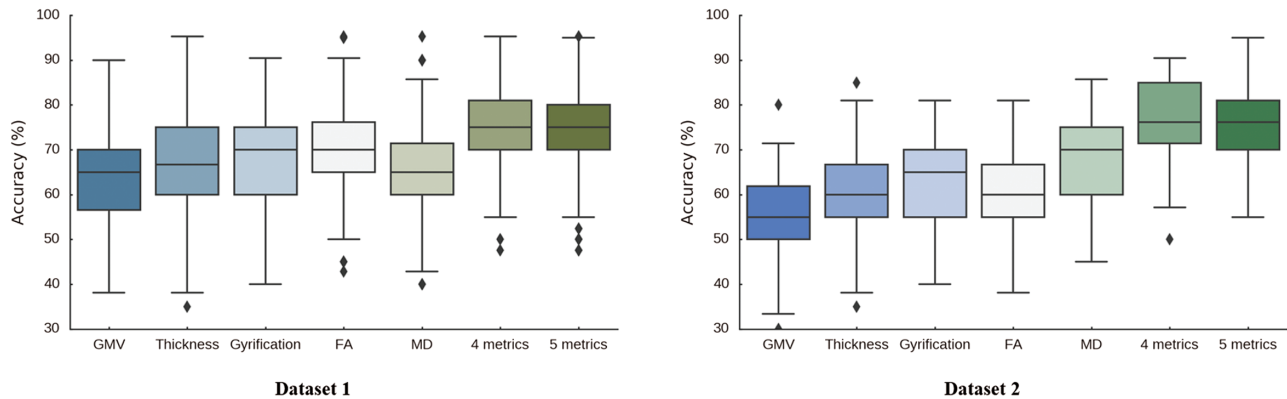
Fig. 3. Comparison for classification accuracy of single metric and multimodal measures in Dataset 1 and Dataset 2. The X-axis represents features from single metric and fused metrics. GMV, gray matter volume; FA, fractional anisotropy; MD, mean diffusivity. The 5 fused metrics included GMV, cortical thickness, gyrification, FA, and MD. The 4 fused metrics included cortical thickness, gyrification, FA and MD. The Y-axis represents accuracy (%).

cross-validated averages of classification performance measures were computed. Validity is further strengthened by the observations that we were able to achieve similar accuracies, when we applied this classification model to an independent cohort yielding multimodal features based on images recorded using a different MRI scanned.

In previous studies, a wide range of accuracy was reported by using different analysis pipeline and methods in various datasets. For example, based on sMRI and DTI, Peruzzo et al[26] reported a 90% accuracy to distinguish schizophrenia with the multiple kernel learning in 23 patients and 23 HCs. Kambeitz et al[28] performed a meta-analysis and found that brain functional and structural alternations could differentiate schizophrenia with the accuracy of 80%. Multisite analysis with support vector machine (SVM) provided a robust neuroanatomical signature of schizophrenia detectable across diverse patient populations.[29] With same algorithm (SVM) but different neuroanatomical pattern, the accuracy of 73.4% was obtained to classify patients with schizophrenia from HCs.[30]

It has been noted that differences in pipelines for image processing, feature selection, feature extraction, dimensionality reduction, and pattern recognition methods, could at least partly explain the above discrepancies across studies.[30] Moreover, the potential factor that might contribute for this is the biological and clinical heterogeneity of schizophrenia.[31] The homogenous small samples may achieve more encouraging accuracies but are prone to have higher risk of overfitting. In this case, it is difficult to generalize the extraordinary accuracy in a specific research dataset to other datasets, especially the clinical ones.[32]

In this study, cortical thickness, gyrification, fractional anisotropy, and mean diffusivity were concatenated as classification features. The integrated characteristics achieved greater classification performance than either metric alone, which probably demonstrates that the 4 measures are complementary when used for predicting schizophrenia from HC. Notably, the mean accuracy

based on the 4 measures was higher and more robust than that of the fused characteristics including cortical GMV, thickness, gyrification, fractional anisotropy, and mean diffusivity. This discrepancy may be related to an intrinsically inappropriate integration of those 5 neuroimaging metrics. Additionally, this may be due to increased noise information merged into the classification model when more neuroimaging measures are included.

In this study, the cortical thickness of transverse temporal gyrus had the highest feature importance for sMRI in the multimodal feature integration. This may suggest that transverse temporal gyrus has more relevance for predictions of schizophrenia. The transverse temporal gyrus as well as Heschl's gyrus located in the area of primary auditory cortex is responsible for processing auditory perception and language preprocessing.[33] Previous studies found individuals with schizophrenia characterized with auditory hallucination had structural abnormalities in the transverse temporal gyrus as altered with thinner cortex and less GMV.[34,35] The possible dysregulation of cortical plasticity was associated with the auditory verbal hallucinations of patients with schizophrenia.[36] In schizophrenia, reduced pyramidal cell body volume and reduced dendritic density in cortical layer 3 has been observed in primary auditory cortex and the auditory association cortex,[37,38] which may reflect neuropathology underlying cortical thinning in patients with auditory hallucinations.

As one of the most discriminative brain areas, cortical thickness of parahippocampal gyrus overlapped in the fused multimodal characteristics and the single sMRI measures. This demonstrates that the parahippocampal gyrus may have greater importance in individually discriminating patients with schizophrenia. As part of the limbic system, the parahippocampal gyrus, connecting with the hippocampus and amygdala, is involved in spatial memory and visual context.[39,40] The right parahippocampal gyrus may have greater involvement in the self-face recognition.[41] Previous studies reported that

patients with schizophrenia had smaller volumes in the parahippocampal gyrus, which could secondarily affect the hippocampus.[42,43] Additionally, the reality distortion in schizophrenia was positively correlated with increased right hippocampal responses to neutral face recognition.[44] In patients with psychotic disorder, dysfunction of the parahippocampal gyrus may also trigger deactivation of right language areas during auditory hallucinations.[45]

In relation to WM assessment, fractional anisotropy of the left corticospinal tract had highest feature importance in multimodal feature integration. Individuals with schizophrenia had greater involuntary movement, referred to as motor overflow, which was likely due to bilateral corticospinal tracts activity.[46] Impaired motor skills associated with FES were related to reduced corticospinal tract and anterior thalamic radiation FA values.[47] This observation may reflect pathognomonic corticospinal tract integrity, related to delayed maturation of WM, specific to adolescent-onset schizophrenia.[48] Auditory hallucinations of schizophrenia were also positively associated with disruption of corticospinal tract integrity, and alteration of this tract may relate to inner speech and efferent copy mechanisms.[49,50]

The external capsule was one of the most frequently selected brain regions in both the fused multimodal features and single DTI measures. The external capsule, situated lateral to the internal capsule, contains association fiber tracts including the superior longitudinal fasciculus and inferior fronto-occipital fasciculus and commissural fibers.[19] In schizophrenia, the impaired WM integrity in the external capsule is most likely associated with reduced axonal myelination and/or alterations in the axonal cytoskeleton.[51] Fractional anisotropy values of right external capsule in patients with schizophrenia correlated significantly with category-completed scores of the Wisconsin Card Sorting Test.[52] WM integrity loss in schizophrenia is also related to pronounced expression of risk genes involved in calcium signaling and in the formation of synapses and protein complexes.[53]

Possible study limitations should be taken into account. First, only sMRI and DTI were combined in the present study. Additional neuroimaging modalities, such as resting state functional MRI, can be integrated and may improve classification performance. Second, because of the computational load, this study extracted information related to neuroimaging atlas defined brain regions, without performing whole brain voxel-wise searching. Third, this study did not include contrasts with other psychiatric disorders. Contrasting schizophrenia from not only HCs but also other disorders, such as bipolar disorder with multimodal neuroimaging, will undoubtedly be useful in defining diagnostic specificity.

In conclusion, based on the structural MRI and DTI data, this study implemented a data-driven method using individual subject's neuroimaging features, to predict whether that individual meets criteria for schizophrenia. Patients with FES could be efficiently distinguished from HCs based on the multimodal feature integration, which

may offer a means of advancing the current diagnosis of schizophrenia. The findings of this study also indicate that schizophrenia is characterized by GM abnormalities and WM disruptions that have discriminative power and may reflect relevant pathological changes in the brain. These results highlight the benefits of multimodal data fusion for identifying schizophrenia.

Supplementary Material

Supplementary data are available at *Schizophrenia Bulletin* online.

Acknowledgments

References

1. Seitz J, Zuo JX, Lyall AE, et al. Tractography analysis of 5 white matter bundles and their clinical and cognitive correlates in early-course schizophrenia. *Schizophr Bull*. 2016;42:762–771.

2. Chang M, Womer FY, Edmiston EK, et al. Neurobiological commonalities and distinctions among three major psychiatric diagnostic categories: a structural MRI study. *Schizophr Bull*. 2018;13:65–74.

3. Borgwardt S, Koutsouleris N, Aston J, et al. Distinguishing prodromal from first-episode psychosis using neuroanatomical single-subject pattern recognition. *Schizophr Bull*. 2013;39:1105–1114.

4. Sun D, van Erp TG, Thompson PM, et al. Elucidating a magnetic resonance imaging-based neuroanatomic biomarker for psychosis: classification analysis using probabilistic brain atlas and machine learning algorithms. *Biol Psychiatry*. 2009;66:1055–1060.

5. Squarcina L, Castellani U, Bellani M, et al; GET UP Group. Classification of first-episode psychosis in a large cohort of patients using support vector machine and multiple kernel learning techniques. *Neuroimage*. 2017;145:238–245.

6. Guo S, Iwabuchi S, Balain V, Feng J, Liddle P, Palaniyappan L. Cortical folding and the potential for prognostic neuroimaging in schizophrenia. *Br J Psychiatry*. 2015;207:458–459.

7. Takayanagi Y, Takahashi T, Orikabe L, et al. Classification of first-episode schizophrenia patients and healthy subjects by automated MRI measures of regional brain volume and cortical thickness. *PLoS One*. 2011;6:e21047.

8. Ingalhalikar M, Kanterakis S, Gur R, Roberts TP, Verma R. DTI based diagnostic prediction of a disease via pattern classification. *Med Image Comput Comput Assist Interv*. 2010;13:558–565.

9. Rathi Y, Shenton ME, Westin C-F. Preliminary findings in diagnostic prediction of schizophrenia using diffusion tensor imaging. In: Westin C-F, Vilanova A, Burgeth B, eds. *Visualization and Processing of Tensors and Higher Order Descriptors for Multi-Valued Data*. Berlin, Heidelberg: Springer Berlin Heidelberg; 2014:313–324.

10. Ardekani BA, Tabesh A, Sevy S, Robinson DG, Bilder RM, Szeszko PR. Diffusion tensor imaging reliably differentiates patients with schizophrenia from healthy volunteers. *Hum Brain Mapp*. 2011;32:1–9.

11. Kyriakopoulos M, Frangou S. Recent diffusion tensor imaging findings in early stages of schizophrenia. *Curr Opin Psychiatry*. 2009;22:168–176.

12. Samartzis L, Dima D, Fusar-Poli P, Kyriakopoulos M. White matter alterations in early stages of schizophrenia: a systematic review of diffusion tensor imaging studies. *J Neuroimaging*. 2014;24:101–110.

13. Li Y, Xie S, Liu B, et al. Diffusion magnetic resonance imaging study of schizophrenia in the context of abnormal neurodevelopment using multiple site data in a Chinese Han population. *Transl Psychiatry*. 2016;6:e715.

14. Xie Y, Cui Z, Zhang Z, et al. Identification of amnestic mild cognitive impairment using multi-modal brain features: a combined structural MRI and diffusion tensor imaging study. *J Alzheimers Dis*. 2015;47:509–522.

15. Sui J, Huster R, Yu Q, Segall JM, Calhoun VD. Function-structure associations of the brain: evidence from multimodal connectivity and covariance studies. *Neuroimage*. 2014;102:11–23.

16. Fischl B. FreeSurfer. *Neuroimage*. 2012;62:774–781.

17. Desikan RS, Ségonne F, Fischl B, et al. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage*. 2006;31:968–980.

18. Jenkinson M, Beckmann CF, Behrens TE, Woolrich MW, Smith SM. FSL. *Neuroimage*. 2012;62:782–790.

19. Mori S, Oishi K, Jiang H, et al. Stereotaxic white matter atlas based on diffusion tensor imaging in an ICBM template. *Neuroimage*. 2008;40:570–582.

20. Oguz I, Farzinfar M, Matsui J, et al. DTIPrep: quality control of diffusion-weighted images. *Front Neuroinform*. 2014;8:4.

21. Ritchie GR, Dunham I, Zeggini E, Flicek P. Functional annotation of noncoding sequence variants. *Nat Methods*. 2014;11:294–296.

22. Hastie T, Tibshirani R, Friedman J, Franklin J. The elements of statistical learning: data mining, inference and prediction. *Mathematical Intelligencer*. 2005;27:83–85.

23. Louppe G, Wehenkel L, Sutera A, Geurts P. Understanding variable importances in forests of randomized trees. *Adv Neural Inf Process Syst*. 2013;431–439.

24. Friedman JH. Greedy function approximation: a gradient boosting machine. *Annals of Statistics*. 2001;29:1189–1232.

25. Friedman JH. Stochastic gradient boosting. *Comput Stat Data Anal* 2002;38:367–378.

26. Peruzzo D, Castellani U, Perlini C, et al; PICOS-Veneto Group. Classification of first-episode psychosis: a multimodal multi-feature approach integrating structural and diffusion imaging. *J Neural Transm (Vienna)*. 2015;122:897–905.

27. Meng X, Jiang R, Lin D, et al. Predicting individualized clinical measures by a generalized prediction framework and multimodal fusion of MRI data. *Neuroimage*. 2017;145:218–229.

28. Kambeitz J, Kambeitz-Ilankovic L, Leucht S, et al. Detecting neuroimaging biomarkers for schizophrenia: a meta-analysis of multivariate pattern recognition studies. *Neuropsychopharmacology*. 2015;40:1742–1751.

29. Rozycki M, Satterthwaite TD, Koutsouleris N, et al. Multisite machine learning analysis provides a robust structural imaging signature of schizophrenia detectable across diverse patient populations and within individuals. *Schizophr Bull*. 2017. doi: 10.1093/schbul/sbx137. [Epub ahead of print.]

30. Zanetti MV, Schaufelberger MS, Doshi J, et al. Neuroanatomical pattern classification in a population-based sample of first-episode schizophrenia. *Prog Neuropsychopharmacol Biol Psychiatry*. 2013;43:116–125.

31. Schnack HG. Improving individual predictions: machine learning approaches for detecting and attacking heterogeneity in schizophrenia (and other psychiatric diseases). *Schizophr Res*. 2017. (In Press) doi: 10.1016/j.schres.2017.10.023.

32. Arbabshirani MR, Plis S, Sui J, Calhoun VD. Single subject prediction of brain disorders in neuroimaging: promises and pitfalls. *Neuroimage*. 2017;145:137–165.

33. Liegeois-Chauvel C, Musolino A, Chauvel P. Localization of the primary auditory area in man. *Brain*. 1991;114:139–151.

34. Mørch-Johnsen L, Nesvåg R, Jørgensen KN, et al. Auditory cortex characteristics in schizophrenia: associations with auditory hallucinations. *Schizophr Bull*. 2017;43:75–83.

35. Gaser C, Nenadic I, Volz HP, Büchel C, Sauer H. Neuroanatomy of "hearing voices": a frontotemporal brain structural abnormality associated with auditory hallucinations in schizophrenia. *Cereb Cortex*. 2004;14:91–96.

36. van Swam C, Federspiel A, Hubl D, et al. Possible dysregulation of cortical plasticity in auditory verbal hallucinations-a cortical thickness study in schizophrenia. *J Psychiatr Res*. 2012;46:1015–1023.

37. Sweet RA, Pierri JN, Auh S, Sampson AR, Lewis DA. Reduced pyramidal cell somal volume in auditory association cortex of subjects with schizophrenia. *Neuropsychopharmacology*. 2003;28:599–609.

38. Sweet RA, Henteleff RA, Zhang W, Sampson AR, Lewis DA. Reduced dendritic spine density in auditory cortex of subjects with schizophrenia. *Neuropsychopharmacology*. 2009;34:374–389.

39. Malkova L, Mishkin M. One-trial memory for object-place associations after separate lesions of hippocampus and posterior parahippocampal region in the monkey. *J Neurosci*. 2003;23:1956–1965.

40. Mégevand P, Groppe DM, Goldfinger MS, et al. Seeing scenes: topographic visual hallucinations evoked by direct electrical stimulation of the parahippocampal place area. *J Neurosci*. 2014;34:5399–5405.

41. van Veluw SJ, Chance SA. Differentiating between self and others: an ALE meta-analysis of fMRI studies of self-recognition and theory of mind. *Brain Imaging Behav*. 2014;8:24–38.

42. Wright IC, Rabe-Hesketh S, Woodruff PW, David AS, Murray RM, Bullmore ET. Meta-analysis of regional brain volumes in schizophrenia. *Am J Psychiatry*. 2000;157:16–25.

43. Sim K, DeWitt I, Ditman T, et al. Hippocampal and parahippocampal volumes in schizophrenia: a structural MRI study. *Schizophr Bull*. 2006;32:332–340.

44. Surguladze S, Russell T, Kucharska-Pietura K, et al. A reversal of the normal pattern of parahippocampal response to neutral and fearful faces is associated with reality distortion in schizophrenia. *Biol Psychiatry*. 2006;60:423–431.

45. Diederen KM, Neggers SF, Daalman K, et al. Deactivation of the parahippocampal gyrus preceding auditory hallucinations in schizophrenia. *Am J Psychiatry*. 2010;167:427–435.

46. Hoy KE, Georgiou-Karistianis N, Laycock R, Fitzgerald PB. Using transcranial magnetic stimulation to investigate the cortical origins of motor overflow: a study in schizophrenia and healthy controls. *Psychol Med*. 2007;37:583–594.

47. Pérez-Iglesias R, Tordesillas-Gutiérrez D, McGuire PK, et al. White matter integrity and cognitive impairment in first-episode psychosis. *Am J Psychiatry*. 2010;167:451–458.

48. Douaud G, Smith S, Jenkinson M, et al. Anatomically related grey and white matter abnormalities in adolescent-onset schizophrenia. *Brain*. 2007;130:2375–2386.

49. Bopp MHA, Zöllner R, Jansen A, Dietsche B, Krug A, Kircher TTJ. White matter integrity and symptom dimensions of schizophrenia: a diffusion tensor imaging study. *Schizophr Res*. 2017;184:59–68.

50. Ćurčić-Blake B, Nanetti L, van der Meer L, et al. Not on speaking terms: hallucinations and structural network disconnectivity in schizophrenia. *Brain Struct Funct*. 2015;220:407–418.

51. Seal ML, Yücel M, Fornito A, et al. Abnormal white matter microstructure in schizophrenia: a voxelwise analysis of axial and radial diffusivity. *Schizophr Res*. 2008;101:106–110.

52. Lee SH, Kubicki M, Asami T, et al. Extensive white matter abnormalities in patients with first-episode schizophrenia: a diffusion tensor iimaging (DTI) study. *Schizophr Res*. 2013;143:231–238.

53. Romme IA, de Reus MA, Ophoff RA, Kahn RS, van den Heuvel MP. Connectome disconnectivity and cortical gene expression in patients with schizophrenia. *Biol Psychiatry*. 2017;81:495–502.