Schizophrenia Bulletin vol. 48 no. 1 pp. 241–250, 2021
https://doi.org/10.1093/schbul/sbab110
Advance Access publication September 11, 2021

Subtyping Schizophrenia Patients Based on Patterns of Structural Brain Alterations

Yuan Xiao[1,2,12], Wei Liao[3,12,*], Zhiliang Long[3], Bo Tao[1], Qiannan Zhao[1], Chunyan Luo[1], Carol A. Tamminga[4], Matcheri S. Keshavan[5], Godfrey D. Pearlson[6], Brett A. Clementz[7], Elliot S. Gershon[8], Elena I. Ivleva[4], Sarah K. Keedy[8], Bharat B. Biswal[9], Andrea Mechelli[10,*], Rebekka Lencer[2], John A. Sweeney[1,11], Su Lui*[,1,13], and Qiyong Gong[1,13,*]

[1]Huaxi MR Research Center (HMRRC), Department of Radiology, West China Hospital of Sichuan University, Chengdu, Sichuan, China; [2]Department of Psychiatry, University of Münster, Münster, Germany; [3]Center for Information in BioMedicine, School of Life Science and Technology, University of Electronic Science and Technology, Chengdu, Sichuan, China; [4]Department of Psychiatry, University of Texas Southwestern Medical Center, Dallas, TX, USA; [5]Department of Psychiatry, Beth Israel Deaconess Medical Center and Harvard Medical School, Boston, MA, USA; [6]Departments of Psychiatry and Neurobiology, Yale University and Olin Neuropsychiatric Research Center, Hartford, CT, USA; [7]Department of Psychology, University of Georgia, Athens, GA, USA; [8]Department of Psychiatry, University of Chicago, Chicago, IL, USA; [9]Department of Biomedical Engineering, New Jersey Institute of Technology, Newark, NJ, USA; [10]Department of Psychosis Studies, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK; [11]Department of Psychiatry and Behavioral Neuroscience, University of Cincinnati, Cincinnati, OH, USA; [12]These authors contributed equally to this work as joint first authors; [13]These authors contributed equally to this work as joint senior authors.

*To whom correspondence should be addressed; #37 GuoXue Xiang, Chengdu 610041, China; Tel: 86-28-85423960, Fax: 86-28-85423503; e-mail: lusuwcums@hotmail.com

Schizophrenia is a complex and heterogeneous syndrome. Whether quantitative imaging biomarkers can identify discrete subgroups of patients as might be used to foster personalized medicine approaches for patient care remains unclear. Cross-sectional structural MR images of 163 never-treated first-episode schizophrenia patients (FES) and 133 chronically ill patients with midcourse schizophrenia from the Bipolar and Schizophrenia Network for Intermediate Phenotypes (B-SNIP) consortium and a total of 403 healthy controls were recruited. Morphometric measures (cortical thickness, surface area, and subcortical structures) were extracted for each subject and then the optimized subtyping results were obtained with nonsupervised cluster analysis. Three subgroups of patients defined by distinct patterns of regional cortical and subcortical morphometric features were identified in FES. A similar three subgroup pattern was identified in the independent dataset of patients from the multi-site B-SNIP consortium. Similarities of classification patterns across these two patient cohorts suggest that the 3-group typology is relatively stable over the course of illness. Cognitive functions were worse in subgroup 1 with midcourse schizophrenia than those in subgroup 3. These findings provide novel insight into distinct subgroups of patients with schizophrenia based on structural brain features. Findings of different cognitive functions among the subgroups support clinical differences in the MRI-defined illness subtypes. Regardless of clinical presentation and stage of illness, anatomic MR subgrouping biomarkers can separate neurobiologically distinct subgroups of schizophrenia patients, which represent an important and meaningful step forward in differentiating subtypes of patients for studies of illness neurobiology and potentially for clinical trials.

Key words: Schizophrenia/heterogeneity/subgroup/structural MRI/brain alterations/neuroimaging

Introduction

Schizophrenia syndrome is a clinically heterogeneous and genetically complex illness, as reflected in a large number of independent risk genes for the condition[1–4] and extensive variability in clinical presentation, treatment outcomes, and neurobiological features.[5–12] This heterogeneity suggests that this syndrome may in fact represent a combination of disease entities with an overlap in clinical presentation.[13,14] In recent years, a number of studies have investigated this notion using a range of approaches.[5,15]

Studies of brain anatomy with magnetic resonance imaging (MRI) in schizophrenia have typically utilized case-control designs to identify group-level brain alterations in patient cohorts.[5–11,16–21] This has led to the identification of regional gray matter alterations, such as decreased cortical thickness or area in frontal and temporal lobes, as well as widely distributed white matter alterations. While these group-level findings are of interest, the severity and specific features of neuroanatomical alterations vary widely across patients. It remains unclear whether

neuroanatomic variability is driven by varying degrees of illness severity or by qualitatively distinct schizophrenia subtypes.[5,22,23]

Clinical biomarker strategies using noninvasive MRI have the potential to identify neurobiologically distinct subtypes within complex syndromes such as schizophrenia,[5,24–27] which has the potential to advance diagnostic and therapeutic practice from a full reliance on behavioral features. Previous studies have shown a discrete heterogeneity of gray matter deficits,[22,28,29] though additional issues need to be addressed. For example, antipsychotic drug treatment and illness course, which both can affect brain structure and function[30–33] to influence subgroup identification. Investigating heterogeneity of brain alterations in both first-episode schizophrenia (FES) prior to any treatment and in chronically ill individuals can address these issues.[34,35]

Many different clustering algorithms are used to classify patients into subgroups on the basis of their similarity in symptoms, cognitions, or imaging features,[4,36,37] but no consensus has been reached even on the definition of cluster. A density peak-based clustering (DPC) algorithm[38] was recently proposed, which assumes that the cluster centers should satisfy the following two criteria: 1) higher density than their neighbors and 2) long distance with other high local density data.[38] Typical clustering algorithms, such as, K-means and K-medoids, always assign an observation to the nearest center, and can be less robust in detecting nonspherical clusters. Compared with these typical clustering algorithms, the DPC algorithm is based on the distance between data points, and thus is better able to detect nonspherical clusters and to automatically find the correct number of clusters,[38] without initialization and multiple iterations. Moreover, it is a parameter-free method.

In the current study, we examined cortical (surface area, cortical thickness) and subcortical (volume) morphology, in acutely ill patients with never-treated FES and clinically stable chronic patients with midcourse schizophrenia each with a matched healthy control group. Gray matter was chosen as the subgrouping feature because it comprises neuronal cell bodies, neuropil, glial cells, and synapses, and is well documented to be altered in schizophrenia. The novel clustering algorithm-DPC was used to intuitively classify schizophrenia patients into subgroups with distinct neuroanatomical features.[38] The primary aim of this study was to identify the number and features of discrete subgroups in a sample of FES patients. We then determined whether a similar pattern of neuroanatomically-defined patient subgroups could be found in an independent sample of clinically stable patients with midcourse schizophrenia. Additionally, we tested for differences in clinical features in identified subgroups. We hypothesized that: 1) 2 or 3 subgroups of patients would be found in FES, and the similar clustering solution would be also found in the midcourse patients; 2) the different subgroup of patients would demonstrate different clinic or imaging features.

Methods

Participants and Clinical Measures

The FES sample was comprised of 163 never-treated and acutely ill patients (mean age: 23.37 ± 7.45y, female: 89) and 173 healthy controls (mean age: 24.08 ± 6.39y, female: 83) from West China Hospital of Sichuan University in Chengdu, China (table 1). The second sample included 133 clinically stable patients with midcourse schizophrenia (mean age: 33.77 ± 11.99y, female: 50) and 230 healthy controls (mean age: 33.63 ± 10.42y, female: 120) from the Bipolar and Schizophrenia Network for Intermediate Phenotypes (B-SNIP) consortium collected via a multisite collaboration in the USA (table 1).[6] The study was approved by each local research ethics

Table 1. Demographic and clinical characteristics of never-treated first-episode schizophrenia patients, midcourse schizophrenia patients from the B-SNIP consortium study and the separate healthy controls for the two patient groups

Characteristic	Mean ± SD	Mean ± SD	P
	Antipsychotic-Naïve FES (N = 163)	HC1 (N = 173)	
Gender(M/F)	*N* (74/89)	*N* (90/83)	.22
Age (years)	23.37 ± 7.45	24.08 ± 6.39	.35
Education (years)	12.09 ± 3.09	12.80 ± 3.74	.06
Illness duration (months)	12.21 ± 20.08	–	–
GAF scores	30.10 ± 11.13	–	–
PANSS scores			
Total	95.58 ± 19.33	–	–
Negative	24.71 ± 6.36	–	–
Positive	18.73 ± 7.97	–	–
General	45.89 ± 9.82	–	–
	Midcourse SCZ (N = 133)	HC2 (N = 230)	
Gender (M/F)	*N* (83/50)	*N* (110/120)	<.01
Age (years)	33.77 ± 11.99	33.63 ± 10.42	.91
Education (years)	13.04 ± 2.34	14.98 ± 2.45	<.01
Illness duration (years)	11.72 ± 10.82	–	–
GAF scores	49.27 ± 12.19	–	–
PANSS scores			
Total	69.54 ± 17.40	–	–
Negative	17.00 ± 6.03	–	–
Positive	17.96 ± 5.39	–	–
General	34.62 ± 8.74	–	–

FES, first-episode schizophrenia; HC1, healthy controls from FES study; Midcourse SCZ, treated patients with midcourse schizophrenia from the B-SNIP consortium study; HC2, healthy controls from the B-SNIP consortium study; SD, standard deviation; GAF, Global Assessment of Functioning; PANSS, Positive and Negative Syndrome Scale

committee and written informed consent was obtained from all participants prior to study participation.

Diagnoses of schizophrenia were determined using the Structured Interview for the DSM-IV (SCID-P). Patients' overall function and symptom severity were assessed using the Global Assessment of Functioning (GAF) and Positive and Negative Syndrome Scale (PANSS). Additionally, the Brief Assessment of Cognition in Schizophrenia (BACS)[39] was administered to the patients with midcourse schizophrenia in order to evaluate cognitive functioning. For FES patients, psychiatric evaluations and imaging studies were performed prior to the administration of any lifetime antipsychotic treatment.

Imaging Data Acquisition and Data Preprocessing

All FES patients and healthy controls in the sample from China underwent structural MRI scanning using a 3.0T MR scanner (for scan parameters, see details in Supplementary table 1) at West China Hospital of Sichuan University, Chengdu, China. For each subject, high-resolution T1-weighted structural images were acquired using a spoiled gradient recall sequence.

For midcourse schizophrenia patients and healthy controls from the B-SNIP dataset, all participants underwent MRI scanning on 3.0T scanners across five sites (see details in Supplementary table 1); all subjects at each site were scanned on the same magnet for the duration of the study. T1-weighted Magnetization Prepared Rapid Gradient Echo (MPRAGE) or Inversion Recovery-Prepared Spoiled Gradient-Echo (IR-SPGR) sequences, as appropriate for scanner brands, were administered following the Alzheimer's Disease Neuroimaging Initiative (ADNI1) protocol (http://adni.loni.usc.edu/methods/documents/mri-protocols/). The sequence parameters are detailed in Supplementary table 1.

MRI images were independently inspected by an experienced neuroradiologist to check for distortions and artifacts affecting image quality and to exclude patients with visible cerebral abnormalities of neuroradiological significance.

Structural image processing was carried out to reconstruct the brain's cortical surfaces and included the following steps: (1) spatial noise removal using the classical nonlocal means of filtering, which averaged vertices weighted by the similarity of their neighborhoods, to deal with magnetic resonance images with spatially varying noise levels (e.g., Gaussian distributed noise); (2) cortical surface reconstruction; (3) tissue segmentation of cerebrospinal fluid (CSF), white matter (WM), and gray matter (GM); (4) triangular mesh tessellation over the GM-WM boundary and mesh deformation; (5) correction of topological defects on the surface; (6) individual surface mesh inflation into a sphere; (7) estimation of the deformation between the resulting spherical mesh and a common spherical coordinate system.

Surface-Based Features and Subcortical Feature Generation

For subject-wise neuroanatomical profiling, we computed 3 features previously reported as abnormal in schizophrenia[40] using the FreeSurfer 6.0 software package. Individual surface maps were projected onto the fsaverage cortical surface grid. "Cortical thickness" was defined as the distance of corresponding vertices between the gray-white and pial boundary. "Cortical surface area" was defined as the area of triangles surrounding a vertex along the white matter interface. Subcortical volumes were obtained from seven regions: thalamus, caudate, putamen, pallidum, hippocampus, amygdala, and nucleus accumbens. Cortical thickness and cortical surface area measures were extracted for 34 gray matter regions in each hemisphere (68 ROIs from the Desikan-Killiany atlas, neuroanatomic features presented in Supplementary table 2). In this case, 150 (68*2+7*2) neuroanatomical features were obtained for each participant.

Subtype Analysis

Neuroanatomical subtyping analysis was performed in each of the two patient samples independently. First, a 150 x N matrix (N = number of patients) was generated. Then the effects of age, gender and education level were regressed out after generating the matrix. We next identified relevant, nonredundant neuroanatomic features for clustering individuals, reasoning that a low-dimensional representation would yield more valid/replicable findings. Principal component analysis was employed to obtain eigenvalues and corresponding principal components. The component with the largest variance accounted for was used for the next step that involved calculating the dissimilarity (or distance) between any pair of patients using "*pdist*.m" with Mahalanobis distance in Matlab. This resulted in a N x N dissimilarity matrix, in which each cell provided a measure of dissimilarity. A higher value denoted lower similarity of anatomical patterns between patients. This dissimilarity matrix was then used for the subsequent cluster analysis. A density peak-based clustering (DPC) algorithm[38] was employed to classify schizophrenia patients into subgroups with distinct neuroanatomical features. This DPC algorithm assumes that the cluster centers should satisfy the following two criteria: 1) higher density than their neighbors and 2) long distance with other high local density data (see figure 4 in ref 39).[38]

Specifically, given a data point i (e.g., one patient), the local density ρ_i was defined as:

$$\rho_i = \sum_j \chi(d_{ij} - d_c)$$

Where, $\chi(d_{ij} - d_c) = \begin{cases} 1 \ if \ (d_{ij} - d_c) < 0 \\ 0 \ if \ (d_{ij} - d_c) = 0 \end{cases}$, and the d_c is a constant. The ρ_i was estimated using a Gaussian kernel. The algorithm is robustly independent of the d_c for a large dataset because it is sensitive only to the relative magnitude of local density ρ_i in different points.

The δ_i was defined as:

$$\delta_i = \min_{j:\rho_j > \rho_i}(d_{ij})$$

After computing ρ_i and δ_i for all data, those data were considered as cluster centers if they had high ρ_i value and δ_i values, followed by assigning the remaining data to the same cluster as its nearest neighbor of higher density. Unlike a traditional cluster algorithm such as the K-means approach with a predefined number of clusters, this algorithm can determine the number of clusters intuitively with cluster centers characterized by high density (ρ_i) and large distance (δ_i).

To explore illness-related heterogeneity, neuroanatomical schizophrenia subtyping analysis using DPC was first applied to the cohort of never-treated FES patients (n = 163). Then, to explore whether the illness heterogeneity also existed in later course of illness in schizophrenia, the same DPC subtyping analysis was performed in midcourse schizophrenia patients (B-SNIP study, n = 133).

To explore whether the clustering solutions for FES and midcourse schizophrenia samples were similar, we combined the two datasets, and performed the subtyping analysis once more on the pooled sample.

Distance within (density) or/and between clusters was estimated to assess the similarity of subgroups between FES dataset and the combined dataset. Additionally, correlation analysis was also performed to explore the morphometric alteration pattern similarity between FES subgroups and midcourse patient subgroups.

Group Comparison Analysis

Pairwise group comparison analyses of cerebral structural measures were conducted between each subgroup of patients and healthy controls covarying for age, gender and site (for B-SNIP sample). For surface area and volume analyses, total intracranial volume was an additional covariate. Comparison of subgroups was corrected for multiple hypothesis testing using the FDR procedure for all 150 features. Comparison of subgroups was corrected for multiple hypothesis testing using the FDR procedure for all 150 features (FDR rate was set as q = .05).

Results

Neuroanatomical Delineation of Schizophrenia Subtypes

Three clusters or subgroups for FES were determined by the result that no more than three clusters satisfied both criteria of high density and large distance criterion. As we can see from the figure below, there are three colored data points that were selected and determined as three clusters (cut-off) value based on the rules mentioned above (see figure 1). A similar three subgroup solution pattern was identified in the independent dataset of midcourse schizophrenia patients from the multi-site B-SNIP consortium (see figure 2). After analyzing the FES and midcourse schizophrenia samples separately, we performed a third cluster analysis on the pooled samples. This third analysis also yielded a three subgroup solution (see figure 2). Importantly, allocation of patients to groups 1–3 in the pooled data solution was almost identical to the allocation of patients to groups (1–3) when the two patient samples were considered separately (figure 2). Only 5 of 296 patients across the two samples were not allocated to the same groups in the pooled sample analysis as they were in the original single-sample cluster analyses. In subgroup 1, no patient was reclassified in either patient sample. Healthy comparison subjects could not be classified into subtypes using this same algorithm, suggesting a relatively homogeneous pattern of brain anatomic features within the control samples. The distance within and between clusters in FES dataset was similar to that in the combined dataset and the longer-term ill sample (Supplementary table 4).

Patterns of Structural Brain Alterations in Schizophrenia Subgroups Relative to Healthy Controls

We explored the brain alteration patterns in each FES and midcourse schizophrenia patients separately. FES patients in subgroup 1 showed decreased surface area, thickness, and volume, mainly in cortical-thalamic-cortical circuitry and increased thickness in left rostral anterior cingulate gyrus (figure 3, Supplementary table 2), while FES patients in subgroup 2 and subgroup 3 showed no significant cortical or subcortical alteration.

In midcourse schizophrenia patients from the B-SNIP study, relative to healthy controls, patients in subgroup 1 showed widespread gray matter deficits in all lobes, and in insular cortex and bilateral hippocampus, while showing increased gray matter volume in bilateral pallidum (figure 3, Supplementary table 3). Patients in subgroup 2 showed decreased gray matter volume in left hippocampus (figure 3, Supplementary table 3). Patients in subgroup 3 showed no significant brain alteration (figure 3, Supplementary table 3).

Brain Structural Similarities and Differences Between FES and Midcourse Schizophrenia Samples

Next, comparison of cerebral alterations of FES and midcourse schizophrenia patients was performed. For subgroup 1, patients from the two datasets shared similar cortical surface area deficits in bilateral superior frontal, bilateral rostral middle frontal, bilateral middle

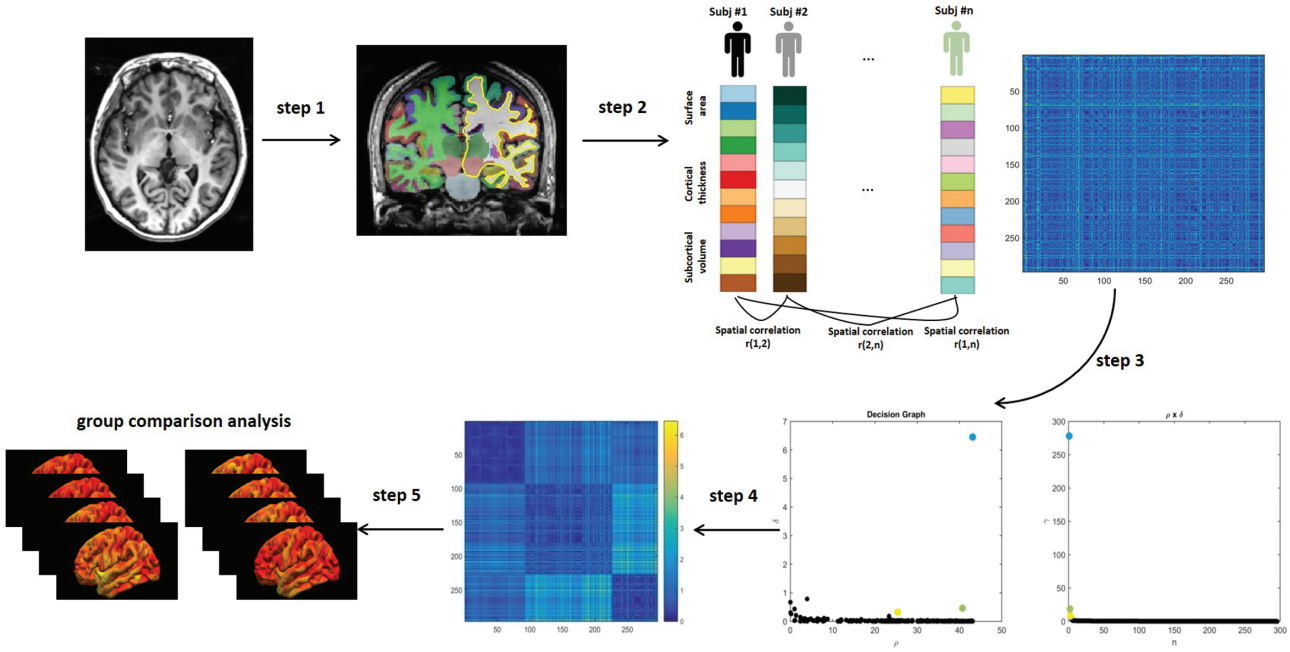
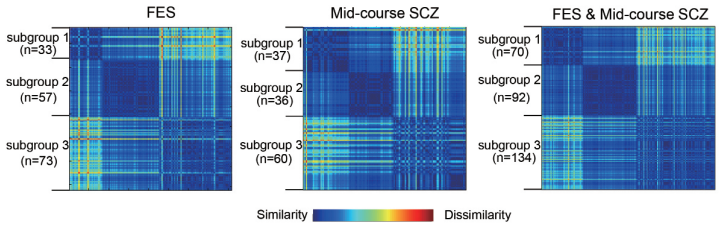
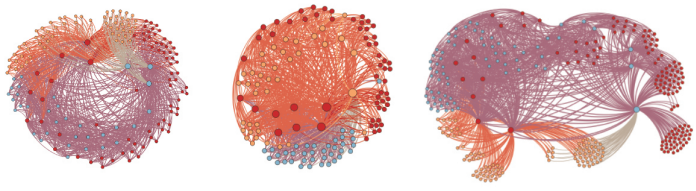


Fig. 1. Schematic diagram of subtyping schizophrenia patients using density peak-based clustering analysis. Step 1: We first obtained the cortical and subcortical structure measures from high-resolution 3D T1-weighted images. Step 2: Surface area, cortical thickness, and subcortical volume measurements from each subject are concatenated into one eigenvector. Then, the correlation of the feature vectors between the two subjects is calculated to represent the similarity between subjects. The process was repeated until the similarity between any two subjects is obtained, and the dissimilarity matrix of the subjects can be obtained (n is the number of participants). Step 3: Subtype analysis: A density peak-based clustering (DPC) algorithm was employed to intuitively classify schizophrenia patients into subgroups with distinct neuroanatomical features. This DPC algorithm assumes that the cluster centers should satisfy the following two criteria: 1) higher density than their neighbors and 2) long distance with other high local density data. Decision graph for the pooled patients showing the 3 clusters solution. (Left panel) The x-axis denotes the local density ρ_i and the y-axis denotes the minimal distance δ_i for all patients. (Right panel) The value of $\gamma_i = \rho_i \delta_i$in decreasing order for all patients. As we can see, there are three clusters or subgroups for this study. Step 4: The re-ordered dissimilarity (Mahalanobis distance) matrix for pooled patients according to three cluster labels. Note the three identified subgroups in dark blue shown aligned along the principal diagonal. Step 5: Group comparison analyses between each subgroup of patients and healthy controls.





Samples		Subgroup 1	Subgroup 2	Subgroup 3
FES	Number	33	57	73
	Percentage	20.20%	35.00%	44.80%
Mid-course SCZ	Number	37	36	60
	Percentage	27.80%	27.10%	45.10%
Pooled data	Number	70	92	134
	Percentage	23.60%	31.10%	45.30%
	FES:Mid-course SCZ	33:37	55:37	75:59

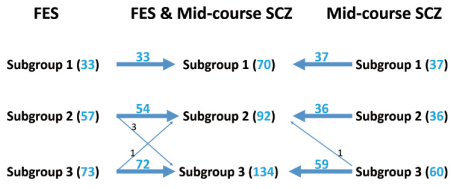


Fig. 2. Clustering results. (left upper panel) Clustering results illustrated by dissimilarity matrices for treatment naïve first episode schizophrenia patients (FES), midcourse schizophrenia patients from Bipolar and Schizophrenia Network for Intermediate Phenotypes (B-SNIP) study and the combined sample. (left lower panel) Force-directed graph of patients with schizophrenia created using Gephi software (https://gephi.org/). The nodes colored differently denote the patients with schizophrenia belonging to different subgroups. Subtyping analysis revealed 3 subgroups of patients in FES, midcourse schizophrenia patients and the pooled analyses of these two samples of schizophrenia patients. (right upper panel) Clustering results showing the number and percentage of patients sorted into each of the three clusters in FES patients, midcourse schizophrenia patients and the combined samples. (right lower panel) Individuals in subgroups 1–3 in both patient samples were sorted into the same respective subgroup (1–3) in the pooled group solution with the exception of only 5 of 296 patients shifting between subgroups 2 and 3 relative to their grouping in the single patient group analyses.

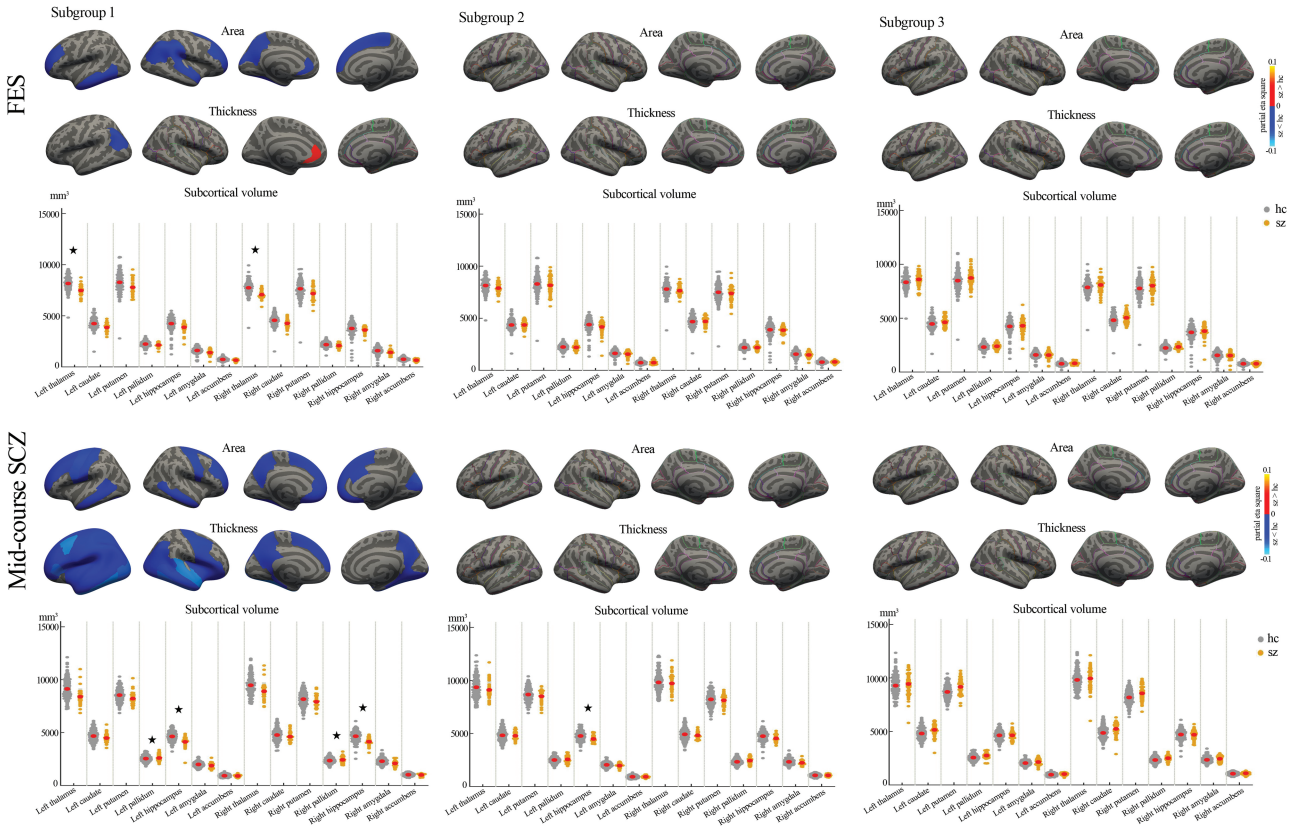


Fig. 3. Cortical/subcortical structural alterations in the three subgroups of FES and midcourse schizophrenia patients compared with healthy controls (red color: increased cortical measure in patients; blue color: decreased cortical measure in patients; asterisk * represents a significant group difference with FDR correction).

temporal, left rostral anterior cingulate, left fusiform, and left precuneus cortex. Further, patients with midcourse schizophrenia showed more widespread cortical/subcortical alterations than FES patients, especially decreased thickness throughout the whole brain and increased bilateral pallidum volume (see Supplementary Materials and Supplementary tables 2–3). Pearson correlation analysis demonstrated that these two subgroup 1 morphometric alteration patterns were highly consistent (Pearson's $r = .83$, $P < .001$, Supplementary figure 1). For subgroup 2, patients with midcourse schizophrenia showed more volume deficits in left hippocampus than FES patients. For subgroup 3, they both showed no significant neuroanatomical alteration. Thus, while the basic structure of subgroup classification was retained in later-course patients, differences between the two patient samples indicate that certain aspects of neuroanatomy that may be influenced by illness progression or antipsychotic treatment as suggested previously.[31–33,35,41]

Clinical Features in the Three Subgroups

Detailed demographic and clinical features of three subgroups were presented in Supplementary table 5. Subgroup 2 of FES showed a higher ratio of female patients relative to other subgroups. Subgroup 3 of midcourse schizophrenia demonstrated higher level of education years than other subgroups. Exploratory analyses in the B-SNIP sample revealed that BACS scores differed among the three subgroups ($F = 4.32$, $P = .02$), with patients from subgroup 1 showing lower Composite-Z scores than subgroup 3 ($P < .05$, Tukey's post hoc test, figure 4). Analysis of subtest scores revealed significant differences between subgroups 1 and 3 on digit sequencing and verbal fluency tests ($P < .05$, Tukey's post hoc test). There were no significant differences in PANSS scores between the three identified subgroups in either the FES or the midcourse schizophrenia patients.

There were no significant differences in daily chlorpromazine-equivalent antipsychotic medication dosage among the three subgroups of midcourse patients.

Relationships Between the Three Subgroup Solution and the Three Biotype Solution for the Midcourse Schizophrenia Patients

The B-SNIP consortium previously identified a 3 group (Biotype) differentiation of psychotic disorder patients using cognitive and ERP data.[5] We calculated the overlap in classification of our 3-group solution using MRI data

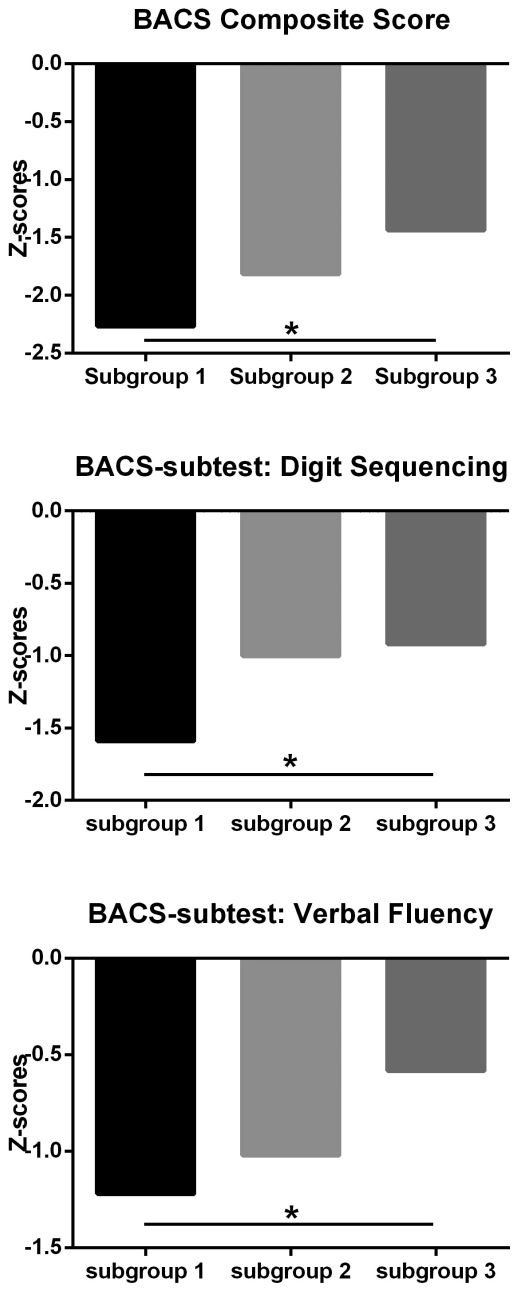


Fig. 4. Cognitive performance among three subgroups of patients with midcourse schizophrenia. Score of zero = mean of performance of healthy individuals as determined by the healthy group used to norm the test; BACS composite Z-scores (or subtests on digit sequencing/verbal fluency) (Y axis) presented as standard deviation (SD) for each patient in midcourse schizophrenia from mean performance of test norm value, using SD for calculations. (asterisk * represents a significant group difference at $P < .05$).

and the 3 Biotype solution of the B-SNIP program using Dice coefficients (Supplementary tables 6–7). Then, permutation testing (10 000 times) was performed to test whether the Dice coefficients were different from those when patients were randomly distributed into three sets. Results show that patients of subgroup 1 characterized by widespread dystrophic structural brain changes overlapped with Biotype 1, which is characterized by pronounced cognitive and electrophysiological alterations ($P < .01$, Supplementary table 8).

Discussion

Using structural MRI in a large sample of acutely ill never-treated FES patients and a second group of clinically stable, midcourse patients, we demonstrated that patients can be classified into three subgroups defined by distinct patterns of cortical and subcortical anatomic features. This clustering of patients using MRI data was remarkably consistent across the two patient samples, suggesting that our reported patient classification was robust to variations in ethnicity, stage of illness, and treatment status. These findings and the analyses of B-SNIP biotype classification are consistent with discrete patterns of heterogeneity of schizophrenia, suggesting distinct neurobiological bases for illness manifestations despite overall similarity in symptom expression. Clinical/behavioral features have been shown over many years to have limited utility in stratifying schizophrenia patients for the purpose of clinical trials and personalized treatment. The present findings represent an important step towards the stratification of patients on the basis of neurobiological rather than behavioral features as a strategy for advancing psychiatric patient care.[42]

The distinct patterns of gray matter alterations in the three patient subgroups reveals the particular brain features that were leveraged to distinguish the patient subgroups. Midcourse schizophrenia patients in subgroup 3 with a pattern of no obvious brain alteration showed better cognitive function than subgroup 1 of patients who demonstrated widespread reductions in cortical and subcortical measurements. These findings suggest a behavioral relevance to the MR subgroups, and by inference a relevance for general functional ability. MRI and postmortem studies of schizophrenia at the group level have consistently shown decreases in these cortical and subcortical parameters,[23,43] so it is novel to demonstrate that a substantial subgroup of patients have no obvious alteration in brain anatomy. The remaining patients, especially in subgroup 1, showed more widespread neuroanatomic abnormalities relative to healthy controls.

Patients in subgroup 1 showed widespread decreases of cortical surface area and thickness and subcortical volume. Such alterations may result from pathologically extended pruning processes[41] or sublethal apoptotic activity.[44] Our finding of this subgroup is not novel, but now with the improved MR acquisitions and development of statistical tools for classification and subtyping, we could quantify the size of the subgroup and its separateness from other patients meeting DSM criteria for schizophrenia as a global syndrome. In contrast to the general pattern of dystrophic changes, increased regional cortical

thickness was seen in left rostral anterior cingulate in subgroup 1 of FES patients but not in midcourse patients. That change could result from the early course of illness pathology, such as preapoptotic osmotic changes,[45] or hypertrophy of synaptic connections.[46] In subgroup 1 patients with midcourse schizophrenia, all lobes showed decreased alterations, especially for the thickness in frontal and temporal lobes. The greater and more widespread pattern of deficits in the midcourse patients (figure 3 and Supplementary tables 2–3) could result from a widening expression of atrophic changes related to illness course. Specifically, patients with midcourse schizophrenia were more likely to show widely reduced cortical thickness than patients with FES, reflected in the fact that 8% more patients and 8% fewer patients were classified into subgroups 1 and 2 in midcourse than FES patients. This pattern may result from brain changes previously reported to be associated with illness duration or antipsychotic treatment.[32,35,43,47–50]

Patients in subgroup 2 with FES showed no significant cortical or subcortical alteration, while patients with midcourse schizophrenia exhibited more decrease gray matter volume in left hippocampus than FES. This pattern may result from ill duration and antipsychotic treatment. The moderate morphological alteration in midcourse schizophrenia was consistent with moderate cognitive deficits among the three subgroups, which support the morphological basis for cognition.

The consistency of the classification achieved via machine learning across our two patient samples not only provides evidence for the validity of the pattern of patient subgrouping, but shows that despite differences in ethnicity, antipsychotic treatment, and illness duration, the core structure of neuroanatomic features differentiating our three subgroups was consistent. To the best of our knowledge, this is the first study using neuroimaging data (combined cortical and subcortical gray matter features) and a data-driven analytic method to identify reliable biologically-defined subtypes of the schizophrenia syndrome across never-treated FES and midcourse schizophrenia patients.

It is noteworthy that subgroup 1 patients in midcourse schizophrenia with widespread gray matter deficits showed greater cognitive deficits, especially in working memory and verbal fluency, while those who displayed no obvious brain alteration showed less impaired cognition. While showing a relation of brain atrophy to cognitive impairment is not surprising, it does demonstrate clinical relevance of the patient subgrouping and supports the validity of the anatomic subtyping-based classification. The overlap with the B-SNIP biotypes based on ERP and cognitive measures provides parallel validation.

There are certain limitations to our study that merit comment. First, given the cross-sectional design, our study cannot trace possible natural dynamic alterations of heterogeneity patterns over the course of illness. But by including both patients with FES prior to any treatment and treated patients with midcourse schizophrenia, our results suggest that the observed patterns appear to be similar in early and midlife adults. Second, we only recruited schizophrenia patients and healthy subjects, so future studies will be needed to establish the similarity of clustering patterns across the broader spectrum of psychotic disorders. Third, while our study is large and demonstrates informative clustering patterns across four discrete patient and control cohorts, and the relevance of drug treatment effects for MRI-based patient classification approaches, the clinical significance of our findings for prognosis and treatment planning remains to be established. Forth, cognition information is not available for the FES patients, which results in the lack of direct comparison of cognition between early stage and midcourse of illness.

The question of whether there are neurobiologically discrete subgroups of schizophrenia patients is a crucial one for personalized medicine, for drug discovery, and for discovery of illness etiology. Delineated forms of illness within the schizophrenia syndrome, identified in clinical research and then mechanistically understood, could dramatically change drug development and the treatment landscape of schizophrenia patient care. Failure to establish and effectively utilize such heterogeneity based on biomarkers independent of diagnostic behavioral features may be an important factor in the failure to develop novel effective treatments for the disorder in recent decades despite the many advances in clinical neuroscience, as different subgroups of patients may require different therapeutic strategies.[23,37,51–55] Here, analyzing structural neuroimaging data with a novel machine learning algorithm,[38] we demonstrate the existence of three reliable neurobiological subtypes irrespective of treatment and stage of illness. This finding represents a significant step forward toward developing biologically informed classification based on target organ abnormalities to complement behavioral analysis currently guiding psychiatric diagnostic practice. Findings such as ours may also help speed progress in gene and drug discovery, providing a framework for personalized medicine to improve outcomes for patients suffering from serious mental illness.[56–61]

Supplementary Material

Supplementary material is available at *Schizophrenia Bulletin*.

Funding

This work was supported by the National Natural Science Foundation of China (Grant Nos. 81901705, 82120108014, 81621003, 81671664, 81761128023, 81820108018, 81901702), the US National Institute of Mental Health (NIMH) grants (MH077862 [to JAS], MH077851 [to CAT], MH103366 [to BAC], MH078113

[to MSK], MH077945 [to GDP], MH096900 [to BAC] and MH077852 [to Gunvant K. Thaker]), and grants from the Humboldt Foundation (JAS, YX, SL), China Postdoctoral Science Foundation (2019M663513), Sichuan Science and Technology Program (2020YFS0116), the Postdoctoral Interdisciplinary Research Project of Sichuan University, and 1.3.5 project for disciplines of excellence, West China Hospital, Sichuan University (ZYJC18020, ZYYC08001). The funders had no further role in study design, data collection, data analysis, data interpretation, or writing of the report.

Acknowledgments

We acknowledge the many staff members who assisted in the collection of brain images and performed clinical assessments. We thank all participants for their generous contributions to our projects. GDP has served on an advisory panel for Bristol-Meyer Squibb. CAT serves on an advisory board for Karuna and KyNexis, and an advisory panel for Sunovion, Astellas, and Merck. MSK has received research support from Sunovion and GlaxoSmithKline. JAS consulted to VeraSci. The remaining authors report no competing interests.

References:

1. Fabbri C, Serretti A. Role of 108 schizophrenia-associated loci in modulating psychopathological dimensions in schizophrenia and bipolar disorder. *Am J Med Genet B Neuropsychiatr Genet.* 2017;174:757–764.
2. Zhao H, Nyholt DR. Gene-based analyses reveal novel genetic overlap and allelic heterogeneity across five major psychiatric disorders. *Hum Genet.* 2017;136:263–274.
3. Igolkina AA, Armoskus C, Newman JRB, et al. Analysis of gene expression variance in schizophrenia using structural equation modeling. *Front Mol Neurosci.* 2018;11:192.
4. Alnæs D, Kaufmann T, van der Meer D, et al.; Karolinska Schizophrenia Project Consortium. Brain heterogeneity in schizophrenia and its association with polygenic risk. *JAMA Psychiatry.* 2019;76:739–748.
5. Clementz BA, Sweeney JA, Hamm JP, et al. Identification of distinct psychosis biotypes using brain-based biomarkers. *Am J Psychiatry.* 2016;173:373–384.
6. Tamminga CA, Ivleva EI, Keshavan MS, et al. Clinical phenotypes of psychosis in the Bipolar-Schizophrenia Network on Intermediate Phenotypes (B-SNIP). *Am J Psychiatry.* 2013;170:1263–1274.
7. Malhotra AK. Dissecting the heterogeneity of treatment response in first-episode schizophrenia. *Schizophr Bull.* 2015;41:1224–1226.
8. Tsutsumi A, Glatt SJ, Kanazawa T, et al. The genetic validation of heterogeneity in schizophrenia. *Behav Brain Funct.* 2011;7:43.
9. Molina V, Blanco JA. A proposal for reframing schizophrenia research. *J Nerv Ment Dis.* 2013;201:744–752.
10. Nenadic I, Gaser C, Sauer H. Heterogeneity of brain structural variation and the structural imaging endophenotypes in schizophrenia. *Neuropsychobiology.* 2012;66:44–49.
11. Meda SA, Clementz BA, Sweeney JA, et al. Examining functional resting-state connectivity in psychosis and its subgroups in the bipolar-schizophrenia network on intermediate phenotypes cohort. *Biol Psychiatry Cogn Neurosci Neuroimaging.* 2016;1:488–497.
12. Kinon BJ. The group of treatment resistant schizophrenias. Heterogeneity in Treatment Resistant Schizophrenia (TRS). *Front Psychiatry.* 2018;9:757.
13. Takahashi S. Heterogeneity of schizophrenia: genetic and symptomatic factors. *Am J Med Genet B Neuropsychiatr Genet.* 2013;162B:648–652.
14. Weickert CS, Weickert TW, Pillai A, Buckley PF. Biomarkers in schizophrenia: a brief conceptual consideration. *Dis Markers.* 2013;35:3–9.
15. Ivleva EI, Clementz BA, Dutcher AM, et al. Brain structure biomarkers in the psychosis biotypes: findings from the bipolar-schizophrenia network for intermediate phenotypes. *Biol Psychiatry.* 2017;82:26–39.
16. Fettes P, Schulze L, Downar J. Cortico-striatal-thalamic loop circuits of the orbitofrontal cortex: promising therapeutic targets in psychiatric illness. *Front Syst Neurosci.* 2017;11:25.
17. Lui S, Zhou XJ, Sweeney JA, Gong Q. Psychoradiology: the frontier of neuroimaging in psychiatry. *Radiology.* 2016;281:357–372.
18. Kelly S, Jahanshad N, Zalesky A, et al. Widespread white matter microstructural differences in schizophrenia across 4322 individuals: results from the ENIGMA Schizophrenia DTI Working Group. *Mol Psychiatry.* 2018;23(5):1261–1269.
19. 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. *Molecular Psychiatry*. 2016;21:547–553.
20. Xiao Y, Lui S, Deng W, et al. Altered cortical thickness related to clinical severity but not the untreated disease duration in schizophrenia. *Schizophr Bull.* 2015;41(1):201–210.
21. Godwin D, Alpert KI, Wang L, Mamah D. Regional cortical thinning in young adults with schizophrenia but not psychotic or non-psychotic bipolar I disorder. *Int J Bipolar Disord.* 2018;6:16.
22. Honnorat N, Dong A, Meisenzahl-Lechner E, Koutsouleris N, Davatzikos C. Neuroanatomical heterogeneity of schizophrenia revealed by semi-supervised machine learning methods. *Schizophr Res*. 2017;214:43–50.
23. Williams MR, Chaudhry R, Perera S, et al. Changes in cortical thickness in the frontal lobes in schizophrenia are a result of thinning of pyramidal cell layers. *Eur Arch Psychiatry Clin Neurosci.* 2013;263:25–39.
24. Lerch JP, van der Kouwe AJ, Raznahan A, et al. Studying neuroanatomy using MRI. *Nat Neurosci.* 2017;20:314–326.
25. Drysdale AT, Grosenick L, Downar J, et al. Resting-state connectivity biomarkers define neurophysiological subtypes of depression. *Nat Med.* 2017;23:28–38.
26. Pan Y, Pu W, Chen X, et al. Morphological profiling of schizophrenia: cluster analysis of MRI-based cortical thickness data. *Schizophr Bull*. 2020;46:623–632.
27. Chand GB, Dwyer DB, Erus G, et al. Two distinct neuroanatomical subtypes of schizophrenia revealed using machine learning. *Brain*. 2020;143:1027–1038.
28. Mothi SS, Sudarshan M, Tandon N, et al. Machine learning improved classification of psychoses using clinical and biological stratification: update from the bipolar-schizophrenia network for intermediate phenotypes (B-SNIP). *Schizophr Res.* 2019;214:60–69.

29. Xiao Y, Yan Z, Zhao Y, et al. Support vector machine-based classification of first episode drug-naïve schizophrenia patients and healthy controls using structural MRI. *Schizophr Res.* 2019;214:11–17.
30. Reilly JL, Harris MS, Keshavan MS, Sweeney JA. Adverse effects of risperidone on spatial working memory in first-episode schizophrenia. *Arch Gen Psychiatry.* 2006;63:1189–1197.
31. Lui S, Li T, Deng W, et al. Short-term effects of antipsychotic treatment on cerebral function in drug-naive first-episode schizophrenia revealed by "resting state" functional magnetic resonance imaging. *Arch Gen Psychiatry.* 2010;67:783–792.
32. Keshavan MS, Bagwell WW, Haas GL, Sweeney JA, Schooler NR, Pettegrew JW. Changes in caudate volume with neuroleptic treatment. *Lancet.* 1994;344:1434.
33. Meng L, Li K, Li W, Xiao Y, Lui S, Sweeney JA, Gong Q. Widespread white-matter microstructure integrity reduction in first-episode schizophrenia patients after acute antipsychotic treatment. *Schizophr Res*. 2018;204:238–244.
34. Cropley VL, Klauser P, Lenroot RK, et al. Accelerated gray and white matter deterioration with age in schizophrenia. *Am J Psychiatry.* 2017;174:286–295.
35. Fusar-Poli P, Smieskova R, Kempton MJ, Ho BC, Andreasen NC, Borgwardt S. Progressive brain changes in schizophrenia related to antipsychotic treatment? A meta-analysis of longitudinal MRI studies. *Neurosci Biobehav Rev.* 2013;37:1680–1691.
36. Dickinson D, Zaidman SR, Giangrande EJ, Eisenberg DP, Gregory MD, Berman KF. Distinct polygenic score profiles in schizophrenia subgroups with different trajectories of cognitive development. *Am J Psychiatry.* 2020;177:298–307.
37. Dickinson D, Pratt DN, Giangrande EJ, et al. Attacking heterogeneity in schizophrenia by deriving clinical subgroups from widely available symptom data. *Schizophr Bull.* 2018;44(1):101–113.
38. Rodriguez A, Laio A. Machine learning. Clustering by fast search and find of density peaks. *Science.* 2014;344:1492–1496.
39. Hill SK, Reilly JL, Keefe RS, et al. Neuropsychological impairments in schizophrenia and psychotic bipolar disorder: findings from the Bipolar-Schizophrenia Network on Intermediate Phenotypes (B-SNIP) study. *Am J Psychiatry.* 2013;170:1275–1284.
40. Rimol LM, Nesvåg R, Hagler DJ Jr, et al. Cortical volume, surface area, and thickness in schizophrenia and bipolar disorder. *Biol Psychiatry.* 2012;71:552–560.
41. Doucet GE, Moser DA, Luber MJ, Leibu E, Frangou S. Baseline brain structural and functional predictors of clinical outcome in the early course of schizophrenia. *Mol Psychiatry*. 2018;25:863–872.
42. Tregellas JR. Neuroimaging biomarkers for early drug development in schizophrenia. *Biol Psychiatry.* 2014;76(2):111–119.
43. Dietsche B, Kircher T, Falkenberg I. Structural brain changes in schizophrenia at different stages of the illness: a selective review of longitudinal magnetic resonance imaging studies. *Aust N Z J Psychiatry.* 2017;51:500–508.
44. Collin G, Seidman LJ, Keshavan MS, et al. Functional connectome organization predicts conversion to psychosis in clinical high-risk youth from the SHARP program. *Mol Psychiatry*. 2018;25(10):2431–2440.
45. Ren W, Lui S, Deng W, et al. Anatomical and functional brain abnormalities in drug-naive first-episode schizophrenia. *Am J Psychiatry.* 2013;170:1308–1316.
46. Cui LB, Liu L, Wang HN, et al. Disease definition for schizophrenia by functional connectivity using radiomics strategy. *Schizophr Bull.* 2018;44:1053–1059.
47. Dorph-Petersen KA, Pierri JN, Perel JM, Sun Z, Sampson AR, Lewis DA. The influence of chronic exposure to antipsychotic medications on brain size before and after tissue fixation: a comparison of haloperidol and olanzapine in macaque monkeys. *Neuropsychopharmacology.* 2005;30:1649–1661.
48. Ho BC, Andreasen NC, Ziebell S, Pierson R, Magnotta V. Long-term antipsychotic treatment and brain volumes: a longitudinal study of first-episode schizophrenia. *Arch Gen Psychiatry.* 2011;68:128–137.
49. Konopaske GT, Dorph-Petersen KA, Pierri JN, Wu Q, Sampson AR, Lewis DA. Effect of chronic exposure to antipsychotic medication on cell numbers in the parietal cortex of macaque monkeys. *Neuropsychopharmacology.* 2007;32:1216–1223.
50. Lesh TA, Tanase C, Geib BR, et al. A multimodal analysis of antipsychotic effects on brain structure and function in first-episode schizophrenia. *JAMA Psychiatry.* 2015;72:226–234.
51. Picardi A, Viroli C, Tarsitani L, et al. Heterogeneity and symptom structure of schizophrenia. *Psychiatry Res.* 2012;198:386–394.
52. Ahmed AO, Strauss GP, Buchanan RW, Kirkpatrick B, Carpenter WT. Schizophrenia heterogeneity revisited: clinical, cognitive, and psychosocial correlates of statistically-derived negative symptoms subgroups. *J Psychiatr Res.* 2018;97:8–15.
53. Seaton BE, Goldstein G, Allen DN. Sources of heterogeneity in schizophrenia: the role of neuropsychological functioning. *Neuropsychol Rev.* 2001;11:45–67.
54. Weinberg D, Lenroot R, Jacomb I, et al. Cognitive subtypes of schizophrenia characterized by differential brain volumetric reductions and cognitive decline. *JAMA Psychiatry.* 2016;73(12):1251–1259.
55. Mirzakhanian H, Singh F, Cadenhead KS. Biomarkers in psychosis: an approach to early identification and individualized treatment. *Biomark Med.* 2014;8(1):51–57.
56. Goff DC, Romero K, Paul J, Mercedes Perez-Rodriguez M, Crandall D, Potkin SG. Biomarkers for drug development in early psychosis: current issues and promising directions. *Eur Neuropsychopharmacol.* 2016;26:923–937.
57. Roffman JL, Lamberti JS, Achtyes E, et al. Randomized multicenter investigation of folate plus vitamin B12 supplementation in schizophrenia. *JAMA Psychiatry.* 2013;70(5):481–489.
58. Rodrigues-Amorim D, Rivera-Baltanás T, López M, Spuch C, Olivares JM, Agís-Balboa RC. Schizophrenia: a review of potential biomarkers. *J Psychiatr Res.* 2017;93:37–49.
59. Dazzan P. Neuroimaging biomarkers to predict treatment response in schizophrenia: the end of 30 years of solitude? *Dialogues Clin Neurosci.* 2014;16(4):491–503.
60. Forray C, Buller R. Challenges and opportunities for the development of new antipsychotic drugs. *Biochem Pharmacol.* 2017;143:10–24.
61. Keshavan MS, Lawler AN, Nasrallah HA, Tandon R. New drug developments in psychosis: challenges, opportunities and strategies. *Prog Neurobiol.* 2017;152:3–20.