Schizophrenia Bulletin vol. 48 no. 3 pp. 563–574, 2022
https://doi.org/10.1093/schbul/sbac030
Advance Access publication March 30, 2022

Application of a Machine Learning Algorithm for Structural Brain Images in Chronic Schizophrenia to Earlier Clinical Stages of Psychosis and Autism Spectrum Disorder: A Multiprotocol Imaging Dataset Study

Yinghan Zhu[1], Hironori Nakatani[1,2], Walid Yassin[3], Norihide Maikusa[1], Naohiro Okada[4,5], Akira Kunimatsu[6,7], Osamu Abe[6], Hitoshi Kuwabara[8], Hidenori Yamasue[8], Kiyoto Kasai[4,5,9,10], Kazuo Okanoya[1,4,9,10], and Shinsuke Koike*[,1,4,9,10]

[1]Center for Evolutionary Cognitive Sciences, Graduate School of Arts and Sciences, The University of Tokyo, 3-8-1 Komaba, Meguro-ku, Tokyo 153-8902, Japan; [2]Department of Information Media Technology, School of Information and Telecommunication Engineering, Tokai University, 2-3-23, Takanawa, Minato-ku, Tokyo 108-8619, Japan; [3]Department of Child Neuropsychiatry, Graduate School of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan; [4]The International Research Center for Neurointelligence (WPI-IRCN), Institutes for Advanced Study (UTIAS), University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8654, Japan; [5]Department of Neuropsychiatry, Graduate School of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan; [6]Department of Radiology, Graduate School of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan; [7]Department of Radiology, IMSUT Hospital, The Institute of Medical Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan; [8]Department of Psychiatry, Hamamatsu University School of Medicine, 1-20-1 Handayama, Higashi-ku, Hamamatsu City, Shizuoka 431-3192, Japan; [9]University of Tokyo Institute for Diversity & Adaptation of Human Mind (UTIDAHM), 3-8-1 Komaba, Meguro-ku, Tokyo 153-8902, Japan; [10]University of Tokyo Center for Integrative Science of Human Behavior (CiSHuB), 3-8-1 Komaba, Meguro-ku, Tokyo 153-8902, Japan

*To whom correspondence should be addressed; Center for Evolutionary Cognitive Sciences, Graduate School of Arts and Sciences, The University of Tokyo, 3-8-1 Komaba, Meguro-ku, Tokyo 153-8902, Japan; tel: +81-3-5454-4327, fax: +81-3-5454-4327, e-mail: skoike-tky@umin.ac.jp

Background and Hypothesis: **Machine learning approaches using structural magnetic resonance imaging (MRI) can be informative for disease classification; however, their applicability to earlier clinical stages of psychosis and other disease spectra is unknown. We evaluated whether a model differentiating patients with chronic schizophrenia (ChSZ) from healthy controls (HCs) could be applied to earlier clinical stages such as first-episode psychosis (FEP), ultra-high risk for psychosis (UHR), and autism spectrum disorders (ASDs).** ***Study Design:*** **Total 359 T1-weighted MRI scans, including 154 individuals with schizophrenia spectrum (UHR, *n* = 37; FEP, *n* = 24; and ChSZ, *n* = 93), 64 with ASD, and 141 HCs, were obtained using three acquisition protocols. Of these, data regarding ChSZ (*n* = 75) and HC (*n* = 101) from two protocols were used to build a classifier (training dataset). The remainder was used to evaluate the classifier (test, independent confirmatory, and independent group datasets). Scanner and protocol effects were diminished using ComBat.** ***Study Results:*** **The accuracy of the classifier for the test and independent confirmatory datasets were 75% and 76%, respectively. The bilateral pallidum and inferior frontal gyrus pars triangularis strongly contributed to classifying ChSZ. Schizophrenia spectrum individuals were more likely to be classified as ChSZ compared to ASD (classification rate to ChSZ: UHR, 41%; FEP, 54%; ChSZ, 70%; ASD, 19%; HC, 21%).** ***Conclusion:*** **We built a classifier from multiple protocol structural brain images applicable to independent samples from different clinical stages and spectra. The predictive information of the classifier could be useful for applying neuroimaging techniques to clinical differential diagnosis and predicting disease onset earlier.**

Key words: support vector machine/classification/structural MRI/voxel-based morphometry/multisite study/harmonization

Introduction

Several case–control studies have reported on the characteristics of brain anatomical differences in patients with chronic schizophrenia (ChSZ) compared to healthy controls (HCs), such as a reduction in gray matter volume in the frontal and temporal cortices, hippocampus, thalamus, and nucleus accumbens.[1–4] Mega studies have also identified brain characteristics among patients with

schizophrenia including an increase in the volume of the putamen and pallidum from the Enhancing Neuro Imaging Genetics through Meta-Analysis (ENIGMA, n = 4568)[5] and the Cognitive Genetics Collaborative Research Organization in Japan (COCORO, n = 2,564).[4,6] Gray matter loss in the frontotemporal regions is observed in different stages of schizophrenia, such as first-episode psychosis (FEP) and the ultra-high risk for psychosis (UHR).[7–19] Previous studies reported a progressive decrease in gray matter volume in the superior temporal lobe during the transition period among UHR individuals[9,15,16,18] and several years after disease onset.[9,18,19] Although approximately two-thirds of UHR cases did not develop a psychotic disorder (UHR-NP), some neuroanatomical alterations in the frontotemporal regions were seen in UHR overall and UHR-NP.[7,20] On the other hand, the effect size of the volume reduction in the triangular part of the inferior frontal gyrus (IFG) showed similarity among individuals with UHR, FEP, and ChSZ.[21–23] These results suggest that schizophrenia spectrum disorders have specific brain characteristics before onset, which are also present in individuals with subthreshold psychotic symptoms, and these characteristics become more disease-specific with progression from the first psychotic episode to the chronic stage.

Previous studies reported that neuroanatomical alterations may be partially shared among individuals with schizophrenia and autism spectrum disorders (ASDs) in the frontal lobes, anterior cingulate cortex, insula, basal ganglia, and cerebellum.[24–27] The overlap of structural alterations in the insular cortex was observed in individuals with UHR and FEP.[24,28] Although no large sample studies have directly compared the structural brain characteristics of ASDs and schizophrenia, the findings of ENIGMA suggested that ASD individuals showed less volume loss compared to patients with schizophrenia[27] and greater cortical thickness in the superior frontal gyrus and frontal pole compared to individuals with typical development.[29] We also reported that the volume reduction in the pars triangularis of the IFG is specific to the schizophrenia spectrum, while that in the pars opercularis of the IFG is specific to ASD.[21,23,30] Therefore, investigating the common and spectrum-specific neuroanatomical alterations in schizophrenia and ASD may provide new biological insights beyond case–control studies and render them applicable to possible biological markers in clinical settings.

Recently, machine learning approaches have been applied to structural brain imaging to determine the classification pattern of patients with psychiatric disorders.[3,12,31–34] We previously built a three-class machine learning classifier differentiating ChSZ, ASD, and HC, and showed that UHR and FEP individuals were classified into the ChSZ and HC groups but not into the ASD group.[35] Therefore, the machine learning approach can be informative for disease classification applicable to different clinical stages of psychosis. However, several limitations should be addressed in its interpretation and clinical application. First, only a few studies have evaluated a model that achieved good performance in terms of overall accuracy with independent data.[2,31,36,37] Lack of generalizability to unseen data is prone to information leaks between the training and test datasets.[38] To minimize data leakage when building models, some strategies can be used, such as repeating the preprocess of brain images for each cross-validation fold or holding out a validation dataset. The former strategy is meant to preprocess the images in every fold while building the classifiers, which requires a high calculation cost for neuroimaging data. Holding out a validation dataset is applied in the classifier evaluation step using the data that were not used in the model building process by holding out parts of the data in advance. Since the data collected from the same site and procedure still includes potential information leakage, independent confirmatory data outside of the site and procedure for building a classifier or newly measured data after building the classifier will be needed. Second, previous studies have used large samples from multiple sites and datasets; however, the differences in measurement protocols and magnetic resonance imaging (MRI) equipment have been neglected.[31,39,40] A recent multisite resting-state functional connectivity study showed that disease-derived information from functional images was smaller compared to machine- and protocol-derived information.[41] The machine learning approach requires a large sample size; however, one machine and protocol from a single site is limited in sampling from various psychiatric disorders. Diminishing the machine- and protocol-derived differences in MRI data should be considered before machine learning classifiers are applied to a clinical setting.

Here, we intend to develop a support vector machine (SVM) classifier to differentiate between individuals with ChSZ and HCs, and test whether the classifier applies to individuals with earlier schizophrenia spectrum, such as UHR and FEP, and other disorders such as ASD. Although SVM and logistic regression similarly perform using brain MRI in clinical psychiatry,[35] SVM was the most popular algorithm.[31,32,36,37] Furthermore, SVM can generate coordinates for each data point as predictive performance for unknown data. To overcome the limitations associated with previous multisite/multiprotocol datasets, we proposed the use of ComBat,[42] a batch-effect correction tool for harmonizing voxel-wise data collected with multiple protocols. ComBat is a suitable method for application to multidataset gray matter volumes or that of cortical thickness. Furthermore, we investigated the neuroanatomical alterations between the ChSZ and HC groups. To prevent information leakage from the data used in building a classifier, we applied a two-step approach using an independent confirmatory dataset and other clinical stages and spectrum data to evaluate the performance of the classifier.

First, we hypothesized that the performance of the classifiers would be retained for independent confirmatory

datasets. Second, the classifier would discriminate the various clinical stages of psychosis (UHR and FEP) as schizophrenia, and ASD as HCs. Third, the characteristic of schizophrenia would correlate positively with the clinical stages, as greater anatomical alterations would be observed for patients with longer duration of illness. An SVM classifier was employed with the following purposes: (1) to ascertain how distinguishable ChSZ patients and HCs are from each other using T1-weighted MRI data, (2) to describe patterns of morphological features/neuroanatomical alterations contributing to the classification of psychosis, and (3) to evaluate the performance of the classifier in predicting the decision scores of unseen data such as those of UHR, FEP, and ASD individuals.

Methods

Participants

A total of 359 participants, including 154 individuals with schizophrenia spectrum disorder (UHR, n = 37; FEP, n = 24; and ChSZ, n = 93), 64 individuals with ASD, and 141 HCs were enrolled in this study (table 1). The HC group was matched for sex, age, and premorbid intelligence quotient (IQ) to the ChSZ group ($P > .05$).

We applied a two-step approach to evaluate the performance of the models by dividing the data into four datasets: training, test, independent confirmatory, and independent group datasets (figure 1). First, the training and test datasets comprised the data of individuals with

Table 1. Demographic Characteristics of Study Participants

	HC	UHR	FEP	ChSZ	ASD	Statistical value	*P*-value
Participants							
Total, *n*	141	37	24	93	64		
Protocol1, *n*	58	27	20	34	37		
Protocol2, *n*	55	2	0	49	5		
Protocol3, *n*	28	8	4	10	22		
Sex, male/female	83/58	20/17	18/6	55/38	61/3	$\chi 2 = 32.89$	<.001
Age, *mean (SD)*	29.66 (7.88)	20.59 (3.35)	24.46 (5.88)	31.44 (10.03)	29.44 (6.67)	$F = 39.20$ HC>FEP>UHR ASD>FEP>UHR	<.001
Handedness (right), %	85.71	86.49	95.83	77.42	82.81		
IQ							
JART IQ, *mean (SD)*	104.82 (7.42)	105.02 (9.79)	106.27 (10.46)	101.68 (10.33)	NA	$F = 3.61$ HC>ChSZ	.018
FIQ, *mean (SD)*	NA	NA	NA	NA	105.79 (12.26)	NA	NA
VIQ, *mean (SD)*	NA	NA	NA	NA	112.55 (13.56)	NA	NA
PIQ, *mean* (*SD*)	NA	NA	NA	NA	94.98 (16.18)	NA	NA
DOI, *mean* (*SD*)	NA	NA	0.21 (0.17)	7.28 (6.91)	NA	$t = 9.96$	<.001
CPeq, *mean* (*SD*)	NA	154.0 (306.0)	452.4 (420.2)	698.3 (582.1)	NA	$F = 23.14$ ChSZ>FEP>UHR	<.001
PANSS							
Positive symptom, *mean* (*SD*)	NA	12.5 (3.88)	13.48 (4.57)	16.47 (5.28)	NA	$F = 8.77$ ChSZ>UHR	<.001
Negative symptom, *mean* (*SD*)	NA	16.17 (6.37)	18.1 (4.45)	20.38 (6.69)	NA	$F = 6.57$ ChSZ>UHR	.002
General psychopathology, *mean* (*SD*)	NA	31.2 (8.81)	33.76 (7.95)	36.41 (9.78)	NA	$F = 4.42$ ChSZ>UHR	.014
ADI-R							
Social, *mean* (*SD*)	NA	NA	NA	NA	15.63 (5.16)	NA	NA
Com, *mean* (*SD*)	NA	NA	NA	NA	9.45 (4.58)	NA	NA
RRB, *mean* (*SD*)	NA	NA	NA	NA	3.89 (1.88)	NA	NA
Total, *mean* (*SD*)	NA	NA	NA	NA	27.79 (9.54)	NA	NA

Note: HC, healthy control; UHR, ultra-high risk for psychosis; FEP, first-episode psychosis; ChSZ, chronic schizophrenia; ASD, autism spectrum disorder; SD, standard deviation; DOI, duration of illness; CPeq, chlorpromazine equivalent doses; PANSS, Positive and Negative Syndrome Scale; ADI-R, Autism Diagnostic Interview-revised; Social, Social interaction issues; Com, Communication and language skills; RRB, Restricted and repetitive behavior.

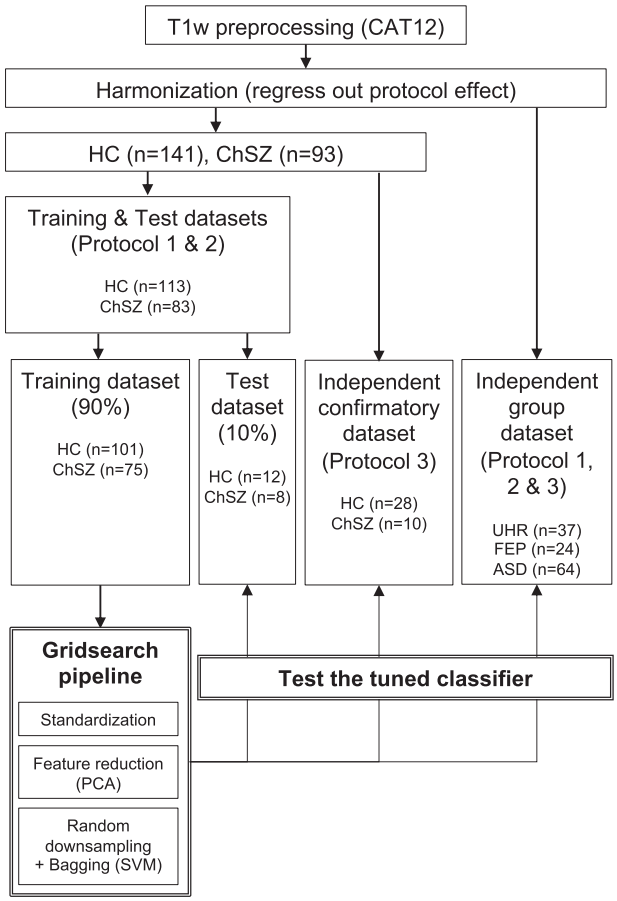

Fig. 1. Pipeline employed for the processing and analysis of T1-weighted images. HC, healthy control; UHR, ultra-high risk for psychosis; FEP, first-episode psychosis; ChSZ, chronic schizophrenia; ASD, autism spectrum disorder.

ChSZ and HC collected using protocols 1 and 2. Ninety percent of the data were randomly sorted as the training dataset, and the remaining 10% were sorted as the test dataset. The independent confirmatory dataset comprised the data of the same groups collected using protocol 3, which was completely excluded from the training partition, to perform an independent first-step evaluation without site information leakage. To evaluate the classifier for earlier clinical stages of psychosis as the second step, we defined the independent group dataset as the data of individuals with UHR, FEP, and ASD collected using any protocol.

The participants were recruited from the outpatient and inpatient units of the University of Tokyo Hospital, University of Tokyo Health Service Center, psychiatry clinics, and internet referrals. Individuals with ChSZ were diagnosed with schizophrenia by designated psychiatrists. Individuals with ASD were diagnosed according to the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision (DSM-IV-TR)[43] with more than 2 months of follow-up examinations by an experienced psychiatrist (H.Y.). The diagnoses were further confirmed by a certified psychiatrist (H.K.) using the Japanese version of the Autism Diagnostic Interview – Revised (ADI-R). The inclusion criteria were ages 15–40 years for FEP and 15–30 years for UHR, nonreceipt of antipsychotic medications for more than 16 cumulative weeks, and continuous psychotic symptoms within the past 6 months. All eligible participants in the UHR and FEP groups were assessed using the Structured Interview for Prodromal Symptoms (SIPS)[44,45] and evaluated using the UHR or psychosis criteria (supplementary materials). The SIPS criteria for psychosis were the same as those of psychotic disorders in the DSM-IV-TR. HCs were not diagnosed with ASD, schizophrenia, or any other psychiatric disorder, and were screened for neuropsychiatric disorders using the Structured Clinical Interview for DSM-IV, Nonpatient Edition.[46]

The exclusion criteria were as follows: (1) previous and/or present severe brain injury and/or neurological illness; (2) a previous history of electroconvulsive therapy; (3) a premorbid IQ of ≤70 as assessed using the 25-item version of the Japanese Adult Reading Test (JART25)[47,48] for the schizophrenia spectrum groups and full scale of the Wechsler Adult Intelligence Scale Revised Japanese version (WAIS-R) for the ASD group; (4) previous and/or present alcohol addiction; and (5) previous and/or present continuous substance use. For the schizophrenia spectrum groups, we also excluded participants with clearly comorbid ASD according to the DSM-IV criteria.[49] The detailed inclusion and exclusion criteria for the UHR and FEP groups are described in the protocol paper.[49]

Symptom severity for schizophrenia spectrum groups was assessed using the Positive and Negative Syndrome Scale (PANSS)[50] and designated using the positive, negative, and general psychopathology subscales. For the ASD group, the ADI-R subtypes (social, communication, and restricted and repetitive behavior [RRB]) were assessed. The chlorpromazine equivalent dose was calculated for medications received at the time of scanning.[51]

The study protocol was approved by the ethics committee of the Faculty of Medicine, University of Tokyo (approval nos. 397, 629, 630, and 2226). All participants provided written informed consent to participate in the measurements after receiving a complete explanation of the experiment.

MRI Data Acquisition and Preprocessing

T1-weighted images were obtained from the three datasets using different scanners and protocols. All structural MRI images were acquired using 3 Tesla General Electric scanners. The scanning for protocol 1 was performed using an 8-channel head coil on SIGNA HDx (3D-FSPGR sequence, 176 axial slices, slice thickness: 1.0 mm; supplementary table S1), for protocol 2 using a 24-channel head coil on DISCOVERY MR750w (Sag_IR-FSPGR sequence, 196 sagittal slices, slice thickness:

1.2 mm), and for protocol 3 using a 32-channel head coil on DISCOVERY MR750w (3D-FSPGR sequence, 172 axial slices, slice thickness: 1.0 mm). All T1-weighted images were first corrected for intensity nonuniformity with N4BiasFieldCorrection,[52] distributed using ANTs 2.2.0 (RRID: SCR_004757).[53] The intensity corrected images were then segmented to produce images of different tissue types (gray matter, white matter, and cerebrospinal fluid) using CAT12 software (http://www.neuro.uni-jena.de/cat/). We used the Diffeomorphic Anatomical Registration Through Exponentiated Lie Algebra (DARTEL)[54] option in the CAT12 toolbox to normalize the segmented scans into a standard Montreal Neurological Institute (MNI) space. Default parameters were used for this preprocessing. Smoothing was applied later in the harmonization process. To assess gray matter segment homogeneity and identify possible outliers, the "Check Data Quality" module of CAT12 was used. No participants were excluded from this step. To retain maximum image information, voxels from gray matter images were smoothed using a Gaussian smoothing kernel of 2-mm full width at half-maximum as features in further classification. All features from the smoothed, modulated, and normalized gray matter images were transformed to a two-dimension matrix (participants × features) using Niftimasker, a component of Nilearn (https://github.com/nilearn/nilearn).

ComBat Harmonization

ComBat[42] is a harmonization method used to remove scanner and protocol effects based on the adjusted general linear model harmonization method. ComBat uses Bayesian criteria to improve the estimation for small sample size data (supplementary materials). Further analyses were conducted using Python version 3.7.4. We applied the transformed two-dimensional data with participants' age and sex as covariables, along with protocol effects (supplementary figure S1).

Support vector Machine

The main idea behind SVMs is to separate two groups using a contrasted hyperplane (supplementary materials).[55] In building a classifier, we applied standardization and dimension reduction using principal component analysis (PCA) to the dataset, which included 554 992 features. The usage of value standardization, optimization of the number of PCA components, and hyperparameters of the classifier (penalty parameter C and kernel parameter gamma) were tuned using GridSearchCV implemented in the "scikit-learn" module (version 0.21.3) in Python (https://scikit-learn.org/stable/whats_new/v0.21.html). We plotted the weights of the classifier to determine the importance of the features for generalization (supplementary materials).[56,57] The classifier was optimized using a 10-fold cross-validated grid search over a defined parameter grid. Data from the HC group were randomly downsampled to the same ratio as the ChSZ group in each fold. To reduce downsampling bias, downsampling and grid search were repeated 1000 times and stratified 10-fold for the training data. Then, we applied 10-fold cross-validation and 1000 permutations to evaluate the significance of the cross-validation scores of the model with the best hyperparameters for the training dataset. The best cross-validation accuracy score was averaged across 1000 repeats. Permutation tests were conducted by shuffling the labels in the training data, and the permutation-based *P*-value was calculated.[58] The final model with the best hyperparameters was trained using the entire training dataset. One hundred and fifty PCA components were used as features. Finally, the trained classifier was applied to the test set and the independent confirmatory dataset with the best parameters tuned by grid search. The predictive performance of the classifier was evaluated using an independent group dataset (UHR, FEP, and ASD data collected using any protocol).

Statistical Analysis

Evaluation metrics First, the classifier was evaluated using the test, independent confirmatory, and independent group datasets separately by the given scores of the tuned classifier using the training dataset. We calculated the confusion matrix, macro, and weighted average accuracies to evaluate the classifier because the data used were imbalanced (supplementary materials). We also reported the area under the curve (AUC) of the receiver operator characteristic.

Predictive performance of the classifier The predictive performance of the classifier and the chi-squared test were applied to the classified labels of the test, independent confirmatory, and independent group datasets. Since we compared 10 pairs of groups, Bonferroni's correction was applied to post-hoc comparisons ($P < .05/10 = .005$). Decision scores generated by the SVM were tested using an analysis of variance separately for all samples corresponding to the hyperplane. Bonferroni's correction was applied to post-hoc comparisons ($P < .05/10 = .005$).

Correlations between decision scores and clinical severity To determine the relationship between the decision score and symptom severity, Pearson's correlation analyses were performed using the PANSS subscores for the schizophrenia spectrum groups and the ADI-R subscale scores for the ASD group. Bonferroni's correction was applied to the subscores ($P < .05/3 = .016$). To determine the potential effect of medication on the classification, we also tested the correlation between the decision score and medication dose for the schizophrenia spectrum groups using Spearman's rank correlation (uncorrected $P < .05$). For the UHR group, we also tested

the difference in decision scores between those with and without medication using a *t*-test.

Results

Model Evaluation

The best cross-validation accuracy within the training dataset was 74% (±0.68). The permutation test showed that it was significantly higher than that attributable to chance (50%, $P < .001$). The accuracy with the best estimator for the test and independent confirmatory datasets were 75% (AUC = 0.88) (table 2, supplementary figure S2A) and 76% (AUC = 0.82) (supplementary figure S2B), respectively.

The voxel space feature weights of the SVM showed that the clusters including the IFG pars triangularis, superior frontal gyrus, cuneus, superior occipital gyrus, putamen, and pallidum contributed to identifying ChSZ (figure 2A, table 3). Clusters including the inferior parietal gyrus, inferior occipital gyrus, superior parietal gyrus, and middle frontal gyrus contributed to the identification of HCs.

Predictive Performance of the Classifier for the Test, Independent Confirmatory, and Independent Group Datasets

A chi-squared test showed a significant difference within the classified labels for the test, independent confirmatory, and independent group datasets, respectively ($X^2(1, n = 20) = 5.69$, $P < .05$; $X^2(1, n = 38) = 7.72$, $P < .01$; and $X^2(2, n = 118) = 11.25$, $P < .01$). Further residual analysis showed that the HC group was significantly more likely to be classified as HCs than the ChSZ group in the independent confirmatory dataset (79% vs. 21%, corrected $P < .01$). For the independent group dataset, the ASD and UHR groups were significantly more likely to be classified as HCs (classification rate to HC: 81% and 59%, respectively, corrected $P < .01$), while the FEP group as ChSZ (46%, corrected $P < .01$; figure 2B). Compared to the ASD group, UHR was more likely to be classified as ChSZ (corrected $P < .01$). A chi-squared test showed no difference between HCs from the independent confirmatory dataset and ASD participants from the independent group dataset.

Using the decision score generated by the SVM for all groups, we found a significant main effect of the decision score for the five groups ($F = 161.99$, $P < .001$). Multiple comparisons showed that UHR was close to FEP but different from HC, ASD, and ChSZ (HC < ASD < UHR and FEP < ChSZ; corrected $P < .001$) (figure 2C).

Table 2. Classification Report of the Test and Independent Datasets

	Precision	Recall	F1-score	*N* for prediction
Test dataset				
HC	0.89	0.67	0.76	12
ChSZ	0.64	0.88	0.74	8
macro avg	0.76	0.77	0.75	20
weighted avg	0.79	0.75	0.75	20
Independent confirmatory dataset				
HC	0.88	0.79	0.83	28
ChSZ	0.54	0.70	0.61	10
macro avg	0.71	0.74	0.72	38
weighted avg	0.79	0.76	0.77	38

Since the samples from protocol 1 and 2 were used for Test dataset (HC = 12, ChSZ = 8) and those from protocol 3 for Independent confirmatory dataset (HC = 28, ChSZ = 10), the metrics except for weighted average accuracy (weighted avg) may be affected by sampling imbalance (Supplementary Materials).

Correlations Between Decision Scores and Clinical Severity

No significant correlations were found between PANSS subscores and decision scores. In the schizophrenia spectrum groups, significant correlations between chlorpromazine equivalent dose and decision scores were observed in the UHR group ($rho = 0.44$, $P < .05$). However, no significant differences in decision scores between medicated and nonmedicated UHR participants were found ($t = -1.78$, not significant). No significant correlation was found between the ADI-R subscale scores and decision scores in the ASD group. No significant correlations were found between JART IQ and decision scores in each group.

Discussion

To the best of our knowledge, the current study is the first to apply machine learning to the classifier for HC and ChSZ groups with multi-protocol structural MRI and multidisease spectrum and clinical stages. To evaluate the classifier, we confirmed a two-step approach using an independent confirmatory dataset obtained via a different protocol from the ones used in building the classifier and the earlier clinical stage and ASD dataset. We successfully achieved a 74% accuracy in the 2-class classification within training dataset. The decision scores from the classifier indicated that the characteristics of ChSZ became pronounced in the UHR and FEP groups, which was different from those in the HC and ASD groups.

In this study, we discriminated between the HC and ChSZ groups with 74% and 75% accuracy in the training and test sets, respectively. The performance of the classifier on the independent confirmatory dataset achieved 76% accuracy. To avoid downsampling bias, we performed random execution and 1000 repeats, which may cause a slightly lower training accuracy than the test accuracy (≤1%). Our results are compatible with the accuracy of 72%–77% in a previous study.[37] As Rozycki et al.[37] treated MRI scanners as a variable in the multivariate analysis, we used ComBat to harmonize the effects of MRI protocols

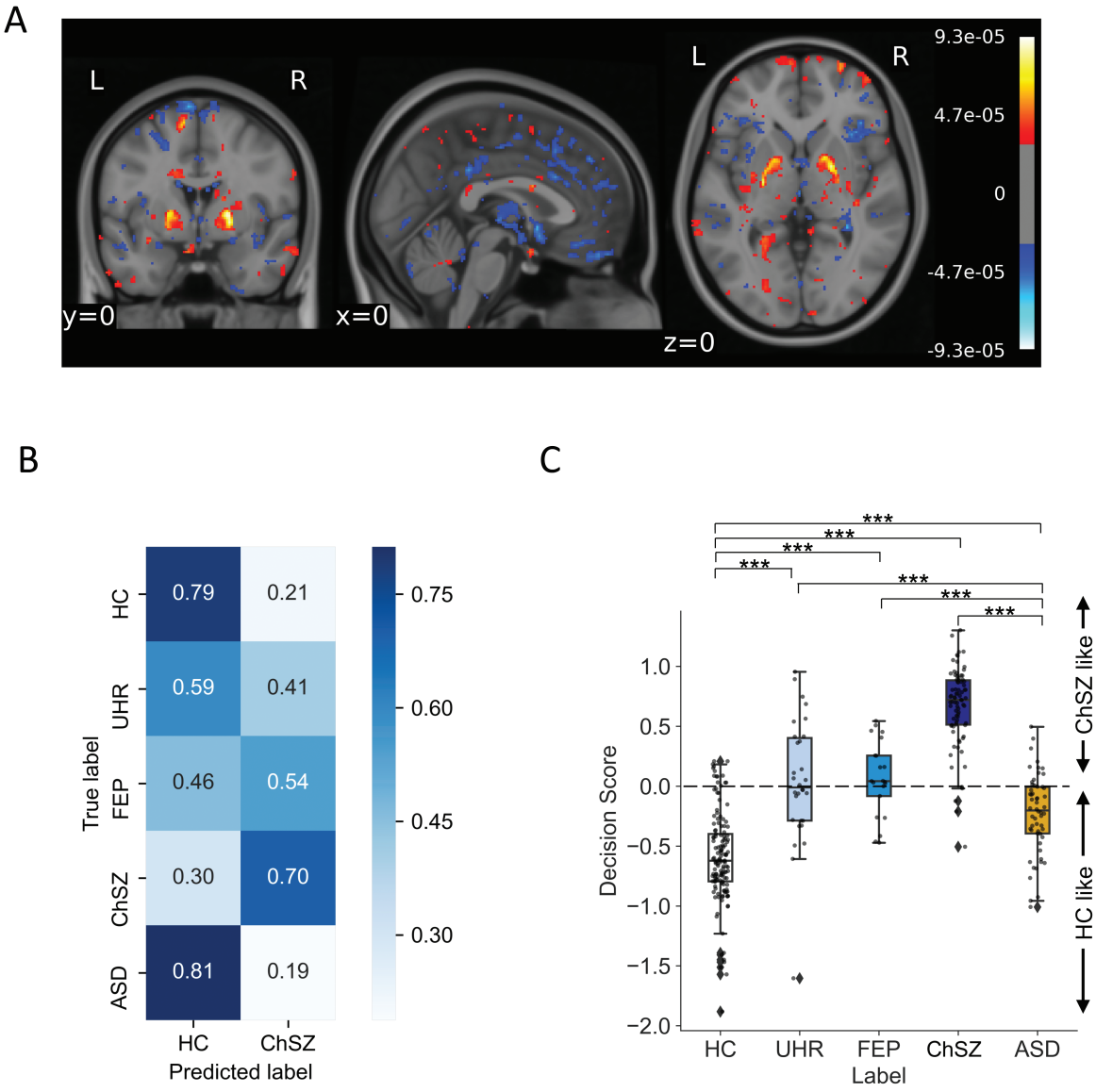

Fig. 2. Voxel feature contributions and predictive performance comparisons of support vector machine (SVM). (A) Weighted features of SVM classification in the voxel space. Note that positive scores indicate the regions of feature contribution for identifying patients with chronic schizophrenia, while negative score for healthy controls. (B) Predictive performance of HC and ChSZ groups was evaluated using the independent confirmatory dataset, and UHR, FEP, and ASD groups using the independent group dataset. (C) Box and scatter plot of decision scores of support vector machines. *P*-values of post hoc comparisons were corrected using the Bonferroni method (***P < .001, **P < .01, *P < .05).

or scanners. ComBat is superior at preserving within-site biological variability and improving the consistency and replicability of the voxels associated with age.[59–61] By applying the two-step approach using multi-protocol samples, we obtained less biased results than evaluating the classifier using data from the same dataset. We obtained a model with a solid predictive performance for new data, unlike previous studies that did not test the performance of the trained classifier.[2,37] As expected, a majority of ASD patients were classified as HCs. The decision scores also indicated that ASD and UHR patients share some characteristics that differ from those of the HC or ChSZ group. Moreover, no significant associations were found between IQ or chlorpromazine equivalent doses and decision scores, indicating that the brain characteristics of schizophrenia found in the present study may serve as biomarkers for improving methods for differential diagnosis.

In line with previous studies of volumetric alterations in the schizophrenia spectrum,[1–6,13] we found that a pattern of morphological features, including the IFG pars triangularis, putamen, and pallidum contributed to the identification of ChSZ patients. Using a manual tracing method within the IFG, we previously found that the volume reduction in the IFG pars triangularis is a disease-specific feature in the schizophrenia spectrum, such as UHR,[21] FEP,[21] and ChSZ,[23] and the extent of

Table 3. Top 10 Weighted Features of Regions Contributed to SVM Classification Included in AAL Atlas.

	Peak			Contribution values			Brain regions	
Cluster ID	x	y	z	Cluster mean weight ($\times10^{-5}$)	Peak weight ($\times10^{-5}$)	Volume (mm^3)	Cluster	Peak
Top 10 by cluster mean weight for healthy controls								
361	−51	−55.5	54	−3.09	−4.42	104.625	90.32% Parietal_Inf_L; 9.68% no_label	Parietal_Inf_L
48	−36	−73.5	−7.5	−3.08	−5.53	1090.12	75.85% Occipital_Inf_L; 13.00% Occipital_Mid_L; 5.88% Fusiform_L; 5.26% no_label	Occipital_Inf_L
504	25.5	−25.5	37.5	−3.04	−3.84	74.25	100.00% no_label	no_label
63	40.5	13.5	61.5	−3.03	−6.24	772.875	96.51% Frontal_Mid_2_R	Frontal_Mid_2_R
392	12	−60	60	−3.03	−5.27	101.25	80.00% Precuneus_R; 20.00% Parietal_Sup_R	Precuneus_R
148	−4.5	−91.5	36	−3.02	−4.57	273.375	69.14% Cuneus_L; 19.75% no_label; 11.11% Occipital_Sup_L	no_label
105	28.5	−37.5	72	−3.01	−4.47	381.375	100.00% Postcentral_R	Postcentral_R
333	19.5	−61.5	−64.5	−2.98	−5.06	114.75	100.00% no_label	no_label
586	−16.5	−43.5	−52.5	−2.98	−4.47	57.375	88.24% Cerebelum_9_L; 11.76% no_label	Cerebelum_9_L
74	−12	−63	72	−2.96	−5.41	631.125	90.37% Precuneus_L; 5.35% Parietal_Sup_L	Precuneus_L
Top 10 by peak weight for healthy controls								
1	7.5	46.5	−24	−2.89	−6.98	118770	23.13% no_label	Rectus_R
26	−19.5	−51	76.5	−2.87	−6.45	1859.62	54.99% Parietal_Sup_L; 22.87% Precuneus_L; 14.34% Postcentral_L; 5.44% Paracentral_Lobule_L	Parietal_Sup_L
63	40.5	13.5	61.5	−3.03	−6.24	772.875	96.51% Frontal_Mid_2_R	Frontal_Mid_2_R
12	28.5	−7.5	−42	−2.81	−5.83	2504.25	47.04% no_label; 18.06% Fusiform_R; 15.09% Temporal_Pole_Mid_R; 12.67% Temporal_Inf_R; 6.74% ParaHippocampal_R	Fusiform_R
5	−27	−10.5	46.5	−2.87	−5.69	5055.75	31.38% Precentral_L; 27.10% Frontal_Mid_2_L; 26.44% Frontal_Sup_2_L; 14.35% no_label	no_label
21	−45	7.5	28.5	-2.73	−5.66	2052	43.75% Precentral_L; 39.47% Frontal_Inf_Oper_L; 11.84% Frontal_Inf_Tri_L	Frontal_Inf_Oper_L
69	57	−31.5	−21	−2.94	−5.57	678.375	90.55% Temporal_Inf_R; 7.46% no_label	Temporal_Inf_R
161	−15	−73.5	63	−2.91	−5.54	239.625	73.24% Precuneus_L; 18.31% no_label; 8.45% Parietal_Sup_L	Precuneus_L
48	−36	−73.5	−7.5	−3.08	−5.53	1090.12	75.85% Occipital_Inf_L; 13.00% Occipital_Mid_L; 5.88% Fusiform_L; 5.26% no_label	Occipital_Inf_L
11	−37.5	−49.5	−52.5	−2.76	−5.49	2575.12	65.92% Cerebelum_8_L; 21.10% no_label; 10.62% Cerebelum_7b_L	Cerebelum_8_L
Top 10 by cluster mean weight for chronic schizophrenia								
432	−49.5	46.5	4.5	3.73	6.02	91.125	96.30% Frontal_Inf_Tri_L	Frontal_Inf_Tri_L
253	−12	51	46.5	3.55	5.73	155.25	58.70% Frontal_Sup_2_L; 32.61% Frontal_Sup_Medial_L; 8.70% no_label	Frontal_Sup_2_L
72	−9	−84	27	3.26	6.75	648	83.85% Cuneus_L; 16.15% Occipital_Sup_L	Cuneus_L
516	−10.5	−33	−28.5	3.25	5.53	70.875	100.00% no_label	no_label
138	18	−52.5	72	3.18	5.63	293.625	100.00% Parietal_Sup_R	Parietal_Sup_R
224	54	40.5	7.5	3.13	5.01	178.875	100.00% Frontal_Inf_Tri_R	Frontal_Inf_Tri_R
297	58.5	24	25.5	3.09	4.97	131.625	87.18% Frontal_Inf_Tri_R; 12.82% Frontal_Inf_Oper_R	Frontal_Inf_Tri_R
79	−9	0	64.5	3.08	6.39	563.625	79.64% Supp_Motor_Area_L; 16.17% no_label	Supp_Motor_Area_L
396	−49.5	−45	−45	3.06	4.02	97.875	100.00% Cerebelum_Crus2_L	Cerebelum_Crus2_L
1308	12	−48	22.5	3.04	3.37	16.875	60.00% Cingulate_Post_R; 40.00% Precuneus_R	Cingulate_Post_R
Top 10 by peak weight for chronic schizophrenia								
2	22.5	−4.5	3	2.97	9.32	17465.6	48.10% no_label; 14.40% Putamen_R; 9.80% Pallidum_R; 5.84% Caudate_R; 5.10% Temporal_Sup_R	Pallidum_R
3	−19.5	−1.5	4.5	3.01	8.66	11292.8	35.51% no_label; 19.75% Putamen_L; 16.44% Pallidum_L; 8.70% Parietal_Inf_L	Pallidum_L
72	−9	−84	27	3.26	6.75	648	83.85% Cuneus_L; 16.15% Occipital_Sup_L	Cuneus_L
4	6	70.5	4.5	2.98	6.72	9966.38	17.78% Frontal_Sup_Medial_L; 15.92% Frontal_Sup_2_R; 14.83% no_label; 12.87% Frontal_Sup_Medial_R; 11.38% Frontal_Sup_2_L; 9.75% Frontal_Mid_2_R	Frontal_Sup_Medial_R

Table 3. Continued

Cluster ID	Peak x	y	z	Contribution values: Cluster mean weight ($\times 10^{-5}$)	Peak weight ($\times 10^{-5}$)	Volume (mm^3)	Brain regions: Cluster	Peak
79	−9	0	64.5	3.08	6.39	563.625	79.64% Supp_Motor_Area_L; 16.17% no_label	Supp_Motor_Area_L
80	15	−76.5	42	2.87	6.18	560.25	65.06% Cuneus_R; 19.88% Precuneus_R; 7.83% Parietal_Sup_R	Cuneus_R
432	−49.5	46.5	4.5	3.73	6.02	91.125	96.30% Frontal_Inf_Tri_L	Frontal_Inf_Tri_L
64	42	−6	21	2.83	5.74	752.625	47.09% no_label; 42.60% Rolandic_Oper_R; 7.62% Insula_R	Rolandic_Oper_R
45	64.5	1.5	27	2.77	5.74	1194.75	64.41% Precentral_R; 14.97% Postcentral_R; 11.02% Frontal_Inf_Oper_R; 9.32% Rolandic_Oper_R	Postcentral_R
253	−12	51	46.5	3.55	5.73	155.25	58.70% Frontal_Sup_2_L; 32.61% Frontal_Sup_Medial_L; 8.70% no_label	Frontal_Sup_2_L

The 10 largest clusters that contributed to the classification by cluster mean and peak of the brain regions are shown. All features are listed in supplementary tables S2-3. By transforming the weights of each PCA component back into the original feature (voxel) output, a weighted image was generated in a standard MNI space, which including the coordinates of the weights. AAL atlas was used to generate a coordinates table and region labels based on the information given by the weighted image.

structural alteration was associated with brain activity during a verbal fluency test in FEP.[22] These findings suggest that a decrease in gray matter volume in the region is a trait characteristic of the schizophrenia spectrum. On the other hand, the IFG pars opercularis is related to a disease-specific feature in ASD.[30] In this study, the SVM classifier used the structural features of the IFG pars triangularis as one of the contributory regions and is a reasonable selection according to previous studies.

Previous large-scale studies from ENIGMA[5] and COCORO[6] found increased volumes of the putamen and pallidum in patients with schizophrenia. Recently, our multi-site study showed that individuals with UHR exhibited increased volumes of the left caudate and pallidum[62]; however, the ENIGMA clinical high risk for psychosis study failed to replicate these findings.[20] This suggests that an altered striatum and pallidum volume is another disease-specific feature of the schizophrenia spectrum. These volumes are inconsistent in terms of FEP.[63–65] Nevertheless, the findings of the present study suggest a major contribution of the putamen and pallidum for classifying ChSZ, and these regions may contribute to some disease-specific pathologies in schizophrenia.

The decision score given by the classifier increased according to the clinical stages of psychosis, suggesting that the gradual transitions of neuroanatomical characteristics were distinguishable among the HC, UHR, FEP, and ChSZ groups. This further implies that prominent symptoms (anatomical alterations) appear after FEP and progressively change after the initial period of psychosis onset (HC < ASD < UHR and FEP < ChSZ) (figure 2C). Although a partial overlap of anatomical alterations has been reported between the ASD and ChSZ groups,[66,67] the classifier in this study weighted otherwise. As ASD is mostly classified as HC compared to other schizophrenia spectrum groups, the neuroanatomical changes used in the classifier during early clinical stages are specific to schizophrenia and weighed both differences in trait and progressions in brain pathology. Recent studies including whole brain (gray matter, white matter, and ventricular cerebrospinal fluid volumes)[33] or diffusion MRI as features reported a high performance rate[32,34]; with more features, it is possible to achieve a higher accuracy with relatively smaller samples than this study.

Our study had several limitations. First, we did not build a classification system differentiating ASD from other groups because of the smaller sample size. More samples from multiple sites will yield a multiclass classifier that can be used to learn more about disease-specific structural brain characteristics, which may apply to earlier clinical stages of the schizophrenia spectrum. Moreover, although ComBat was applied protocol-wise, there is little risk of information leak since the data used in the current study were acquired from a single site. However, ComBat was unable to distinguish the sampling bias from measurement bias.[41,60,68] While applying combat harmonization to multiple site datasets, the sampling effects could be a potential information leak. Second, due to the limited samples of UHR individuals who later converted to psychosis, we were unable to evaluate whether the classifier could differentiate a later onset of psychosis. Third, although we showed that the classification was slightly influenced by the medication dose, the potential effect of medication on the structural characteristics should be considered. Fourth, although previous machine learning classification studies did not control features for intracranial volume,[31,35,36] we need to see whether the features adjusted for intracranial volume would increase the sensitivity in further studies.

The present study compared a 2-class SVM classifier (HC vs. ChSZ) using the multiprotocol voxel-based morphometry datasets, which showed a good predictive performance for the unknown data of the UHR, FEP, and ASD groups. The classifier indicated that the characteristic was gradually modified in the UHR, FEP, and ChSZ groups. This method could be the next step to apply brain MRI machine learning algorithms to clinical settings. Further elaboration of the method applied herein may contribute to the early discovery and clinical diagnosis of schizophrenia and ASDs in the future.

Supplementary Material

Supplementary material is available at *Schizophrenia Bulletin*.

Acknowledgments

This work was supported by the Agency for Medical Research and Development (AMED) under grant numbers JP21dm0307001, JP21dm0307004, and JP21dm0207069; the Japan Society for the Promotion of Science (JSPS) KAKENHI grant numbers JP19H03579, JP20KK0193, JP21H02851, JP21H05171, and JP21H05174; and the Japan Science and Technology Agency (JST) Moonshot R&D grant number JPMJMS2021. This study was also supported by the University of Tokyo Center for Integrative Science of Human Behavior (CiSHuB), the International Research Center for Neurointelligence (WPI-IRCN) at the University of Tokyo Institutes for Advanced Study (UTIAS), and the Joint Usage/Research Program of the Medical Institute of Development Disabilities Research, Showa University.
The authors have declared that there are no conflicts of interest in relation to the subject of this study.

References

1. Chand GB, Dwyer DB, Erus G, et al. Two distinct neuroanatomical subtypes of schizophrenia revealed using machine learning. *Brain.* 2020;143(3):1027–1038.
2. Nemoto K, Shimokawa T, Fukunaga M, et al. Differentiation of schizophrenia using structural MRI with consideration of scanner differences: a real-world multisite study. *Psychiatry Clin Neurosci.* 2020;74(1):56–63.
3. Schwarz E, Doan NT, Pergola G, et al.; IMAGEMEND Consortium, Karolinska Schizophrenia Project (KaSP) Consortium. Reproducible grey matter patterns index a multivariate, global alteration of brain structure in schizophrenia and bipolar disorder. *Transl Psychiatry.* 2019;9(1):12.
4 Koshiyama D, Miura K, Nemoto K, et al. Neuroimaging studies within Cognitive Genetics Collaborative Research Organization aiming to replicate and extend works of ENIGMA. *Hum Brain Mapp.* 2022;43(1):182–193.
5. van Erp TG, Hibar DP, Rasmussen JM, et al. Subcortical brain volume abnormalities in 2028 individuals with schizophrenia and 2540 healthy controls via the ENIGMA consortium. *Mol Psychiatry*. 2016;21(4):547–553.
6. Okada N, Fukunaga M, Yamashita F, et al. Abnormal asymmetries in subcortical brain volume in schizophrenia. *Mol Psychiatry.* 2016;21(10):1460–1466.
7. Cannon TD, Chung Y, He G, et al.; North American Prodrome Longitudinal Study Consortium. Progressive reduction in cortical thickness as psychosis develops: a multisite longitudinal neuroimaging study of youth at elevated clinical risk. *Biol Psychiatry.* 2015;77(2):147–157.
8. D'Ambrosio E, Jauhar S, Kim S, et al. The relationship between grey matter volume and striatal dopamine function in psychosis: a multimodal 18F-DOPA PET and voxel-based morphometry study. *Mol Psychiatry.* 2021;26(4):1332–1345.
9. Takahashi T, Wood SJ, Yung AR, et al. Progressive gray matter reduction of the superior temporal gyrus during transition to psychosis. *Arch Gen Psychiatry.* 2009;66(4):366–376.
10. Wang YM, Yang ZY, Wang Y, et al. Grey matter volume and structural covariance associated with schizotypy. *Schizophr Res.* 2020;224:88–94.
11. Fusar-Poli P, Borgwardt S, Crescini A, et al. Neuroanatomy of vulnerability to psychosis: a voxel-based meta-analysis. *Neurosci Biobehav Rev.* 2011;35(5):1175–1185.
12. Koutsouleris N, Dwyer DB, Degenhardt F, et al.; PRONIA Consortium. Multimodal machine learning workflows for prediction of psychosis in patients with clinical high-risk syndromes and recent-onset depression. *JAMA Psychiatry.* 2021;78(2):195–209.
13. Koike S, Uematsu A, Sasabayashi D, et al. Recent advances and future directions in brain MR imaging studies in schizophrenia: toward elucidating brain pathology and developing clinical tools. *Magn Reson Med Sci*. Published online 2021. doi:10.2463/mrms.rev.2021-0050
14. Olabi B, Ellison-Wright I, McIntosh AM, Wood SJ, Bullmore E, Lawrie SM. Are there progressive brain changes in schizophrenia? A meta-analysis of structural magnetic resonance imaging studies. *Biol Psychiatry.* 2011;70(1):88–96.
15. Pantelis C, Velakoulis D, McGorry PD, et al. Neuroanatomical abnormalities before and after onset of psychosis: a cross-sectional and longitudinal MRI comparison. *Lancet.* 2003;361(9354):281–288.
16. Pantelis C, Yücel M, Bora E, et al. Neurobiological markers of illness onset in psychosis and schizophrenia: the search for a moving target. *Neuropsychol Rev.* 2009;19(3):385–398.
17. Sami M, Worker A, Colizzi M, et al.; Collaborators. Association of cannabis with glutamatergic levels in patients with early psychosis: evidence for altered volume striatal glutamate relationships in patients with a history of cannabis use in early psychosis. *Transl Psychiatry.* 2020;10(1):111.
18. Takahashi T, Wood SJ, Yung AR, et al. Insular cortex gray matter changes in individuals at ultra-high-risk of developing psychosis. *Schizophr Res.* 2009;111(1–3):94–102.
19. Kasai K, Shenton ME, Salisbury DF, et al. Progressive decrease of left superior temporal gyrus gray matter volume in patients with first-episode schizophrenia. *Am J Psychiatry.* 2003;160(1):156–164.
20. Jalbrzikowski M, Hayes RA, Wood SJ, et al.; ENIGMA Clinical High Risk for Psychosis Working Group. Association of structural magnetic resonance imaging measures with psychosis onset in individuals at clinical high risk for developing psychosis: an ENIGMA working group mega-analysis. *JAMA Psychiatry.* 2021;78(7):753–766.
21. Suga M, Yamasue H, Abe O, et al. Reduced gray matter volume of Brodmann's Area 45 is associated with severe

psychotic symptoms in patients with schizophrenia. *Eur Arch Psychiatry Clin Neurosci.* 2010;260(6):465–473.

22. Iwashiro N, Suga M, Takano Y, et al. Localized gray matter volume reductions in the pars triangularis of the inferior frontal gyrus in individuals at clinical high-risk for psychosis and first episode for schizophrenia. *Schizophr Res.* 2012;137(1–3):124–131.
23. Iwashiro N, Koike S, Satomura Y, et al. Association between impaired brain activity and volume at the sub-region of Broca's area in ultra-high risk and first-episode schizophrenia: a multimodal neuroimaging study. *Schizophr Res.* 2016;172(1–3):9–15.
24. Díaz-Caneja CM, Schnack H, Martínez K, et al. Neuroanatomical deficits shared by youth with autism spectrum disorders and psychotic disorders. *Hum Brain Mapp.* 2019;40(5):1643–1653.
25. Edgar JC. Identifying electrophysiological markers of autism spectrum disorder and schizophrenia against a backdrop of normal brain development. *Psychiatry Clin Neurosci.* 2020;74(1):1–11.
26. Mitelman SA, Bralet M-C, Haznedar MM, et al. Diametrical relationship between gray and white matter volumes in autism spectrum disorder and schizophrenia. *Brain Imaging Behav*. 2016;11(6):1823–1835.
27. Patel Y, Parker N, Shin J, et al. Virtual histology of cortical thickness and shared neurobiology in 6 psychiatric disorders. *JAMA Psychiatry*. 2021;78(1):47–63.
28 Parellada M, Pina-Camacho L, Moreno C, et al. Insular pathology in young people with high-functioning autism and first-episode psychosis. *Psychol Med.* 2017;47(14):2472–2482.
29. van Rooij D, Anagnostou E, Arango C, et al. Cortical and subcortical brain morphometry differences between patients with autism spectrum disorder and healthy individuals across the lifespan: results from the ENIGMA ASD working group. *Am J Psychiatry.* 2018;175(4):359–369.
30. Yamasaki S, Yamasue H, Abe O, et al. Reduced gray matter volume of pars opercularis is associated with impaired social communication in high-functioning autism spectrum disorders. *Biol Psychiatry.* 2010;68(12):1141–1147.
31. Nunes A, Schnack HG, Ching CRK, et al.; ENIGMA Bipolar Disorders Working Group. Using structural MRI to identify bipolar disorders—13 site machine learning study in 3020 individuals from the ENIGMA Bipolar Disorders Working Group. *Mol Psychiatry.* 2020;25(9):2130–2143.
32. Lee J, Chon MW, Kim H, et al. Diagnostic value of structural and diffusion imaging measures in schizophrenia. *Neuroimage Clin.* 2018;18:467–474.
33. Davatzikos C, Shen D, Gur RC, et al. Whole-brain morphometric study of schizophrenia revealing a spatially complex set of focal abnormalities. *Arch Gen Psychiatry.* 2005;62(11):1218–1227.
34. Elad D, Cetin-Karayumak S, Zhang F, et al. Improving the predictive potential of diffusion MRI in schizophrenia using normative models—towards subject-level classification. *Hum Brain Mapp.* 2021;42(14):4658–4670.
35. Yassin W, Nakatani H, Zhu Y, et al. Machine-learning classification using neuroimaging data in schizophrenia, autism, ultra-high risk and first-episode psychosis. *Transl Psychiatry.* 2020;10(1):278.
36. Nieuwenhuis M, van Haren NE, Hulshoff Pol HE, Cahn W, Kahn RS, Schnack HG. Classification of schizophrenia patients and healthy controls from structural MRI scans in two large independent samples. *Neuroimage.* 2012;61(3):606–612.
37. 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.* 2018;44(5):1035–1044.
38. Vieira S, Pinaya WH, Mechelli A. Using deep learning to investigate the neuroimaging correlates of psychiatric and neurological disorders: methods and applications. *Neurosci Biobehav Rev.* 2017;74(Pt A):58–75.
39. Isobe M, Redden SA, Keuthen NJ, et al. Striatal abnormalities in trichotillomania: a multi-site MRI analysis. *Neuroimage Clin.* 2018;17:893–898.
40. Nielsen JA, Zielinski BA, Fletcher PT, et al. Multisite functional connectivity MRI classification of autism: ABIDE results. *Front Hum Neurosci.* 2013;7:599.
41. Yamashita A, Yahata N, Itahashi T, et al. Harmonization of resting-state functional MRI data across multiple imaging sites via the separation of site differences into sampling bias and measurement bias. *PLoS Biol.* 2019;17(4):e3000042.
42. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics.* 2007;8(1):118–127.
43. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 4th ed. Washington, DC: American Psychiatric Press; 1994.
44. Miller TJ, McGlashan TH, Woods SW, et al. Symptom assessment in schizophrenic prodromal states. *Psychiatr Q.* 1999;70(4):273–287.
45. Kobayashi H, Nozaki S, Mizuno M. Reliability of the structured interview for prodromal syndromes Japanese version (SIPS-J). *Jpn Bull Soc Psychiatry.* 2007;15:168–174.
46. First MB, Spitzer RL, Gibbon M, et al. *Structured Clinical Interview for DSM-IV Axis I Disorders, Non-patient Edition*. New York, NY: Biometrics Research Department, New York State Psychiatric Institute; 1997 (Japanese translation: Kitamura T, Okano T. Tokyo, Japan: Nihon Hyoronsha Publishers; 2003).
47. Hirata-Mogi S, Koike S, Toriyama R, Matsuoka K, Kim Y, Kasai K. Reliability of a paper-and-pencil version of the Japanese Adult Reading Test short version. *Psychiatry Clin Neurosci.* 2016;70(8):362.
48. Matsuoka K, Kim Y. *Japanese Adult Reading Test (JART)*. Shinkou-Igaku Publ Tokyo. 2006.
49. Koike S, Takano Y, Iwashiro N, et al. A multimodal approach to investigate biomarkers for psychosis in a clinical setting: the integrative neuroimaging studies in schizophrenia targeting for early intervention and prevention (IN-STEP) project. *Schizophr Res.* 2013;143(1):116–124.
50. Kay SR, Fiszbein A, Opler LA. The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophr Bull.* 1987;13(2):261–276.
51. Inada T, Inagaki A. Psychotropic dose equivalence in Japan. *Psychiatry Clin Neurosci.* 2015;69(8):440–447.
52. Tustison NJ, Avants BB, Cook PA, et al. N4ITK: improved N3 bias correction. *IEEE Trans Med Imaging.* 2010;29(6):1310–1320.
53. Avants BB, Epstein CL, Grossman M, Gee JC. Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. *Med Image Anal.* 2008;12(1):26–41.
54. Ashburner J. A fast diffeomorphic image registration algorithm. *Neuroimage.* 2007;38(1):95–113.
55. Cortes C, Vapnik V. Support-vector networks. *Mach Learn*. 1995;20(3):273–297.

56. Guyon I, Elisseeff A. An introduction to variable and feature selection. *J Mach Learn Res.* 2003;3:1157–1182.
57. Guyon I, Weston J, Barnhill S, et al. Gene selection for cancer classification using support vector machines. *Mach Learn*. 2002;46(1–3):389–422.
58. Ojala M, Garriga GC. Permutation tests for studying classifier performance. *J Mach Learn Res*. 2010;11(6).
59. Fortin JP, Parker D, Tunç B, et al. Harmonization of multi-site diffusion tensor imaging data. *Neuroimage.* 2017;161:149–170.
60. Radua J, Vieta E, Shinohara R, et al.; ENIGMA Consortium collaborators. Increased power by harmonizing structural MRI site differences with the ComBat batch adjustment method in ENIGMA. *Neuroimage.* 2020;218:116956.
61. Fortin JP, Cullen N, Sheline YI, et al. Harmonization of cortical thickness measurements across scanners and sites. *Neuroimage.* 2018;167:104–120.
62. Sasabayashi D, Takayanagi Y, Takahashi T, et al. Subcortical brain volume abnormalities in individuals with an at-risk mental state. *Schizophr Bull.* 2020;46(4):834–845.
63. Cuesta MJ, Lecumberri P, Cabada T, et al. Basal ganglia and ventricle volume in first-episode psychosis. A family and clinical study. *Psychiatry Res Neuroimaging.* 2017;269:90–96.
64. Cuesta MJ, Lecumberri P, Moreno-Izco L, et al. Motor abnormalities and basal ganglia in first-episode psychosis (FEP). *Psychol Med.* 2021;51(10):1625–1636.
65. Fan F, Xiang H, Tan S, et al. Subcortical structures and cognitive dysfunction in first episode schizophrenia. *Psychiatry Res Neuroimaging.* 2019;286:69–75.
66. King BH, Lord C. Is schizophrenia on the autism spectrum? *Brain Res.* 2011;1380:34–41.
67. Mazefsky CA, Oswald DP, Day TN, Eack SM, Minshew NJ, Lainhart JE. ASD, a psychiatric disorder, or both? Psychiatric diagnoses in adolescents with high-functioning ASD. *J Clin Child Adolesc Psychol.* 2012;41(4):516–523.
68. Maikusa N, Zhu Y, Uematsu A, et al. Comparison of traveling-subject and ComBat harmonization methods for assessing structural brain characteristics. *Hum Brain Mapp.* 2021;42(16):5278–5287.