Two Patterns of White Matter Abnormalities in Medication-Naive Patients With First-Episode Schizophrenia Revealed by Diffusion Tensor Imaging and Cluster Analysis

Huaqiang Sun, PhD; Su Lui, MD, PhD; Li Yao, MS; Wei Deng, MD; Yuan Xiao, MS; Wenjing Zhang, MS; Xiaofei Huang, MD, PhD; Junmei Hu, MD; Feng Bi, MD, PhD; Tao Li, MD, PhD; John A. Sweeney, PhD; Qiyong Gong, MD, PhD

IMPORATNCE  Accumulating evidence supports the hypothesis that cerebral white matter abnormalities are involved in the pathophysiology of schizophrenia; however, findings from in vivo neuroimaging studies have been inconsistent. Besides confounding factors, including age, illness duration, and medication effects, an additional cause for the inconsistent results may be heterogeneity in the nature of white matter alterations associated with the disorder.

OBJECTIVE  To investigate whether different patterns of white matter abnormalities exist in a large cohort of medication-naive patients with first-episode schizophrenia and the relationship between such patterns and clinical parameters.

DESIGN, SETTING, AND PARTICIPANTS  A cross-sectional diffusion tensor imaging study of 113 medication-naive patients with first-episode schizophrenia and 110 demographically matched healthy control individuals. The study was conducted in the mental health center of West China Hospital, Sichuan University, Chengdu, China, from January 2006 to June 2014.

MAIN OUTCOMES AND MEASURES  The patterns of white matter abnormalities revealed by tract-specific analysis in conjunction with hierarchical clustering.

RESULTS  With diffusion features extracted from 18 fiber tracts, cluster analysis revealed 2 patterns of abnormalities. One pattern (42.5% of patient sample) showed widespread white matter abnormalities compared with matched healthy control individuals, while another pattern (57.5% of patient sample) only showed circumscribed regional white matter abnormalities, mainly in the left superior longitudinal fasciculus. Patients in these subgroups did not differ in demographic features; however, negative symptoms were more severe in patients with widespread white matter abnormalities.

CONCLUSIONS AND RELEVANCE  Two distinct patterns of white matter abnormalities exist at the early phase of schizophrenia, with those having global abnormalities experiencing more severe negative symptoms. The finding that distinct subgroups of patients with schizophrenia have different forms of white matter pathology may reflect qualitatively distinct genetic influences or neurodevelopmental alterations and thus represents a promising strategy for resolving neurobiological heterogeneity in the schizophrenia syndrome.
Schizophrenia has long been hypothesized as a syndrome involving insufficient or ineffective communication between brain regions. As cerebral white matter comprises densely packed, mostly myelinated axons that interconnect gray matter regions, pathology in white matter has been proposed to be a contributing factor in the pathophysiology of schizophrenia. This hypothesis has been supported by postmortem and genetic studies, which have demonstrated abnormalities in myelination, oligodendrocytes, and axons in patients with schizophrenia.

Using diffusion tensor imaging (DTI), magnetic resonance imaging studies targeting regions of interest (ROIs) and using voxel-based morphometry and tract-based spatial statistics have reported decreased fractional anisotropy (FA), a putative index that reflects white matter integrity, in patients with schizophrenia. However, the location and range of deficits varied considerably across studies. For instance, most ROI-based studies found FA reduction in the uncinate, cingulum, fornix, corpus callosum, and inferior longitudinal fasciculus (ILF), whereas negative results in the targeted ROIs have also been reported. Results from voxel-based morphometry studies have likewise shown different regional patterns of FA reductions in patients with schizophrenia. Thus far, tract-based spatial statistic results have shown localized FA reduction within the superior longitudinal fasciculus (SLF) or widespread white matter abnormalities across the whole brain in first-episode patients, and in patients with chronic disease, deficits have been seen in the corpus callosum, ILF, inferior fronto-occipital fasciculus (IOF), and SLF.

It is widely accepted that schizophrenia is a heterogeneous illness clinically, neurobiologically and genetically. Thus, a possible explanation for the inconsistent findings is the existence of different underlying pathophysiologies that lead to different patterns of white matter abnormalities among subtypes of patients with schizophrenia. The potential for such heterogeneity is suggested by some early studies that have reported discrete or more generalized white matter pathology. For example, Kittis et al reported FA reduction in the left uncinate fasciculus and Rowland et al found lower FA localized to the right superior longitudinal fasciculus and middle frontal white matter in deficit syndrome schizophrenia, while Voineskos et al reported more diffuse alterations in patients with deficit schizophrenia. Symptom features have also been linked to white matter pathology as Guo et al reported right lateralized white matter abnormalities only in paranoid schizophrenia. However, symptom-based subtyping strategies have been criticized for their instability over time and the inability to ensure that patients in any clinical subtype share similar neural system pathology. Thus, an alternative approach resolving schizophrenia heterogeneity using neurobiological parameters, such as diffusion properties of white matter tracts, may represent an advantageous strategy for identifying subgroups or subtypes of the illness.

In addition, minimizing the influence of potential confounding factors is particularly important in addressing questions of biological heterogeneity. Most previous DTI studies included chronic and medicated patients, and it is conceivable that long-term antipsychotic medication or secondary factors related to course of illness effects may contribute to abnormalities observed in DTI. Thus, studying medication-naive first-episode patients is particularly attractive for investigating heterogeneity in the clinical neurobiology of schizophrenia.

The current study aimed to characterize diffusion properties of major white matter tracts in a relatively large sample of patients with medication-naive first-episode schizophrenia relative to matched healthy control individuals in order to use cluster analysis to identify subgroups of patients based on the pattern of white matter alterations in diffusion properties across fiber tracts and to identify clinical differences between identified subgroups of patients.

Methods

Participants
In this study conducted from January 2006 to June 2014, participants included 113 right-handed patients with first-episode schizophrenia (56 men; 57 women; mean age, 23.8 years; age range, 16-46) and 110 matched healthy control individuals (58 men; 52 women; mean age, 23.4; age range, 18-41) (see the eAppendix in the Supplement for a detailed description of the demographics). Diagnosis of patients was determined using the Structured Clinical Interview for DSM-IV, and clinical symptoms were assessed using the Positive and Negative Syndrome Scale (PANSS).

This study was approved by the ethics committee of Sichuan University, Chengdu, Sichuan, China; written informed consent was obtained from each participant.

Magnetic Resonance Imaging Acquisition
All magnetic resonance imaging scans were performed using a GE Signa EXCITE 3.0-T scanner (GE Healthcare). Diffusion tensor imaging (20 diffusion-encoding directions) and high-resolution T1 data were acquired from each participant (see the eAppendix in the Supplement for detailed scanning parameters).

DTI Processing and Automatic Tracts Identification
Whole-data processing steps are illustrated in eFigure 1 in the Supplement (see the eAppendix in the Supplement for a detailed description). For routine DTI processing, head motion, and eddy current correction, brain extraction and tensor model fitting were all performed using FSL software (FMRIB Software Library, FMRIB).

We used automated fiber quantification software to identify 18 white matter tracts in each participant’s brain. The identification procedure included 3 primary steps: (1) whole-brain deterministic fiber tractography, (2) waypoint ROI-based tract segmentation, and (3) probability map-based fiber refinement (see the eAppendix in the Supplement for a detailed description of the automated fiber quantification processing). The identified 18 tracts were bilateral anterior thalamic radiation, corticospinal tract (CST), cingulum cingula...
late, cingulum hippocampus, IFOF, ILF, SLF, uncinate, genu, and splenium of the corpus callosum (eFigure 2 in the Supplement).

Feature Extraction, Cluster Analysis, and Cluster Validation
After tract identification, the diffusion measurement along the tract core, defined as the tract profile, was extracted from each fiber tract. In addition to FA, the tract profile of mean diffusivity (MD), a summative measure that describes average total diffusivity in a given voxel, was also evaluated. The resulting tract profiles were visually inspected to exclude patients with obvious calculation error in fiber reconstruction or identification and then smoothed using 10-point moving average to reduce local dramatic variation caused by imaging noise. The integration of FA and MD values along the full length of tract profiles were defined as the features of a tract (Figure 1). In this way, each tract had 2 features and each participant would have 36 features to depict their global white matter status (see the eAppendix in the Supplement for a detailed description about feature extraction).

Agglomerative hierarchical clustering was performed on patients with schizophrenia using the 36 white matter features extracted above. The optimal cluster number was determined using Silhouette, Dunn, and connectivity indices, which reflect the compactness, separation, and connectedness of the generated clusters. The stability of the cluster solution was also tested using a subsampling technique (see the eAppendix in the Supplement for a detailed description of the cluster analysis and cluster validation).

Statistical Analysis
Demographic characteristics, including age, sex distribution, and years of education, were compared between healthy control individuals and patient subgroups identified by the cluster analysis using 1-way analysis of variance or χ² test. Clinical symptom scores (PANSS positive and negative scores) and illness duration were similarly compared between patient subgroups using 2-tailed t test (2 subgroups) or analysis of variance (>2 subgroups). All statistical test results were considered significant if corrected P values were less than .05.

Fractional anisotropy and MD profiles were compared between healthy control individuals and patient subgroups in a pointwise manner. The tract profiles from a participant were arranged in a single matrix. All these matrices were fed into permutation-based statistical analysis with 10,000 permutations using the FSL Randomize program, with age and illness duration as covariates. The statistical results were subject to familywise error correction for multiple comparisons following threshold-free cluster enhancement and thresholded at P < .05 (see eFigure 3 in the Supplement for the flowchart of pointwise comparison).

Additional exploratory analysis of the correlation between diffusion properties and PANSS positive/negative symptoms scores were performed within whole patient group and within each subgroup using partial Pearson correlation, with age as a covariate.

Results

Hierarchical Clustering
The result of hierarchical clustering is shown as a combination of dendrogram and heat map illustration in Figure 2. According to the definition of used validating indices, the maximum is preferred for both the Silhouette and Dunn indices to determine the optimal number of clusters, while the minimum is preferred for the connectivity index. Our cluster result achieved the maximum for the Silhouette index (0.51) and Dunn index (19.25) and the minimum for connectivity (2.93) when the cluster number was equal to 2 (eFigure 4 in the Supplement). The stability test also suggested the 2-cluster solution is most stable (eFigure 5 in the Supplement). Thus, the optimal cluster number that best

---

Copyright 2015 American Medical Association. All rights reserved.
represents the data structure was determined to 2. Subsequent analysis mainly focused on these 2 subgroups (subgroup 1 and subgroup 2 indicated in Figure 2). Forty-eight patients (42.5%) were placed in subgroup 1 and 65 patients (57.5%) were placed in subgroup 2. From visual inspection of the dendrogram, the patients in subgroup 1 showed relative...
Abbreviation: NA, not applicable.

of the right SLF (lumhippocampus and right ILF, and the frontallobe portion of the left thalamic radiation, the inferior portion of the posterior portion of the SLF and othersmall segments in the posterior portion of the bilateral cingulum hippocampus, bilateral ILF, bilateral SLF, and genu of the corpus callosum, as well as in the posterior portion of the bilateral thalamic radiation, the inferior portion of the bilateral CST, occipital lobe portion of the splenium of the corpus callosum, both ends of the left cingulum cingulate and left uncinate, the temporal lobe portion of the right uncinate, and the intermediate component of the bilateral IFOF. Subgroup 2 showed regional FA reduction mainly in the left SLF and other small segments in the posterior portion of the left thalamic radiation, the inferior portion of the bilateral CST, the temporal lobe portion of the bilateral cingulum hippocampus and right ILF, and the frontal lobe portion of the right SLF (Figure 3).

Pointwise comparison of MD profiles between healthy control individuals and subgroup 1 showed extensive elevation in the entire bilateral cingulum hippocampus, bilateral ILF, bilateral IFOF, bilateral SLF, bilateral uncinate, genu of the corpus callosum, the left CST, both ends of the splenium of the corpus callosum and most of the bilateral thalamic radiation, bilateral cingulum cingulate, and the right CST. Subgroup 2 showed regional MD elevation in the inferior portion of the CST and posterior portion of the cingulum cingulate (Figure 4). A summary on average FA and MD of each identified fiber tract from healthy control individuals and patient subgroups was appended in the Supplement (eTables 1 and 2).

**Subgroup Differences in Tract Profile of FA and MD**

In pointwise comparison of FA profiles, subgroup 1 showed widespread FA reduction relative to healthy control individuals in the entire bilateral cingulum hippocampus, bilateral ILF, bilateral SLF, and genu of the corpus callosum, as well as in the posterior portion of the bilateral thalamic radiation, the inferior portion of the bilateral CST, occipital lobe portion of the splenium of the corpus callosum, both ends of the left cingulum cingulate and left uncinate, the temporal lobe portion of the right uncinate, and the intermediate component of the bilateral IFOF. Subgroup 2 showed regional FA reduction mainly in the left SLF and other small segments in the posterior portion of the left thalamic radiation, the inferior portion of the bilateral CST, the temporal lobe portion of the bilateral cingulum hippocampus and right ILF, and the frontal lobe portion of the right SLF (Figure 3).

In whole patient group partial correlation analysis between PANSS positive/negative and FA/MD of each fiber tract, we found moderate negative correlations between negative symptom score and the average FA of the following tracts: left anterior thalamic radiation ($r = -0.36; P < .001$), left CST ($r = -0.33; P = .005$), genu of the corpus callosum ($r = -0.42; P < .001$), left IFOF ($r = -0.41; P < .001$), right IFOF ($r = -0.32; P = .002$), left ILF ($r = -0.43; P < .001$), and right ILF ($r = -0.45; P < .001$). Positive correlations were also found between negative symptom score and average MD of the following fiber tracts: left anterior thalamic radiation ($r = 0.39; P < .001$), right anterior thalamic radiation ($r = 0.34; P = .001$), left CST ($r = 0.41; P < .001$), right CST ($r = 0.36; P = .001$), genu of the corpus callosum ($r = 0.32; P = .002$), left IFOF ($r = 0.36; P < .001$), left ILF ($r = 0.32; P = .002$), right ILF ($r = 0.32; P = .002$), right SLF ($r = 0.34; P = .001$), and left uncinate ($r = 0.34; P = .001$). No significant correlation was found between positive symptom score and average FA/MD value of any tract in the whole patient group. In addition, no significant correlation was found between diffusion properties (FA or MD) of fiber tracts and positive or negative symptom scores within either subgroup.

**Discussion**

Although most researchers agree that white matter abnormalities exist in schizophrenia, whether these abnormalities are localized to specific tracts or spread throughout the brain remains controversial. Issues including medication history, varying illness duration, and small sample size have been considered likely to cause these inconsistent findings. Another alternative is that there are neurobiologically distinct subgroups of patients with schizophrenia with discrete forms of white matter pathology. Using a neuroimaging data-driven method, the current study, supporting that second possibil-
ity, identified 2 distinct patterns of white matter abnormality in a large sample of medication-naive patients with first-episode schizophrenia. One of the patterns showed widespread FA reduction and MD elevation across the 18 examined fiber tracts, while another pattern only showed regional FA reduction primarily in the left SLF. The 2 subgroups did not differ in demographic features and illness duration, but the subgroup with extensive white matter pathology did have more severe negative symptoms. The current findings provide new evidence indicating the existence of 2 neurobiologically distinct subgroups of patients with schizophrenia, which may help explain the heterogeneity of findings in prior DTI studies of schizophrenia and also reflect qualitatively distinct genetic influences or neurodevelopmental alterations.

The physical properties of the fiber bundles, such as packing density, myelination, and axon diameter, are all known to influence DTI-derived parameters such as FA and MD.41 Thus, the widespread reduced FA and elevated MD in subgroup 1 may reflect diffused abnormalities of myelin or/and fiber axons in these patients. Two factors may account for such global deficits. From the neurodevelopment view, the widespread FA reduction observed in this subset of our samples could be interpreted as the consequence of impaired myelination caused by genetic factors as genetic studies of schizophrenia have reported dysregulation in genes related to myelin3,4 and oligodendrocytes.5 Another possible factor might be related to the proposed neuroinflammatory influence.42 Multiple lines of evidence suggest an inflammatory process may be involved in schizophrenia.43 Neuroinflammation in the early stage of schizophrenia might increase extracellular volume in the entire brain,44 then resulting in reduced anisotropic diffusivity of white matter.45 This possibility raises the need for future research to examine genetic and inflammatory factors in the subpopulation of patients identified in the present study.

In the subgroup with milder and more localized white matter alterations, the SLF, which is the major white matter connecting prefrontal and parietal cortices, was most affected. According to the macrocircuit theory, specific white matter tracts

Figure 3. The Pointwise Comparison of Fractional Anisotropy (FA) Profile Between Healthy Control Individuals and Patient Subgroups

The plots of FA profiles of 18 identified fiber tracts from healthy control individuals and patient subgroups (green for healthy control individuals, blue for subgroup 1, and red for subgroup 2) in mean (SD) (solid lines for means and shaded areas for SDs). The blue bars under the FA profile indicate the regions of significant difference between subgroup 1 and healthy control individuals; the red bars under the FA profile indicate the regions of significant difference between subgroup 2 and healthy control individuals. The x-axis represents the location between the beginning and termination waypoint regions of interest. L indicates left; R, right.
Figure 4. The Pointwise Comparison of Mean Diffusivity (MD) Profile Between Healthy Control Individuals and Patient Subgroups

The plots of MD profiles of 18 identified fiber tracts from healthy control individuals and patient subgroups (green for healthy control individuals, blue for subgroup 1, and red for subgroup 2) in mean (SD) (solid lines for means, shaded areas for SDs). The blue bar under the MD profile indicates the regions of significant difference between subgroup 1 and healthy control individuals; the red bar under the MD profile indicates the regions of significant difference between subgroup 2 and healthy control individuals. The x-axis represents the location between the beginning and termination waypoint regions of interest. L indicates left; R, right.

It was also noteworthy that patients with widespread and more prominent white matter abnormalities exhibited greater severity in negative symptoms than those with regional deficits. In fact, negative symptoms (ie, anhedonia, amotivation, and social dysfunction) were found to be related to specific white matter deficits in previous studies, notably with alterations in the corpus callosum,\textsuperscript{52} uncinate,\textsuperscript{53} SLF,\textsuperscript{54} and ILF.\textsuperscript{55} Consistent with previous reports, the current study also found moderate correlations between the severity of negative symptoms and diffusion measurements of certain tracts, including the anterior thalamic radiation, CST, genu of the corpus callosum, ILF, IFOF, SLF, and uncinate at the whole patient group level. However, the absence of significant correlation between severity in negative symptoms and diffusion measurements in the subgroup level suggests that the white matter deficit may not be the direct cause of negative symptoms; instead, they may both be downstream phenomena of the same upstream pathogenesis. This hypothesis is partially supported by neurochemical studies that have proven that proinflammatory cytokines can markedly induce anhedonic behavior and social impairments as neuroinflammation is a possible pathogenesis for schizophrenia.\textsuperscript{56}

From a broader perspective, our data-driven approach using measures of brain pathology rather than clinical phe-
nomenology to classify patients may represent a significant advance in the evolution of diagnostic practices in psychiatry to benefit from and use advances from clinical neuroscience. We recognize that our specific observations using DTI data for grouping patients on the basis of the form of their white matter pathology require validation and further exploration of their origins and relevance for the crucial clinical questions about identifying prodromal states and predicting treatment outcomes and illness course. Also, defining the number of subgroups in a patient population remains a statistical challenge, with different approaches yielding somewhat different solutions.5 Yet, our demonstration of the subtypes of patients classified by their neuroimaging measures in the present study is one step forward in the use of such measures to subgroup and eventually help diagnose illness and plan treatments for patients with psychotic disorders. This is consistent with the broad aims of the Research Domain Criteria project from the National Institute of Mental Health58 and the Bipolar and Schizophrenia Network for Intermediate Phenotypes consortium, which have collaborated.

Several issues should be considered with regard to interpreting the current findings. First, because we were limited by the number of diffusion directions acquired, the fiber tractography was based on a tensor model, which cannot solve the crossing-fiber problem in some voxels. Furthermore, to avoid anatomical variations from superficial regions, only the portions between 2 waypoint ROIs before tracts arose to the innervate cortex were used in the analyses. Measuring superficial areas using higher-resolution techniques available now, but that were less established when we began to recruit the study samples, may provide more useful information relevant to the macrocircuit hypothesis in future work.

Although the patient sample used in this study was relatively large and free from medication exposure and course of illness effects, it may be insufficient to fully capture the neurobiological heterogeneity of schizophrenia as seen in white matter tract disturbances. With a larger patient sample, smaller patient subgroups might be detected in cluster analyses or the subgroups we defined might be parsed into smaller groups. Thus, replication studies, especially with larger sample sizes to validate and perhaps improve on our identification of 2 subgroups, are needed to extend the findings from the current study.

Conclusions

Using a purely data-driven analysis scheme, the current study identified 2 subgroups of patients with schizophrenia defined by different patterns of white matter abnormalities. The group with more severe and widespread white matter pathology had more severe negative symptoms, suggesting a clinical relevance to the neurobiologically based subgrouping. The findings suggest that patterns of white matter abnormalities may provide a promising biomarker for subtyping patients with schizophrenia for studies of illness mechanism and as a quantitative phenotype for genetic research.

REFERENCES

1. Friston KJ. The disconnection hypothesis. Schizophr Res. 1998;30(2):115-125.
memory organization in schizophrenia. Schizophr Res. 2008;103(1-3):52-61.


