

Morphological Profiling of Schizophrenia: Cluster Analysis of MRI-Based Cortical Thickness Data

Yunzhi Pan^{1,2}, Weidan Pu³, Xudong Chen¹, Xiaojun Huang¹, Yan Cai^{1,4}, Haojuan Tao¹, Zhiming Xue¹, Michael Mackinley⁵, Roberto Limongi^{5,6,7}, Zhening Liu^{*,1}, and Lena Palaniyappan^{1,3,5,7}

¹Institute of Mental Health, Second Xiangya Hospital, Central South University, Changsha, PR China;

²Robarts Research Institution, University of Western Ontario, London, Canada;

³Medical Psychological Institute, Second Xiangya Hospital, Central South University, Changsha, PR China;

⁴The Second People's Hospital of Hunan Province, Changsha, Hunan, PR China;

⁵Department of Psychiatry, University of Western Ontario, London, Ontario, Canada;

⁶Pontificia Universidad Católica de Valparaíso, Región de Valparaíso, Chile;

⁷Lawson Health Research Institute, London, Ontario, Canada

*To whom correspondence should be addressed; Mental Health Institute, Second Xiangya Hospital of Central South University, Changsha, 410011 PR China, e-mail: zhening.liu@csu.edu.cn

The diagnosis of schizophrenia is thought to embrace several distinct subgroups. The manifold entities in a single clinical patient group increase the variance of biological measures, deflate the group-level estimates of causal factors, and mask the presence of treatment effects. However, reliable neurobiological boundaries to differentiate these subgroups remain elusive. Since cortical thinning is a well-established feature in schizophrenia, we investigated if individuals (patients and healthy controls) with similar patterns of regional cortical thickness form naturally occurring morphological subtypes. K-means algorithm clustering was applied to regional cortical thickness values obtained from 256 structural MRI scans (179 patients with schizophrenia and 77 healthy controls [HCs]). GAP statistics revealed three clusters with distinct regional thickness patterns. The specific patterns of cortical thinning, clinical characteristics, and cognitive function of each clustered subgroup were assessed. The three clusters based on thickness patterns comprised of a morphologically impoverished subgroup (25% patients, 1% HC), an intermediate subgroup (47% patients, 46% HC), and an intact subgroup (28% patients, 53% HC). The differences of clinical features among three clusters pertained to age-of-onset, N-back performance, duration exposure to treatment, total burden of positive symptoms, and severity of delusions. Particularly, the morphologically impoverished group had deficits in N-back performance and less severe positive symptom burden. The data-driven neuroimaging approach illustrates the occurrence of morphologically separable subgroups in schizophrenia, with distinct clinical characteristics. We infer that the anatomical

heterogeneity of schizophrenia arises from both pathological deviance and physiological variance. We advocate using MRI-guided stratification for clinical trials as well as case-control investigations in schizophrenia.

Key words: heterogeneity of schizophrenia/cortical thickness/clustering analysis

Introduction

Schizophrenia is a complex and persistent mental disorder with a variable course, often associated with dramatic deterioration in functioning. Unfortunately, despite decades of clinical research, only one in six patients with schizophrenia achieve rates of clinical and community functioning on par with their premorbid levels.¹ This variability in treatment response has led a number of studies to posit that schizophrenia, rather than a single disorder, may represent a group of distinct entities with overlapping clinical phenotypes.²⁻⁴ The idea of heterogeneity in schizophrenia is not new, with Kraepelin admitting that dementia praecox is “the expression of a single morbid process, though outwardly they often diverge very far from one another.”⁵ However, despite continued effort, focus on subclassifying patients based on strictly clinical presentation has shown little prognostic value. Advances in Magnetic Resonance Imaging (MRI) technology have provided researchers with a widely available and biologically safe method to investigate the posited presence of neurobiologically derived subgroups in patient populations.

In a large-scale multisite study, Clementz et al used multivariate taxometric analyses of MRI to identify specific biotypes of psychosis and found three neurobiologically distinct psychosis biotypes that did not conform to typical diagnostic boundaries between schizophrenia, schizoaffective disorder, and bipolar disorder.⁶ This suggests that biomarker-based stratification could necessitate a reconceptualization of traditional diagnostic classifications.¹⁻³

Recent imaging studies have revealed that schizophrenia patients show regional cortical thinning in several brain areas,⁴ and this has become an area of interest to assess the potential for neurobiological heterogeneity in schizophrenia.⁵⁻⁸ Although this is a promising line of inquiry, identifying the appropriate number of subtypes has proven difficult, with ranges from 3 to 6 groups being identified in previous anatomical MRI studies.⁵⁻⁸ Furthermore, the degree to which ethnicity will affect assessments of neurobiological subgroups should be assessed as many ethnic differences exist in incidence rates, illness severity, and degree of functional recovery.⁹

In most case-control studies, there is an absence of clear biological demarcation between patients and controls due to small effect sizes or a high degree of variance among patients. Nevertheless, most clustering studies seek biological boundaries only among patients, assuming that a natural distinction exists between patients and healthy controls.^{10,11} Thus, the effect of variation in the healthy brain (i.e., normative modeling) has not been utilized fully when studying the heterogeneity of schizophrenia.¹² Healthy controls do not form a neurobiologically homogeneous group that deviates from the patient subgroups, as within-group heterogeneity of features such as cortical thickness in healthy controls is substantial.¹³

Using a data set-based clustering approach for all participants (schizophrenia patients and healthy controls), we aim to resolve the inconsistency around the identity of biologically heterogeneous subtypes of schizophrenia. After clustering, the characteristics of each cluster will be revealed through clinical information such as diagnosis and symptoms and cognitive tasks. We sought to use a data-driven approach based on neurobiological traits to explore the distinct patterns of morphological variation and the nature of schizophrenia subtypes.

Methods

Participants

Patients ($n = 179$) with a diagnosis of schizophrenia (using the Structured Clinical Interview for DSM-IV-patient version [SCID-P]¹⁴) were recruited from the inpatient and outpatient units at Second Xiangya Hospital of Central South University, Changsha, China from 2009 to 2017. All patients: (1) met the DSM-IV diagnostic criteria for schizophrenia; (2) were 12–45 years of age;

(3) right-handed; and (4) had 9 or more years of formal schooling. The exclusion criteria included: (1) diagnosis of a substance-related disorders, neurological disorder, or a serious physical illness; (2) any contraindication for MRI; and (3) previous electroconvulsive therapy.

In addition to our patient population, we recruited $n = 77$ healthy controls (HCs) from a community sample in Changsha city. The inclusion and exclusion criteria were the same as those of the patient group, with the exception that controls were (1) not diagnosed with any mental illness according to the DSM-IV when interviewed using the SCID nonpatient version (2) did not have first-degree relatives with a psychotic illness. All participants gave their written informed consent to participate in our study after a detailed description of the risks and benefits. The study was approved by the ethics committee of the Second Xiangya Hospital, Central South University.

Clinical Assessment

Diagnoses were made by qualified psychiatrists according to DSM-IV criteria. On the same day as the MRI session, the severity of symptoms was evaluated through the Scale for The Assessment of Positive Symptoms (SAPS),¹⁵ the Scale for The Assessment of Negative Symptoms (SANS),¹⁵ and the Schizophrenia Suicide Risk Scale (SSRS).¹⁶ The duration of illness, antipsychotic load (converted into chlorpromazine equivalent per day), and duration of psychotic medication were recorded.

MRI Acquisition

MRI scanning was conducted on a Philips Gyroscan Achieva 3.0 Tesla MRI scanner at the Institute of Mental Health, Second Xiangya Hospital. High-resolution T1-weighted images were also acquired with a three-dimensional spoiled gradient echo (SPGR) pulse sequence from the sagittal plane, scanning parameter: TR = 7.5 ms, TE = 3.7 ms, FA = 8°, 180 slices, matrix = 256 × 200, the field of view (FOV) = 240 × 240 mm², and slices were contiguous with a slice thickness of 1 mm. Importantly, during the T1-weighted image acquisition, participants were asked to remain still, and if any motion-related artifacts were detected, the scans were repeated.

Cognitive Function Assessment

On the same day of the MRI acquisition, Verbal Fluency, N-back task and Contour Integration Test were administered. Verbal Fluency was tested by asking participants to report as many animals as possible within 60 s. The N-back task was widely used in previous studies of our research group.¹⁷⁻¹⁹ All participants performed a parametric n-back task on Nordic Neurolab's fMRI hardware system for 8 minutes and 16 seconds. All stimuli were sequences of white capital letters on a black background, presented centrally (500 ms duration, 1500 ms

inter-stimulus interval) in a pseudo-random order. The task performance, as represented by the reaction time [RT] and accuracy [AC], of each participant was recorded electronically.

The Contour Integration Test is designed to measure perceptual organization.^{20–22} This test is typically a task to recognize a closed contour circle made up of noncontiguous elements, embedded within a display of randomly oriented elements in a card. Further details of the test administration are provided in the [Supplementary Figure 1](#).

Principal component analysis (PCA) was applied across the whole patient group to reduce multiple comparisons by extracting the components that accounted for the majority of variance for each cognitive task. For N-back target accuracy of 0 back, 2 back, for the whole test and N-back nontarget accuracy (error rate of whole test) were entered into PCA; for contour task, the number of total correct, incorrect, and failure for random testing, total correct, incorrect, and failure for standard testing, were entered into PCA; for verbal fluency task, number of correct responses, wrong responses, and repetitions were entered into PCA. We extracted one principal component for the N-back scores (accounted for 59% of variance); two components were extracted for the contour task scores (component 1 accounted for 50% of variance and component 2 accounted for 24% of variance) with four items in component 1 (omitted score) and three items in component 2 (correct committed score). Two components were extracted for the verbal fluency scores (component 1 accounted for 43% of variance and component 2 accounted for 33% of variance) with two items in component 1 (correct response score) and one item in component 2 (noncategory responses score). We used PCA as the original variables within each test are correlated highly with each other and only the latent components of overall test performance were needed for our purpose of correlating with external variable (in this case cluster membership). Furthermore, there are no universally agreed single composite indices for reporting n-back and contour integration test performance.

Preprocessing of MRI Data

A surface-based approach using Free-Surfer (<http://surfer.nmr.harvard.edu>, version 5.3.0) was used to calculate the cortical thickness in the whole brain. Following skull-stripping and intensity correction, the gray–white matter boundary for each cortical hemisphere was determined by tissue intensity and neighborhood constraints. The resulting surface boundary was tessellated to generate multiple vertices across the whole brain before inflating. Using a deformable surface algorithm guided by the gray–CSF intensity gradient, the resulting gray–white interface was expanded to create the pial surface. The inflated surface was then morphed into a sphere followed

by registration to an average spherical surface for optimal sulcogyral alignment. After the above procedures, Desikan–Kiliany Atlas (68 regions) was used to extract cortical thickness of each region using the FreeSurfer software.²³ Topological defects were corrected manually by two members of the research staff.

Statistical Analysis

Using cortical thickness of 68 regions, we used the K-means clustering method and GAP statistics to identify clusters of participants who shared similar patterns of cortical thickness. K-means clustering was applied to all participants, including HCs. We set K number from 1 to 6 and GAP statistics to estimate the optimal number of clusters in our data. Then we chose the smallest K number that conformed to $\text{Gap}(k) \geq \text{Gap}(k+1) - s_{k+1}$ as the solution of cluster analysis based on the 1-standard-error method suggested by Tibshirani.²⁴ Based on the coordinates of each cluster center, we computed the distance from each individual to each of the three centers.

One-way ANOVA (SPSS 20.0) was used to compare morphological, clinicodemographic, and cognitive indices, with Bonferroni correction to address inflated type 1 error. For data with non-normal distribution (e.g., percentile data on the accuracy of N-back), we used nonparametric Kruskal–Wallis test for statistical analysis. We also investigated the correlation between distance from each cluster center and clinical and cognitive scores. At last, a multivariate generalized linear model with the subgroup based on clusters as the fixed factor was used to test the effect size of all factors including morphological data and phenotypic characteristics.

Results

Demographic and Clinical Characteristics of All Participants

A total of 256 participants (179 SCH, 77 HC) were recruited for the study. The demographic and clinical variables of participants are presented in [Table 1](#). Significant differences were found in gender ($P = .014$), Information-WAIS ($P < .00001$), Digit symbol-WAIS ($P < .00001$), and education ($P < .00001$), but not in age ($P = .288$) between the two diagnostic groups. As expected patients showed significant cognitive impairment compared with HC in three cognitive tasks, including contour task (visual integration), verbal fluency task (language fluency), and N-back task (working memory) ([Table 1](#)).

K-Means Clustering and GAP Statistics

We explored the possibility of the existence of one to six clusters and identified the ideal cluster solution based on GAP statistics to be 3 ([Figure 1A](#)). According to the proportion of patients compared with controls in

Table 1. Participant Demographic Information, Symptom, and Cognitive Scores

	SCH (mean ± SD)	HCs (mean ± SD)	P value (uncorrected)
N	179	77	
Age [range]	23.63 ± 5.77 [13–44]	24.52 ± 5.63 [18–42]	.288
Gender (female/male)	61/117	39/38	.014*
Education	11.58 ± 2.42	14.05 ± 2.25	<.00001**
Information-WAIS	15.81 ± 5.30	21.16 ± 4.53	<.00001**
Digit symbol-WAIS	62.30 ± 15.45	89.46 ± 14.53	<.00001**
Duration_of_Medicine (Days)	198 ± 445	—	—
Dosage_of_Medicine (CPZ equivalent)	134 ± 117	—	—
Duration_of_illness (months)	25.42 ± 32.66	—	—
Onset_age	21.59 ± 5.48	—	—
SAPS scores	20.60 ± 15.63	—	—
SANS scores	33.54 ± 26.52	—	—
Cognitive task			
N-back	0.78 ± 0.25	0.92 ± 0.15	.000007**
Textdisplay2_ACC	0.83 ± 0.27	0.94 ± 0.16	.000296**
N-back	0.52 ± 0.25	0.76 ± 0.19	<.00001**
Textdisplay1_ACC	0.86 ± 0.19	0.78 ± 0.22	.001*
N-back Target_ACC	43.75 ± 4.16	46.07 ± 3.69	.001655*
Contour Random total correct	3.32 ± 6.31	2.38 ± 3.90	.2486
Contour Random total wrong	70.77 ± 8.81	77.54 ± 7.74	.00002**
Contour Standard total correct	7.15 ± 10.84	6.49 ± 6.52	.62772
Verbal fluency correct	13.98 ± 5.06	20.37 ± 5.46	<.00001**
Verbal fluency wrong	0.17 ± 0.40	0.10 ± 0.31	.373
Verbal fluency repeat	0.68 ± 0.94	0.81 ± 0.98	.397

Note: After Bonferroni correction, the significant difference level was 4.16e-4.

*P < .05; **P < 4.16e-4.

each cluster, we deduced that cluster 1 (98% are schizophrenia) was “schizophrenia-like,” cluster 2 (67% are schizophrenia) was “HCs-like,” and cluster 3 (59% are schizophrenia) was intermediate (Figure 1B). When the individual features (regional thickness) were examined in patients compared with all HCs, patients had cortical thinning patterns that differed according to their cluster membership. Schizophrenia patients in cluster 2 appeared to be “morphologically intact,” with a pattern of cortical thickness similar to HCs; schizophrenia patients in cluster 1, “morphologically impoverished,” appeared to have a pattern of widespread cortical thinning. Schizophrenia

patients in cluster 3, “intermediate,” showed regional cortical thinning compared with HCs (Figure 1C). The central point (CP) of each cluster supported above different patterns of clusters (Supplementary Figure 5).

Among the three subgroups clusters, the effect size of differences (partial eta squared) in thickness for left parstriangularis area, left temporal pole, right fusiform area, bilateral middle temporal cortex, and bilateral superior temporal cortex was >0.5 (Supplementary Table 3), indicating a critical role for these regions in the observed heterogeneity of schizophrenia.

Characteristics of Each Cluster

There was no significant overall effect of subgrouping on gender, education, and cognition (Tables 2 and 3). However, there was a significant difference in age between clusters ($P = .000$) (Table 2). Post hoc analysis (performed for ANOVA $P < .15$) showed that the participants in cluster 2 were younger compared with those in clusters 1 and 3 ($P = .000$ and $.043$, respectively). Besides, Kruskal–Wallis analysis showed that the participants in cluster 2 had greater performance of N-back ($P = .04$).

In schizophrenia patients, there were no significant effects of cluster on SANS total and SRSS total (Table 2). However, there were significant effects of cluster on duration of medication exposure (DoM) and delusion ($P = .041$ and $.049$, respectively) (Table 2). Post hoc analysis showed that patients in cluster 1 (the “morphologically impoverished” group) had lower score in digit symbol-WAIS ($P = .046$), longer duration of illness (DoI), and DoM ($P = .044$ and $.023$, respectively), but lower scores in SAPS total and delusion ($P = .033$ and $.016$, respectively) (Table 2) compared with those in cluster 2 (the “morphologically intact” group). In addition, patients in cluster 3 (the “intermediate” group) had older onset age compared with those in cluster 2 (the “morphologically intact” group).

The demographic, cognitive, and clinical characteristics were summarized in Figure 2 (also see Supplementary Table 2 for cluster differences among the healthy controls). And the differences among the 3 clusters were observed in Age, N-back, DoI, DoM, Onset age, SAPS total, and Delusions. We also collected 5-year outcome data on positive symptom relapses and education/employment status using telephone interview and hospital chart review for 59 of 179 patients in this study. These results are presented in the Supplementary Figure 4.

Distance From Individuals to Cluster Central Point

Squared Euclidean distance was used to express distance from individual to cluster central point. According to the solution of K-means clustering, there were three central points (CP) corresponding to three clusters (Supplementary Figure 5). Significant differences were

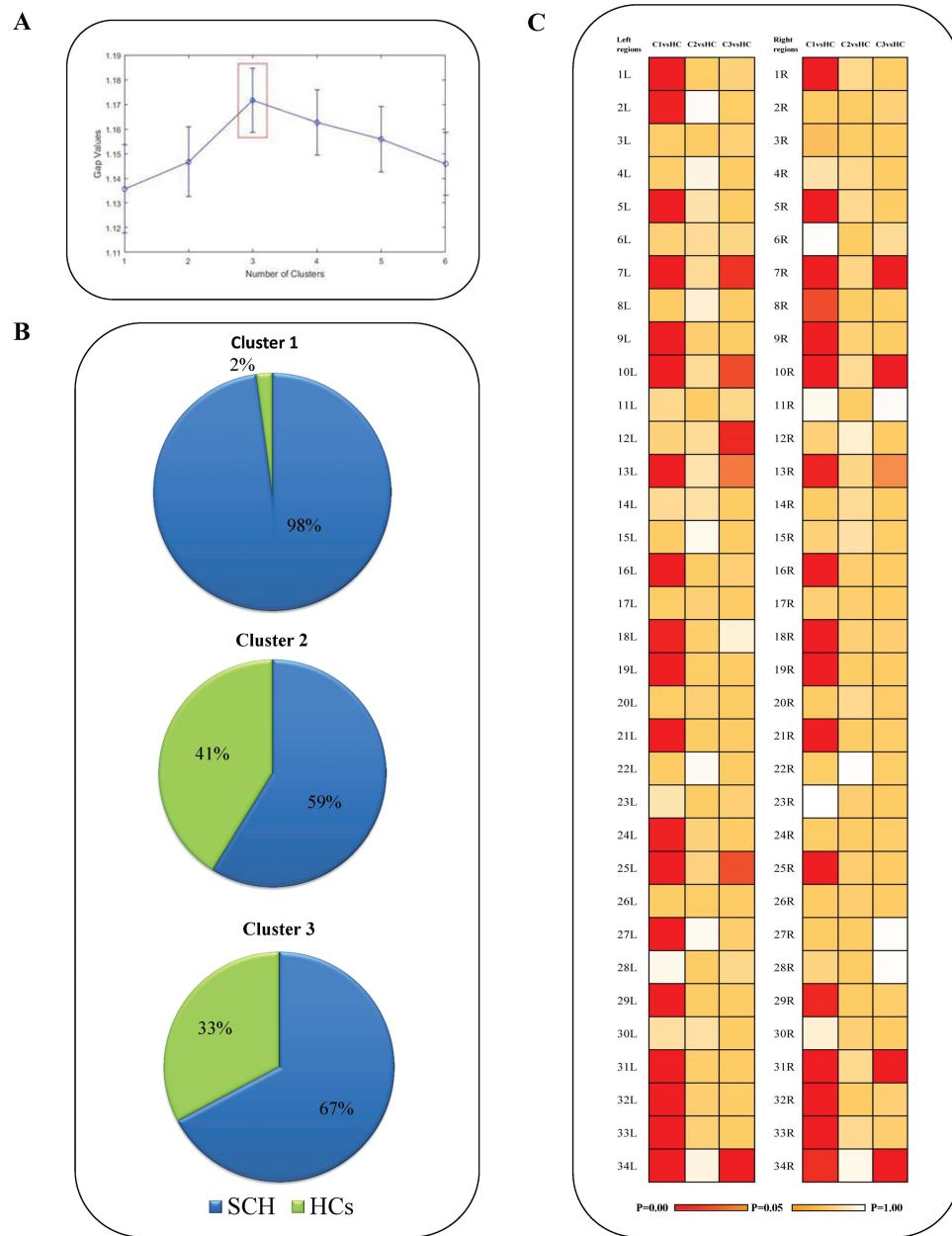


Figure 1. (A) Gap statistic to measure the number of optimum cluster in the data set using K-means clustering. The optimal solution for the morphological data from both patients and controls is the presence of three clusters. (B) The composition of each cluster, with 98%, 59%, and 67% of each cluster being comprised of patients, is shown. (C) Different patterns of cortical thinning in three clusters of patients (C1, C2, and C3). The age- and gender-adjusted differences between patients in each cluster and the total sample of HCs are shown by the coloured cells (with red indicating Bonferroni-adjusted $P < .05$). The name of the corresponding regions from the Desikan–Killiany atlas is shown in [Supplementary Table 3](#).

observed in distance to the CP of each cluster. Distance from schizophrenia patients to CP1 was significantly ($P = .03$) lower than the distance from HCs to CP1; distance from schizophrenia patients to CP2 and CP3 was significantly higher than the distance from HCs to CP2 and CP3 (Figure 3A). Thus, CP1 subgroup more closely reflected the clinical description of schizophrenia while CP2 and CP3 were closer to HCs in their profile. Besides, the correlation between the characteristics of clusters and

distance to CP of all samples is presented in Figure 3B. See [Supplementary Figure 2](#) for the stability of cluster membership to sample size balance, age, and sex.

Discussion

To the best of our knowledge, this is the first study using a data-driven approach to reveal subgroups in relatively early stage of schizophrenia (77.6% patients having

Table 2. Characteristics of Each Cluster

N(HCs/SCH)	Cluster 1		Cluster 2		Cluster 3		F/ χ^2 value	P value (uncorrected)	Post Hoc			
	46(1/45)		85(35/50)		125(41/84)							
	Mean	SD	Mean	SD	Mean	SD						
Age	24.02	6.11	21.95	4.30	25.13	6.03	8.31	.00032*	1>2*, 2<3**			
Gender	1.39	0.49	1.35	0.48	1.41	0.49	0.73	.695	—			
Education	12.13	2.71	12.34	2.49	12.33	2.71	0.12	.889	—			
Information-WAIS	16.40	6.17	18.00	5.88	17.34	5.21	1.069	.345				
Digit symbol-WAIS	64.50	12.74	72.08	19.32	71.10	21.60	2.252	.108	1<2*			
DoI (months)	32.6	40.1	18.6	25.4	25.5	31.5	2.186	.115	1>2*			
DoM (days)	328	665	98	194	186	389	3.263	.041*	1>2*			
Onset_age	21.6	5.2	20.3	5.1	22.4	5.7	2.242	.109	2<3*			
SAPS total	17.69	15.54	24.56	16.96	19.84	14.57	2.501	.085	1<2*			
Hallucinations	1.24	1.65	1.22	1.54	1.34	1.62	0.108	.898	—			
Delusions	1.73	1.59	2.53	1.67	2.02	1.55	3.069	.049*	1<2*			
Bizarre Behavior	0.98	1.35	1.10	1.26	1.17	1.31	0.306	.737	—			
Positive FTD	0.73	1.13	0.91	1.28	0.82	1.17	0.281	.755				
SANS total	29.58	17.78	31.65	24.18	36.76	27.04	1.249	.289				
Affective Flattening	1.29	1.41	1.33	1.36	1.75	1.45	2.171	.117				
Alogia	1.09	1.35	1.43	1.32	1.52	1.45	1.462	.235				
Avolition-Apathy	1.62	1.54	1.90	1.46	2.00	1.58	0.888	.413				
Anhedonia-Asociality	1.82	1.54	2.10	1.56	2.28	1.47	1.380	.254				
Attention	1.16	1.52	1.37	1.30	1.50	1.44	0.855	.427				
SSRS total	13.58	9.54	15.28	9.65	13.53	7.56	0.633	.532	—			

Note: Clinical ratings were administered only to participants with schizophrenia diagnose.

SCH, schizophrenia patients; HCs, healthy controls; DoI, duration of illness; DoM, duration of medication; FTD, formal thought disorder.

After Bonferroni correction, the significant difference level was 2.63e-4.

* $P < .05$; ** $P < 2.63e-4$.

Table 3. Cognitive Comparison Between Clusters

	Cluster 1		Cluster 2		Cluster 3		Kruskal-Wallis χ^2 value	P value (uncorrected)	Post Hoc
	Mean	SD	Mean	SD	Mean	SD			
N-back component	-0.11	0.84	0.21	0.86	-0.07	1.13	4.01	.14	1<2*
Contour task									
Component 1	-0.25	1.08	0.07	1.03	0.114	0.91	5.69	.58	—
Component 2	0.15	0.92	-0.01	0.89	-0.09	1.11	2.46	.29	—
Verbal Fluency									
Component 1	0.06	0.97	-0.04	0.98	-0.03	1.04	0.28	.87	—
Component 2	-0.04	1.07	0.33	1.17	-0.14	0.81	2.03	.36	—

Note: After Bonferroni correction, the significant difference level was 0.005.

* $P < .05$; ** $P < .005$.

<3 years of illness) based on cortical thickness. Applying data-driven clustering to a combined sample of HCs and patients, we identified three subgroups of schizophrenia with distinct patterns of cortical thinning. One subgroup was homogeneously comprised of patients (except for one HC) with widespread reduction in regional cortical thickness. A substantial proportion (33% and 41%) of the other two subgroups comprised of HCs, with patients in cluster 2 exhibiting a highly preserved thickness profile, whereas cluster 3 being intermediate (as shown in Figure 1). This result suggests that anatomical heterogeneity is

not solely an inherent disease feature, but rather representative of variation that can exist in HCs as well. In fact, only 25% of all patients were from the morphologically impoverished group, whereas the rest had cortical thickness features that were shared with healthy controls.

We observe that patients in the morphologically intact subgroup (cluster 2) are more symptomatic (SAPS total, delusions) with a shorter duration of illness and intact cognition (n-back) than the impoverished group (cluster 1), though there were no differences in sex, negative symptom burden, verbal

A

	Age	Onset_age	DoI	DoM	Delusion	SAPS_total	Digit_symbol	N-back
Post-hoc Comparison	1>2*	2<3*	1>2*	1>2*	1<2*	1<2*	1<2*	1<2*
between clusters	2<3*							

B

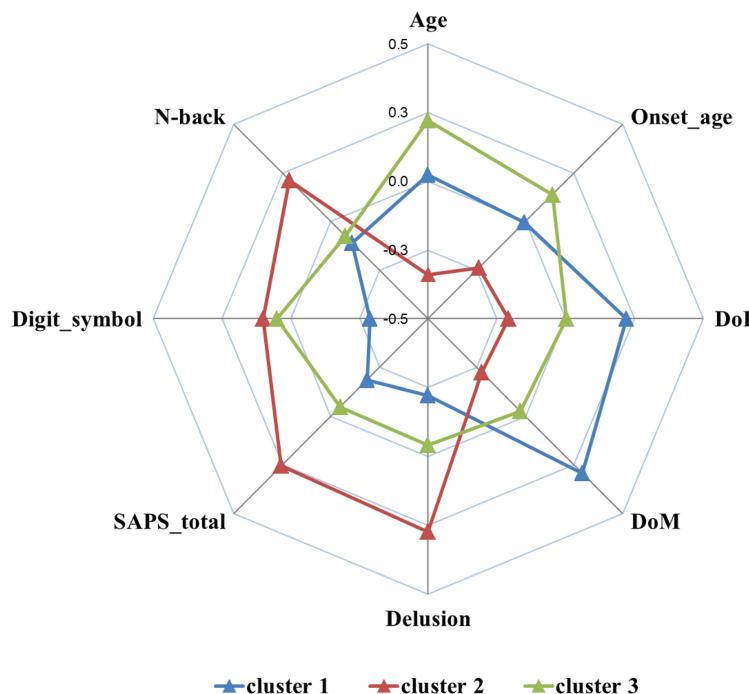
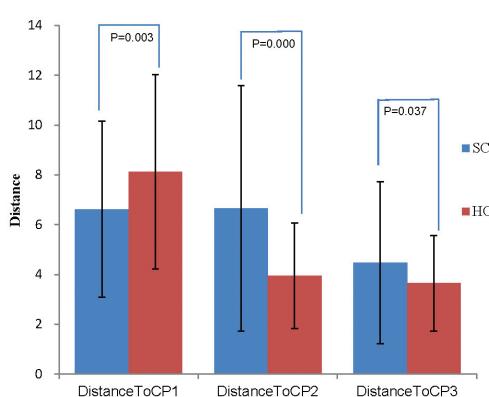


Figure 2. The demographic, cognitive, and clinical characteristics of the three morphological subgroups. (A) Post hoc comparison between clusters in phenotypes. * represents uncorrected $P < .05$; (B) The Y-axis represents the Z-scores of each factor (N-back-axis was the results of PCA). Cluster 1, “morphologically impoverished subgroup,” exhibited older age, lower digit symbol score, worse working memory, and longer DoT and DoM; Cluster 2, “morphologically intact subgroup,” exhibited younger age, higher delusion, and severity of positive symptoms; Cluster 3, “intermediate subgroup,” exhibited older age and onset age.

A



B

Factors	Distance To CP1		Distance To CP2		Distance To CP3	
	Pearson Correlation	P value uncorrected	Pearson Correlation	P value uncorrected	Pearson Correlation	P value uncorrected
Age	-0.305	0.000**	0.030	0.637	-0.219	0.001**
Onset age	-0.275	0.000**	-0.003	0.965	-0.205	0.006**
Duration of Medicine	-0.111	0.145	0.157	0.039*	0.072	0.347
Duration of Illness	-0.184	0.014**	0.075	0.319	-0.047	0.537
Delusion (SAPS20)	0.139	0.065	-0.163	0.030*	-0.060	0.427
Lack intimacy (SANS19)	-0.049	0.516	0.035	0.646	-0.019	0.799
Suicide plan (SSRS21)	-0.056	0.476	-0.011	0.887	-0.038	0.633
N-back main factor	-0.025	0.778	-0.105	0.238	-0.111	0.209

Figure 3. Distance from individuals to cluster central point. (A) presents the difference of distance between SCH (blue) and HCs (red). (B) presents correlations between characteristics of clusters and distance to CP. For diagnosis, 1 = SCH, 2 = HCs. Note: Clinical ratings were available for patients only. After Bonferroni correction, the significant difference level was $P < .0017$. * represent $P < .05$; ** means $P < .0017$.

fluency, and contour recognition. The subgroup with maximal thickness reduction (cluster 1) had the most pronounced cognitive deficits while the subgroup with the least thickness changes (cluster 2) had higher positive symptom burden (especially delusions) and somewhat higher frequency of positive symptom relapses. While initially counterintuitive, these results are indeed consistent with Crow's original dichotomy of a morphologically preserved type-1 schizophrenia with more positive symptom burden and a more chronic, cognitively impaired and structurally altered type-2 schizophrenia with less positive symptoms.²⁵ Furthermore, our results support various studies that reject the notion that positive symptoms per se are neurotoxic (ie, presence of delusions/hallucinations will adversely affect the brain anatomy).^{26,27} Several recent longitudinal studies support the possibility of a cortical reorganization or repair process that ameliorates morphological deficits occurring after the onset of psychosis (see ref. ²⁸ for a review). Our observation suggests that it is likely that such reorganization processes, if present, are more likely in patients with higher degree of positive symptom burden, but lower degree of cognitive impairment. The preservation of cognitive function in the morphologically intact subgroup is consistent with the well-replicated association between cognitive impairment and morphological deficits in schizophrenia.^{29,30}

Although we report three morphologically distinguishable subgroups of schizophrenia based on normative modeling that exploits the variations in healthy brain structure, it is important to note that this does not imply that only three morphological subtypes of schizophrenia exist. Prior studies have identified two¹¹ to six subgroups.⁸ The exact numbers reported vary according to sampling and methodological differences (termed as apparent heterogeneity by Schnack¹²). We recruited medicated patients in a relatively early stage of schizophrenia, all of same ethnicity (Han Chinese), limiting generalizability to more chronic samples from other parts of the world. We also chose to use *k*-means clustering instead of fuzzy clustering, so cluster membership (and clinical distinctions) of individual subjects can be meaningfully interpreted, though the discrete classes thus generated may have less information than fuzzy solutions. We also did not seek a specific number of clusters, and remained agnostic to the number of subgroups. The 3-cluster solution was found optimal based on the data-based gap statistic, which outperforms other cluster solution methods,²⁴ and has the specific advantage of working in combination with an adaptive version of *K*-means clustering in finding elongated clusters.^{24,31}

When studying the effect of clinical phenotype on the cluster membership (distance from cluster centroids), we note that age has a distinct gradient in the most impoverished and intermediate subgroups, with older age indicating more pronounced cortical thinning. Such

a relationship was not seen in the morphologically intact cluster 2, wherein the duration of exposure to antipsychotics was the most influential factor in deviation from the centroid. Furthermore, the cluster solution was stable even when adolescent subjects were excluded from the sample (Supplementary Figure 3 and Supplementary Table 1), indicating that the subgroups may be stable irrespective of the age range of the sample studied. The most morphologically impoverished cluster 1 had highest duration of illness as well as medication exposure, but these factors did not relate to the strength of an individual's cluster membership. This is consistent with the suspected "detrimental" effects of antipsychotics on brain morphology,³² as noted in other clustering studies^{4,8} but suggests that the medication-related variations are likely to be in line with the variability seen in healthy controls.

Limitations

The current study contained several limitations that should be considered. First, despite being agnostic with respect to the diagnostic differences in cortical thickness, our clinical recruitment was based on established clinical criteria for schizophrenia and did not include a broader spectrum of psychosis. A large number of prior observations have indicated that the structural pattern in other psychotic disorders is not qualitatively different but appears to be intermediate between schizophrenia and healthy controls. Secondly, we lacked longitudinal data to confirm the stability of observed clusters. Mechanistic heterogeneity at the individual level may be present across time (ie, different pathways acting at different time points, producing the same phenotype for the individual).³³ Given the cross-sectional nature of most clustering studies to date, the question of stability in cluster solutions remains unknown to date. Third, we were not able to untangle the association between antipsychotic exposure, age of onset, and illness duration, as we lacked a nonmedicated sample of patients. Although antipsychotic confounds are absent in untreated samples, cognitive and clinical symptoms are often unstable in acute stages of psychosis. Nevertheless, caution must be exercised in interpreting medication effects reported here.

We conclude that cortical thickness patterns in a large number of patients with schizophrenia (~75% in this sample) are not deviant but show variations parallel to healthy controls. This raises the interesting question of partitioning the anatomical heterogeneity in schizophrenia to a component of likely pathological perturbation and a component resulting from normative variations in healthy morphology. Given the challenges in reproducing case-control differences, a stratified approach towards identifying distinct sources of variation may be critical in our pursuit of the etiological heterogeneity of schizophrenia. Furthermore, interventional studies that aim to demonstrate structural changes in

schizophrenia are best designed with the consideration of the relative prevalence of subgroups of patients with normal variations as opposed to disease-specific perturbations in morphology.

Supplementary Material

Supplementary material is available at <https://academic.oup.com/schizophreniabulletin/>.

Funding

The China Precision Medicine Initiative (grant number 2016YFC0906300); and the National Natural Science Foundation of China (grant numbers 81561168021, 81671335, 81701325, 81801353).

Conflict of Interest

The authors have declared that there are no conflicts of interest in relation to the subject of this study. L.P. reports personal fees from Otsuka Canada, SPMM Course Limited, UK, Canadian Psychiatric Association; book royalties from Oxford University Press; investigator-initiated educational grants from Janssen Canada, Sunovion and Otsuka Canada outside the submitted work. All other authors report no relevant conflicts.

References

- Clementz BA; (on behalf of the B-SNIP principal investigators). Challenges Facing the Identification of Neurobiologically Distinct Psychosis Subtypes: Response to Neuhaus. *Am J Psychiatry*. 2016;173(8):838–839.
- Clementz BA, Sweeney JA, Hamm JP, et al. Identification of Distinct Psychosis Biotypes Using Brain-Based Biomarkers. *Am J Psychiatry*. 2016;173(4):373–384.
- Meda SA, Ruaño G, Windemuth A, et al. Multivariate analysis reveals genetic associations of the resting default mode network in psychotic bipolar disorder and schizophrenia. *Proc Natl Acad Sci U S A*. 2014;111(19):E2066–E2075.
- Cobia D, Csernansky JG, Ming J, Wang L. Pattern Classification Using Principal Components Analysis Of Cortical Thickness In Neuropsychologically Defined Schizophrenia Subtypes. *Schizophr Bull*. Mar 2011;37:161–161.
- Geisler D, Walton E, Naylor M, et al. Brain structure and function correlates of cognitive subtypes in schizophrenia. *Psychiatry Res*. 2015;234(1):74–83.
- Nenadic I, Sauer H, Gaser C. Distinct pattern of brain structural deficits in subsyndromes of schizophrenia delineated by psychopathology. *Neuroimage*. 2010;49(2):1153–1160.
- Nenadic I, Yotter RA, Sauer H, Gaser C. Patterns of cortical thinning in different subgroups of schizophrenia. *Br J Psychiatry*. 2015;206(6):479–483.
- Sugihara G, Oishi N, Son S, Kubota M, Takahashi H, Murai T. Distinct patterns of cerebral cortical thinning in schizophrenia: a neuroimaging data-driven approach. *Schizophr Bull*. 2017;43(4):900–906.
- Ajnakina O, Lally J, Di Forti M, et al. Patterns of illness and care over the 5 years following onset of psychosis in different ethnic groups; the GAP-5 study. *Soc Psychiatry Psychiatr Epidemiol*. 2017;52(9):1101–1111.
- Dollfus S, Everitt B, Ribeyre JM, Assouly-Besse F, Sharp C, Petit M. Identifying subtypes of schizophrenia by cluster analyses. *Schizophr Bull*. 1996;22(3):545–555.
- Dwyer DB, Cabral C, Kambeitz-Ilankovic L, et al. Brain subtyping enhances the neuroanatomical discrimination of schizophrenia. *Schizophr Bull*. 2018;44(5):1060–1069.
- Schnack HG. Improving individual predictions: Machine learning approaches for detecting and attacking heterogeneity in schizophrenia (and other psychiatric diseases). *Schizophr Res*. 2017;1–9. doi:10.1016/j.schres.2017.10.023.
- Fischl B, Dale AM. Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proc Natl Acad Sci U S A*. 2000;97(20):11050–11055.
- First MB, Spitzer RL, Gibbon M, Williams JBW. *Structured Clinical Interview for DSM-IV Axis I Disorders - Patient Edition*. New York: Biometrics Research Department, New York State Psychiatric Institute; 1997.
- Phillips MR, Xiong W, Wang RW, Gao YH, Wang XQ, Zhang NP. Reliability and validity of the Chinese versions of the Scales for Assessment of Positive and Negative Symptoms. *Acta Psychiatr Scand*. 1991;84(4):364–370.
- Taiminen T, Huttunen J, Heilä H, et al. The Schizophrenia Suicide Risk Scale (SSRS): development and initial validation. *Schizophr Res*. 2001;47(2-3):199–213.
- Zhang H, Wei X, Tao H, et al. Opposite effective connectivity in the posterior cingulate and medial prefrontal cortex between first-episode schizophrenic patients with suicide risk and healthy controls. *PLoS One*. 2013;8(5):e63477.
- Wu G, Wang Y, Mwansisya TE, et al. Effective connectivity of the posterior cingulate and medial prefrontal cortices relates to working memory impairment in schizophrenic and bipolar patients. *Schizophr Res*. 2014;158(1-3):85–90.
- Zhou L, Pu W, Wang J, et al. Inefficient DMN Suppression in Schizophrenia Patients with Impaired Cognitive Function but not Patients with Preserved Cognitive Function. *Sci Rep*. 2016;6:21657.
- Field DJ, Hayes A, Hess RF. Contour integration by the human visual system: evidence for a local “association field”. *Vision Res*. 1993;33(2):173–193.
- Kovács I, Julesz B. A closed curve is much more than an incomplete one: effect of closure in figure-ground segmentation. *Proc Natl Acad Sci U S A*. 1993;90(16):7495–7497.
- Chandna A, Pennefather PM, Kovács I, Norcia AM. Contour integration deficits in anisotropic amblyopia. *Invest Ophthalmol Vis Sci*. 2001;42(3):875–878.
- Desikan RS, Ségonne F, Fischl B, et al. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage*. 2006;31(3):968–980.
- Tibshirani R, Walther G, Hastie T. Estimating the number of clusters in a data set via the gap statistic. *J Roy Stat Soc B* 2001;63:411–423.
- Crow TJ. The two-syndrome concept: origins and current status. *Schizophr Bull*. 1985;11(3):471–486.
- Gould IC, Shepherd AM, Laurens KR, Cairns MJ, Carr VJ, Green MJ. Multivariate neuroanatomical classification of cognitive subtypes in schizophrenia: a support vector machine learning approach. *Neuroimage Clin*. 2014;6:229–236.

27. Li M, Li X, Das TK, et al. Prognostic utility of multivariate morphometry in schizophrenia. *Front Psychiatry*. 2019;10:245.
28. Palaniyappan L. Progressive cortical reorganisation: a framework for investigating structural changes in schizophrenia. *Neurosci Biobehav Rev*. 2017;79:1–13.
29. Cobia DJ, Csernansky JG, Wang L. Cortical thickness in neuropsychologically near-normal schizophrenia. *Schizophr Res*. 2011;133(1-3):68–76.
30. Van Rheenen TE, Cropley V, Zalesky A, et al. Widespread volumetric reductions in schizophrenia and schizoaffective patients displaying compromised cognitive abilities. *Schizophr Bull*. 2018;44(3):560–574.
31. El-Mandouh AM, Mahmoud HA, Abd-Elmegid LA, Haggag MH. Optimized K-means clustering model based on gap statistic. *Int J Adv Comput Sc*. Jan 2019;10(1):183–188.
32. 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(2):128–137.
33. Feczko E, Miranda-Dominguez O, Marr M, Graham AM, Nigg JT, Fair DA. The heterogeneity problem: approaches to identify psychiatric subtypes. *Trends Cogn Sci*. May 29 2019;23(7):584–601.