Neocortical Gray Matter Volume in First-Episode Schizophrenia and First-Episode Affective Psychosis: A Cross-Sectional and Longitudinal MRI Study

Motoaki Nakamura, Dean F. Salisbury, Yoshio Hirayasu, Sylvain Bouix, Kilian M. Pohl, Takeshi Yoshida, Min-Seong Koo, Martha E. Shenton, and Robert W. McCarley

Clinical Neuroscience Division (MN, DFS, SB, TY, M-SK, MES, RWM), Laboratory of Neuroscience, Veterans Affairs Boston Healthcare System, Brockton Division; Department of Psychiatry (MN, DFS, SB, TY, M-SK, MES, RWM), Harvard Medical School, Boston; Psychiatry Neuroimaging Laboratory (MN, SB, TY, M-SK), Brigham and Women's Hospital, Boston; Cognitive Neuroscience Laboratory (DFS), McLean Hospital, Belmont, Massachusetts; Department of Psychiatry (YH), Yokohama City University School of Medicine, Kanagawa, Japan; and Computer Science and Artificial Intelligence Laboratory (KMP), Massachusetts Institute of Technology, Boston, Massachusetts

Abstract

Background—Overall neocortical gray matter (NCGM) volume has not been studied in first-episode schizophrenia (FESZ) at first hospitalization or longitudinally to evaluate progression, nor has it been compared with first-episode affective psychosis (FEAFF).

Methods—Expectation-maximization/atlas-based magnetic resonance imaging (MRI) tissue segmentation into gray matter, white matter (WM), or cerebrospinal fluid (CSF) at first hospitalization of 29 FESZ and 34 FEAFF, plus 36 matched healthy control subjects (HC), and, longitudinally ∼1.5 years later, of 17 FESZ, 21 FEAFF, and 26 HC was done. Manual editing separated NCGM and its lobar parcellation, cerebral WM (CWM), lateral ventricles (LV), and sulcal CSF (SCSF).

Results—At first hospitalization, FESZ and FEAFF showed smaller NCGM volumes and larger SCSF and LV than HC. Longitudinally, FESZ showed NCGM volume reduction (−1.7%), localized to frontal (−2.4%) and temporal (−2.6%) regions, and enlargement of SCSF (7.2%) and LV (10.4%). Poorer outcome was associated with these LV and NCGM changes. FEAFF showed longitudinal NCGM volume increases (3.6%) associated with lithium or valproate administration but without clinical correlations and regional localization.

Conclusions—Longitudinal NCGM volume reduction and CSF component enlargement in FESZ are compatible with post-onset progression. Longitudinal NCGM volume increase in FEAFF may reflect neurotrophic effects of mood stabilizers.

Keywords

Antipsychotics; bipolar disorder; longitudinal volume change; mood stabilizer; neocortex; schizophrenia

Address reprint requests to Robert W. McCarley, M.D., Department of Psychiatry (116A), VA Boston Healthcare System, Brockton Division, Harvard Medical School, 940 Belmont Street, Brockton, MA 02301; robert_mccarley@hms.harvard.edu.

We have no conflicts of interest to disclose.

Supplementary material cited in this article is available online.
Magnetic resonance imaging (MRI) has provided quantitative evidence for structural abnormalities in psychiatric disorders, particularly psychosis (1–3). The accumulating findings in schizophrenia suggest that it may involve two pathophysiologic processes, one occurring early (neurodevelopmental) and, although controversial, one post-onset (1,4–7). A crucial issue is whether progressive deterioration occurs and, if so, whether it is specific to schizophrenic versus affective psychosis.

In evaluating progression and specificity, one must consider which brain regions to evaluate and the methodology to be used. In previous cross-sectional comparisons, the majority of both manually delineated region of interest (ROI) (3) and voxel-based morphometry (VBM) (8–11) studies indicate smaller gray matter (GM) volumes in frontal and temporal neocortex and in limbic regions. In affective psychosis, manual ROI (12) and VBM (13,14) studies indicate deficits in frontal neocortex and limbic regions, although findings are inconsistent (15,16). Automated tissue segmentation of whole brain revealed smaller overall volume of supratentorial GM, including subcortical GM, in schizophrenia (17–20).

In longitudinal studies, Gur et al. (21) reported total frontal lobar volume reduction over about 2.5 years in schizophrenia (first-episode and previously treated patients), and Mathalon et al. (22), in a nonprospective study, reported progressive volume reduction over 4.2 years in right frontal lobe and bilateral posterior superior temporal gyrus in chronic schizophrenia. Ho et al. (23), in a study of recent-onset schizophrenia that used spatial normalization, found sulcal cerebral spinal fluid (SCSF) enlargement and frontal lobe white matter (WM) but not GM volume reduction. Lieberman et al. (24), who also used spatial normalization with warping, reported that haloperidol-treated first-episode patients exhibited significant decreases in GM volume over 1 year, whereas olanzapine-treated patients did not. Cahn et al. (25) measured individual brain tissue classes in first-episode schizophrenia over 1 year and found progressive whole cerebral GM volume reduction (associated with antipsychotic dosage), and increased lateral ventricular volume. A VBM study on first-episode psychosis (26) found gray matter reduction over 2 years in frontotemporal regions and anterior cingulate gyrus in schizophrenia but only in the latter structure within the small bipolar group \((n = 8)\); no control group was available for the longitudinal study.

The strong evidence for neocortical gray matter (NCGM) pathology in schizophrenia (1,3) led us to focus on NCGM volume. Moreover, eliminating basal ganglia, thalamus, medial temporal regions, and all infratentorial structures, although they are also implicated in psychosis, has the key methodologic advantage of allowing a more accurate segmentation because the greater degree of WM interpenetrating the GM of these structures results in a different GM MR signal intensity (27). Also, basal ganglia volumes are affected by antipsychotics (28) and may thus bias whole tissue class measures.

An important issue in MRI studies is the influence of medication on cortical GM. Antipsychotic administration has been reported to be associated with changes in GM volumes in humans (24,25) and GM and WM reductions in nonhuman primates (29). Moreover, mood stabilizers such as lithium and valproate may increase GM volume (30,31). Global measures used in these studies and our study are likely to be more sensitive to such effects than single structure-based measures (gyrus or gyral subdivision) because these have smaller volumes and greater between-subject and intra- and interrater measurement variability; hence they have greater variability per unit volume.

To address these issues and whether structural abnormalities were specific to schizophrenia or similar in affective psychosis, brain tissue volumes were measured cross-sectionally and longitudinally. The initial MRI scan occurred at the first hospitalization for psychosis for first-episode schizophrenia (FESZ) and first-episode affective psychosis (FEAFF) and at protocol
entrance for healthy control subjects (HC). The second scan occurred on average 1.5 years later. Brain structures were segmented using a validated tissue segmentation algorithm (32, 33). Neocortical gray matter (NCGM), cerebral white matter (CWM), SCSF, and lateral ventricles (LV) were measured separately for each hemisphere. Furthermore, to evaluate regional contributions, NCGM was parcellated into frontal, temporal, and parieto-occipital lobes. To our knowledge, this is the first longitudinal tissue segmentation study introducing a direct comparison between FESZ and FEAFF psychotic groups and focusing on NCGM.

**Methods and Materials**

**Subjects**

Consistent with our previous studies (34,35), “first episode” was operationally defined as the first hospitalization for psychosis. Twenty-nine FESZ patients (25 paranoid, 1 disorganized, 1 catatonic, and 2 undifferentiated), 34 FEAFF patients (31 bipolar, 3 unipolar), and 36 HC were compared cross-sectionally. Inclusion criteria for patients and control subjects were age 18 to 45; IQ above 75; and a negative history of seizures, head trauma with loss of consciousness, neurologic disorder, and any lifetime history of alcohol or drug dependence or current abuse. Patients were recruited from inpatient units at McLean Hospital, Belmont, Massachusetts; using inpatients provides highly reliable data on course, symptoms, and medication during the stay. Patient diagnosis was based on the Structured Clinical Interview for DSM (SCID)—Patient Edition (36) and information from medical records and was confirmed at follow-up. Control subjects were recruited from the local community through newspaper advertisement and had no Axis I or II disorder by SCID-Non-Patient Edition (37) and SCID-II interviews (38); there was also no history of any Axis I disorder in their first-degree relatives per report. Seventeen FESZ, 21 FEAFF, and 26 HC subjects were longitudinally rescanned, on average, 1.5 years later.

Groups were matched for age, parental socioeconomic status (PSES) (39), handedness (40), and sex (see Tables 1 and 2). Patients had poorer socioeconomic status (SES) and less education than HC, consistent with the debilitating effects of psychosis. Exclusion of the single mildly left-handed FESZ patient (Edinburgh index = −.07) and single left-handed FEAFF patient (Edinburgh index = −.47) from the subsequent analyses did not change the cross-sectional and longitudinal laterality effects. Patient groups did not differ in antipsychotic medication dosage, total Brief Psychiatric Rating Scale (BPRS) score, or Global Assessment Scale (GAS) score, except for initial BPRS total score in the longitudinal sample, with FESZ relatively more symptomatic ($t_{36} = 5.09, p = .03$; Tables 1 and 2). There were no significant differences between subjects retained or lost to second scan follow-up except for worse GAS scores ($t_{30} = 2.12, p = .04$), but not BPRS scores, in the lost-to-follow-up FEAFF subjects. This study was approved by the McLean Hospital and Harvard Medical School Institutional Review Boards. All subjects gave written informed consent before participation.

**Medication**

Patient report and hospital records were used to assess prehospitalization and between-scan medication history. Daily chlorpromazine equivalent antipsychotic dosage (41) did not correlate with any volumes or volume changes.

**Cross-Sectional Study**—The median duration of antipsychotic medication (typical or atypical antipsychotics) before scanning was 3 weeks for FESZ (range 0–24 weeks) and 1 week for FEAFF (range 0–24 weeks; Table 1).

**Longitudinal Study**—To examine medication effects on longitudinal NCGM volume change, each patient group was evaluated according to medication history and compliance...
before the second MRI scan with respect to antipsychotics and mood stabilizers (MS; lithium, valproate, or both; Table 2). Many patients were on both kinds of medication. In this naturalistic sample, 4 FESZ and 2 FEAFF patients reported medication noncompliance for more than 3 months before the second scan. Patients in the antipsychotic-treated groups were on atypical antipsychotics except for 1 FESZ and 2 FEAFF patients, who received both typical and atypical antipsychotics. The MS group consisted of patients who received MS during the entire interscan period and included 41% of FESZ (7/17) and 72% of FEAFF (13/21).

MRI Processing

The MRI protocol used two pulse sequences on a 1.5-T MRI system (GE Medical Systems, Milwaukee, Wisconsin), as described elsewhere (42). Briefly, a three-dimensional Fourier-transformed spoiled-gradient-recalled (SPGR) acquisition sequence yielded a coronal series of contiguous 1.5-mm images (echo time [TE] = 5 msec, repetition time [TR] = 35 msec, repetition = 1, nutation angle = 45°, field of view [FOV] = 24 cm, acquisition matrix = 256 × 256 × 124, voxel dimension = .9375 × .9375 × 1.5 mm). Next, a double-echo spin-echo yielded 108 contiguous axial double-echo (proton-density-weighted and T2-weighted) slices, with 54 levels, throughout the brain (TE = 30 and 80 msec, TR = 3000 msec, FOV = 24 cm, an interleaved acquisition with 3-mm slice thickness, voxel dimensions = .9375 × .9375 × 3.0 mm).

First, the T2 information from the double-echo spin-echo axial slices was registered to the SPGR images (43). Next, an expectation-maximization (EM) segmentation technique (33) was used to segment the images into three major tissue classes; GM, WM, and CSF. The method simultaneously performs image segmentation and inhomogeneity correction (44), using both SPGR and T2-weighted MR information (33), as well as spatial priors (maps of probability of a tissue class at particular MR image locations) (45). Priors help distinguish ambiguous intensity patterns in nonbrain areas (e.g., skin, skull, and dura matter). The probability map for each tissue class was nonlinearly mapped before segmentation onto each individual's native space, thus ensuring no geometric distortion of voxel location or size in individual brain images, such as might occur with warping to a template, even with subsequent modulation by Jacobian determinants.

Region-of-Interest Definitions

The procedure for isolating neocortex is described elsewhere (46). Briefly, neocortical ROI delineation included all six-layered neocortex and excluded the major portion of nonneocortical cortex, including limbic cortical areas (with the exception of the pyriform cortex) and most of paralimbic cortex, with the exception of portions of cingulate, insula, and temporal pole (for anatomic description: ref. 47). We describe this ROI as NCGM because the included regions of non-six-layer cortex comprise less than 5% of the GM volume in the ROI. Consistent with exclusion of medial temporal gray matter structures, the LV ROI did not include the very small temporal horn portion.

In the longitudinal sample, NCGM was manually parcellated into three lobar ROI, frontal, temporal, and parieto-occipital (Figure 1). This parcellation mainly used sulcal boundaries because these are more faithful to brain anatomy than a purely geometric parcellation. The frontal lobe was separated from the parieto-occipital lobe by the central sulcus on the convexity, a boundary that is constant, easily identifiable, and traceable with little interindividual variation (48). The central sulcus was initially traced on the axial plane, and subsequently these trace lines were used on the coronal plane to separate frontal and parietal lobes. For frontal GM on the medial wall, the posterior terminus was the most posterior coronal slice containing corpus callosum. The frontal lobe was clearly separated from the temporal lobe by the Sylvian fissure and circular insular sulcus. The posterior temporal lobe terminus was geometrically defined as
the most posterior coronal slice where the fornix could be clearly seen along the lateral ventricles, as in our previous studies (34). Occasionally, especially in the right hemisphere, the Sylvian fissure steeply ascended posteriorly through the parieto-occipital region. In this case, the superior boundary separating the temporal lobe from the parieto-occipital lobe was defined as the most superior axial slice where Heschl's gyrus could be seen. The parieto-occipital lobe was automatically defined by its contiguous boundaries with the frontal and temporal lobes.

All manual processes were done by one of the authors (MN) who was blinded to subject group. To assess interrater reliability, three raters blinded to diagnosis evaluated five randomly selected cases for the nonneocortical exclusion region and LV separation from CSF (raters: MN, TY, and GB) and for the NCGM lobar parcellation (raters: MN, UK, and JA). All intraclass correlation coefficients were greater than .99. The extremely high reliability compared with totally manually drawn ROI is likely due to the simplicity of the present ROI boundaries and automatic tissue segmentation.

**Statistical Analysis**

One-way analyses of variance (ANOVAs) assessed group differences in age, SES, PSES, and interscan interval. A chi-square test assessed group differences in gender frequencies. Student t tests tested patient group differences in medication dosages and clinical scales.

For the initial cross-sectional volume comparison, group differences in brain tissue volumes were assessed using a three-factor repeated-measures ANOVA model on tissue volume, with diagnostic group as the between-subjects factor and tissue class and hemisphere as the within-subjects factors. Post hoc analyses used a two-factor ANOVA model for each tissue class, followed by pairwise independent-samples t test with a Bonferroni-corrected cutoff p value of \( \alpha = .0167 \) (two-tailed), which was .05 divided by 3 (number of groups). The intracranial contents (ICC) volume (cm\(^3\)) was derived from the EM atlas segmentation and included all GM, WM, and CSF volumes above the most inferior axial slice containing cerebellum. Relative (%) volume \((\text{absolute volume (cm}^3)/\text{ICC (cm}^3)) \times 100\) was used to control for individual head size, although groups did not differ significantly in ICC \((p = .14)\). All volume comparisons were based on relative volumes. In our dataset, both ICC and age violated the assumption of “homogeneity of regression” (49,50). Thus neither ICC nor age was covaried in the present cross-sectional group comparison. Instead groups were matched precisely on these variables.

For the longitudinal volume comparison, percent volume change was calculated with the following formula: \((\text{relative volume at second scan}) - (\text{relative volume at baseline scan})/ (\text{relative volume at baseline scan}) \times 100 \) (\%), to control for any group difference in initial tissue volumes. Interscan interval was strictly matched among the three groups \((p = .97)\). There was no consistent correlation within each group between tissue volume changes and the interscan interval. Furthermore, the groups differed in the relationship of interscan interval to percent volume changes, indicating that the interscan interval was not appropriate for analysis of covariance. A three-factor repeated-measures ANOVA model was applied to the percent volume changes, with diagnostic group as the between-subjects factor and tissue class and hemisphere as the within-subject factors. We applied the same post hoc tests on percent volume change as the cross-sectional volume comparison.

For NCGM lobar parcellation, the same statistical strategy was used as in the cross-sectional and longitudinal tissue volume comparisons described earlier except that lobes (three levels) were used instead of tissue classes (four levels).

To evaluate the magnitude of group differences, Cohen's \( d \) (51) is provided when pairwise comparisons did not attain significance.
Pearson’s correlations evaluated associations between brain tissue volume change and clinical outcome. Clinical outcome was evaluated by the percent change in BPRS score (total score and four factors) (52): \( \frac{((\text{BPRS score at second scan}) - (\text{BPRS score at baseline scan}))}{\text{BPRS score at baseline scan}} \times 100 \). Although multiple correlations were performed, results were reported as \( p < .05 \) (two-tailed); hence caution in interpretation is needed because of the exploratory nature of the evaluation.

**Results**

**Volumes at First Scan**

Initial scan volumes were computed (see Supplemental Table 1 for descriptive statistics) and graphed with both hemispheres summed (Figure 2). The three-factor ANOVA model showed a tissue by group interaction (\( F_{6,61} = 5.40, p < .0001 \)) without a main effect of group or a tissue by hemisphere by group interaction. Subsequent two-factor ANOVAs for each tissue class revealed that NCGM volume differed between groups (\( F_{2,61} = 7.65, p = .001 \)), without a hemisphere by group interaction. Post hoc pairwise comparisons revealed that, compared with HC, both FESZ (5.8% less, \( t_{63} = 3.66, p = .001 \)) and FEAFF (3.9% less, \( t_{68} = 3.00, p = .004 \)) had smaller (left + right) NCGM but did not differ from each other (\( t_{61} = 1.07, p = .29, \text{Cohen's } d = .27 \)). Groups did not differ in CWM (\( F_{2,61} = .36, p = .70 \)) without a hemisphere by group interaction. Groups differed in SCSF volume (\( F_{2,61} = 5.89, p = .004 \)), without a hemisphere by group interaction. Both FESZ (18.6% more, \( t_{63} = 3.24, p = .002 \)) and FEAFF (14.8% more, \( t_{68} = 2.78, p = .007 \)) showed greater (left + right) SCSF volume than HC but did not differ from each other (\( t_{61} = .65, p = .52, \text{Cohen's } d = .16 \)). Groups showed a marginally significant difference in LV volumes (\( F_{2,61} = 3.10, p = .050 \)), without a hemisphere by group interaction. Both FESZ (24.5% more, \( t_{63} = 2.11, p = .039 \)) and FEAFF (20.0% more, \( t_{68} = 2.08, p = .041 \)) showed greater (left + right) LV volumes than HC. Age, which was group-matched, showed significant negative associations with total NCGM relative volumes in each group (FESZ: \( \beta = -.52, t_{28} = 3.12, p = .004 \), FEAFF: \( \beta = -.60, t_{33} = 4.18, p = .0002 \), HC: \( \beta = -.40, t_{35} = 2.54, p = .016 \)), with an age by group interaction in NCGM volume (\( F_{3,95} = 17.32, p < .0001 \)).

Within each group, there was a negative correlation between NCGM and SCSF (FESZ: \( r = -.75, p < .0001, n = 29 \), FEAFF: \( r = -.48, p = .004, n = 34 \); HC: \( r = -.523, p = .001, n = 36 \)). No group showed a correlation between SCSF and LV.

**Longitudinal Volume Comparison (Figure 3)**

Groups did not differ in initial ICC, did not show ICC change over time, and did not differ in ICC percent change (\( F_{2,61} = 2.17, p = .12 \)), indicating ICC did not indirectly affect relative volume comparisons. Age was not correlated with any longitudinal volume changes. Descriptive statistics are summarized in Supplemental Table 2.

The three-factor ANOVA model showed a main effect of group (\( F_{2,61} = 6.39, p = .003 \)), a tissue by group interaction (\( F_{6,61} = 5.04, p = .0003 \)) and a tissue by hemisphere by group interaction (\( F_{6,61} = 3.57, p = .002 \)). Subsequent two-factor ANOVAs revealed that groups showed different amounts of NCGM volume change (\( F_{2,61} = 13.87, p < .0001 \)), without a hemisphere by group interaction. Post hoc pairwise comparisons on total (left + right) NCGM percent change revealed significant volume reduction in FESZ (−1.7%) compared with HC (+.05%, \( t_{41} = 2.58, p = .014 \)) and FEAFF (+3.6%, \( t_{36} = 4.18, p = .0002 \)). In contrast, FEAFF showed a significant volume increase (+3.6%) compared to HC (+.05%, \( t_{45} = 3.59, p = .001 \)) and FESZ (−1.7%, \( t_{36} = 4.18, p = .0002 \)). Although initially smaller, the relative NCGM volume in FEAFF group was not different from HC at second scan (\( t_{45} = .09, p = .93 \)). Groups did not differ in CWM volume change (\( F_{2,61} = 1.47, p = .24 \)) without the hemisphere by group interaction. Total SCSF volume change showed a trend-level main effect of group (\( F_{2,61} = 3.66, p = .001 \))
Lobar Parcellation of Neocortical Gray Matter (Figure 4)

In the cross-sectional volume comparison, there was no group difference at the initial scan. At the follow-up scan, the FESZ group showed smaller volume in all three lobar regions as compared to the HC group (frontal lobe: t_{41} = 3.15, p = .003, temporal lobe: t_{41} = 2.65, p = .011, parieto-occipital lobes: t_{41} = 2.62, p = .012), and in frontal and temporal lobes compared with the FEAFF group (frontal lobe: t_{36} = 2.84, p = .007, temporal lobe: t_{36} = 2.63, p = .013, parieto-occipital lobes: t_{36} = 2.07, p = .046), without any difference in any lobar region between FEAFF and HC groups.

In the longitudinal volume comparisons, the three-factor ANOVA model showed a main effect of group (F_{2,61} = 12.38, p < .0001). Subsequent two-factor ANOVAs revealed that each lobar region showed a main effect of group (frontal lobe: F_{2,61} = 11.36, p < .0001, temporal lobe: F_{2,61} = 8.96, p = .0004, parieto-occipital lobes: F_{2,61} = 9.73, p = .0002) without a hemisphere by group interaction. Post hoc pairwise comparisons revealed that the NCGM volume reduction observed in the FESZ group was localized to frontal and temporal lobes. The frontal lobe percent decrement in FESZ (−2.4%) was significantly more than HC (−.3%, t_{41} = 3.19, p = .003) and FEAFF (−2.8%, t_{36} = 3.85, p = .0005) and the temporal lobe percent decrement in FESZ (−2.6%) was significantly more than HC (+.5%, t_{41} = 2.79, p = .008) and FEAFF (+3.3%, t_{36} = 3.56, p = .001), whereas the parieto-occipital lobe percent change in FESZ (−7%) was not different from that in HC (+.2%, t_{41} = 1.01, p = .32). In contrast, the NCGM volume increase in the FEAFF group was more diffuse. Percent volume changes in the FEAFF group (frontal lobe: +2.8%, temporal lobe: +3.3%, parieto-occipital lobe: +4.6%) were significantly larger than FESZ (frontal lobe: t_{36} = 3.85, p = .0005, temporal lobe: t_{36} = 3.56, p = .001, parieto-occipital lobe: t_{36} = 3.32, p = .002) and HC (frontal lobe: t_{45} = 2.86, p = .006, temporal lobe: t_{45} = 2.28, p = .028, parieto-occipital lobe: t_{45} = 3.47, p = .001), although the temporal lobe volume increase compared with HC did not quite reach the Bonferroni-corrected p value.

Patient Group Subdivision by Medication History

There were no demographic or clinical differences between the medication-based patient subgroups. There were no volume differences between lithium-treated and valproate-treated patients. The 12 FEAFF on antipsychotics showed significantly more NCGM volume increase than the 9 FEAFF not on antipsychotics (5.4% vs. 1.2%, t_{19} = 2.25, p = .036), and there was no significant difference in NCGM change between FESZ subgroups on and not on antipsychotics (Figure 5). Mood stabilizer medication was not statistically significantly associated with NCGM percent volume change within each group, although effect sizes were moderate to high (Figure 5). Change in SCSF in FESZ with MS (n = 7) was nearly zero (−.
07%), whereas FESZ without MS (n = 10) showed +12.3% SCSF enlargement (t_{15} = 2.63, p = .019).

The FEAFF patients (n = 4) on neither MS nor antipsychotics showed a slight NCGM reduction (−0.2%), whereas the FEAFF patients (n = 8) on both MS and antipsychotics showed a 6.1% NCGM increase, a significant difference (t_{10} = 2.51, p = .031, Cohen's d = 1.73). In this medication contrast, FESZ showed no statistical significance, although the effect size was large (Cohen's d = 1.09).

Combining FESZ and FEAFF groups, the 20 patients on MS showed a significant NCGM increase compared with the 18 patients not on MS (+2.8% vs. −0.6%, t_{36} = 2.35, p = .024), whereas the combined FESZ/FEAFF groups on and not on antipsychotics were not statistically different (moderate effect size, Cohen's d = .60).

**Clinical Correlations with Volume Change over Time**

In FESZ, longitudinal NCGM change was negatively correlated with changes in the BPRS score (Figure 6). Whereas most patients tended to improve in their clinical status, the more the NCGM volume loss, the less the symptom improvement (or the greater the symptom deterioration) in BPRS total score (r = −.57, p = .018), thought disturbance factor (r = −.67, p = .003), and anxiety-depression factor (r = −.73, p = .001). The BPRS thought disturbance and anxiety-depression factors correlations remained significant with a Bonferroni correction for the n = 5 BPRS measures. Within the FESZ group, there was no lobar-specific clinical correlation with lobar NCGM volume changes. Also within the FESZ group, the more the bilateral LV enlargement, the higher the change in the BPRS withdrawal-retardation factor (r = .61, p = .010, Figure 6). In contrast, the FEAFF group did not show any clinical correlations with their NCGM increase (Figure 6), even in the lobar parcellation, and these correlation coefficients differed significantly (Fisher r-to-z transformation) from those of the FESZ.

Changes in CWM and SCSF did not correlate with change in total BPRS score or factors in either patient group.

**Discussion**

**Main Findings**

At the first hospitalization for psychosis, smaller NCGM volume with reciprocally enlarged SCSF space was observed in both FESZ and FEAFF. Despite these similar tissue volume abnormalities at first hospitalization, FESZ and FEAFF exhibited different longitudinal trajectories. FESZ showed progressive volume reduction of NCGM (−1.7%), bilateral LV enlargement (10.4%), and supratentorial SCSF enlargement (+7.2%). In contrast, FEAFF showed a robust NCGM volume increase (+3.6%), whereas CWM, SCSF, and LV did not show significant changes. In terms of regional differences, the FESZ NCGM volume reduction was confined to frontotemporal regions, whereas the FEAFF NCGM volume increase was more diffuse.

To our knowledge, no previous study has prospectively evaluated the total volume of NCGM in isolation from the remaining cerebrum and subcortical region in either patient group, nor has research directly compared FESZ with FEAFF, although previous reports of supratentorial total GM volume deficit in SZ (17–20) are compatible. Both FESZ and FEAFF showed reduced NCGM at first hospitalization, although our own and others' studies suggest a regionally specific pattern of NCGM deficits in SZ not shared by AFF (3).
Longitudinal Volume Changes in FE Schizophrenia

We note the post-onset progressive global NCGM volume reduction of 1.7% is considerably less than the approximately 9%–10% found in left superior temporal gyrus (53), suggesting a greater NCGM volume loss in some regions than in others. Indeed, we found a frontal (−2.4%) and a temporal (−2.6%) but no parieto-occipital NCGM reduction, largely compatible with previous longitudinal lobar parcellation studies done with different methodologies (21,22). Importantly, in FESZ, the NCGM reduction and LV enlargement were associated with smaller improvement or worsening of BPRS clinical features. These findings are similar to Cahn et al.’s (25) prospective longitudinal evaluation of supratentorial GM and LV volume, without the use of spatial normalization, and their relationship to clinical symptoms.

Our data and the cited studies, taken together, provide support for a postonset progression of schizophrenia. What process might account for the progressive loss of NCGM? Postmortem work points to neuropil reduction rather than cell loss as the main cause of gray matter reduction in schizophrenia (54,55). Our laboratory has advanced the hypothesis of a cortical circuit abnormality (deficient recurrent inhibition as a result of γ-aminobutyric acid [GABA]-ergic abnormalities) as a mechanism (56). This hypothesis is supported by preclinical work (57,58), gamma oscillations in schizophrenia (59), and postmortem work (60). In our model, dendritic remodeling and regression occurs as a result of increased excitatory input due to decreased GABAergic recurrent inhibition, which, in turn, results from hypofunction of the recurrent collateral N-methyl-D-aspartate receptor on GABAergic interneurons (61).

According to this model, antipsychotics might result in decreased GM loss by reducing the excess excitatory drive or, through modulation of GABAergic interneurons, by increasing recurrent inhibition. Our data suggest that antipsychotic medications (mainly atypical) might tend to lessen NCGM volume reduction or, at least, do not enhance the NCGM volume reduction (Figure 5). Although GABAergic abnormalities likely also exist in affective psychosis (62), our data do not indicate that they result in progressive NCGM volume reduction.

Cahn et al. (25) did not report information on MS dosage but did report, in contrast to our study, an association of increased GM loss and higher dosage of (typical and atypical) antipsychotics. Lieberman et al. (24) reported that haloperidol-treated patients showed supratentorial GM volume reduction, whereas olanzapine-treated patients showed no or minimal changes; too few patients in our study were on typical antipsychotics to allow comparison. The 10% reduction in both GM and WM after either haloperidol or olanzapine administration in macaques (29) is considerably larger than the effect in our study or any in vivo MRI study on global GM volume change and may represent a different mechanism or a species difference.

Longitudinal Volume Changes in First-Episode Affective Psychosis and Medication Effects

Whereas CWM, SCSF, and LV did not show significant volume changes, FEAFF showed a robust longitudinal NCGM volume increase (+3.6%). This increase might best be explained by MS effects: 86% of the FEAFF patients were on MS at least partly during the interscan period, and 72% were on MS throughout the interscan period. Combining subjects in both patient groups, the 20 patients on MS showed a statistically significant NCGM increase when compared with 18 patients not on MS (+2.8% vs. −.6%), although the same comparison in each patient group alone did not reach significance with the moderate to large effect sizes (Figure 5). The unmedicated FEAFF had almost no NCGM change (−.2%), whereas the MS and antipsychotic group showed a 6.1% increase (p = .031, Cohen’s d = 1.73).

It is of particular note that, unlike the FESZ NCGM reduction–clinical symptom association, the longitudinal NCGM change in FEAFF was not correlated with any clinical features,
suggesting the possibility that the NCGM volume increase might not be an intrinsic feature of affective disorder but perhaps the result of a medication effect that was not associated with clinical improvement. The nonspecific lobar pattern of NCGM volume increase is also compatible with a widespread medication effect. These data further raise the possibility that MS-associated NCGM volume increase could be a major confounding factor in both cross-sectional and longitudinal volume comparisons and possibly might account for at least some of the inconsistency and wide variability of volumetric findings in affective psychosis studies, especially for bipolar disorder (63).

Although the naturalistic study design and small numbers for subgrouping make our conclusions tentative, our data are consistent with a medication-controlled study on lithium in bipolar subjects (30) and an abstract report of volunteers taking lithium (64). Mood-stabilizer-induced GM volume increase might be related to neurotrophic–neuroprotective effects of lithium and valproate (65–67) because both agents increase the cytoprotective protein bcl-2 (B-cell lymphoma/leukemia-2 gene) and inhibit GSK-3β (glycogen synthase kinase 3β), resulting in apoptosis inhibition (68,69).

**Limitations**

There are methodologic limitations of this study. First, the sample size for longitudinal comparisons was not large enough for some patient subgroup analyses. Second, this naturalistic study does not allow medication-controlled comparisons. Therefore, some of the results from the medication subdivision analysis should be regarded as preliminary, needing to be confirmed in future medication-controlled studies. Finally, the clinical symptom correlations need to be confirmed in a future study.

**Conclusions**

At first hospitalization, both FESZ and FEAFF groups showed small overall NCGM with reciprocally enlarged SCSF compared with control subjects. They differed in longitudinal tissue volume changes, however. The FESZ subjects demonstrated progressive overall NCGM volume reduction, localized to the frontotemporal region, and LV and SCSF enlargement, associated with a worse course of clinical symptoms. The FEAFF subjects had a robust NCGM volume increase without any clinical correlation and without regional localization (47,70–72).

**Supplementary Material**

Refer to Web version on PubMed Central for supplementary material.

**Acknowledgments**

This study was supported in part by the Department of Veterans Affairs Research Enhancement Award Program (RWM, MES), Merit Awards (MES, RWM), and a Middleton Award (RWM); grants from the National Institute of Health (Grant Nos. K02 MH 01110 and R01 MH 50747 to MES, R01 MH 40799 to RWM, and R01 MH58704 to DFS); the MIND foundation (Albuquerque, New Mexico, RWM), the Mental Health Research Association (NARSAD; to DFS); and the Welfide Medicinal Research Foundation, Japan (MN). Some of the data were presented at annual meetings of Biological Psychiatry, Atlanta, Georgia, May 19, 2005; the Society for Neuroscience, Washington DC, November 16, 2005; and at the biennial Winter Workshop on Schizophrenia Research, Davos, Switzerland, February 4, 2006. We thank Marie Fairbanks for administrative support and Erin Connor, Georgia Bushell, Laura Rosow, Usman Khan, Jorge Alvarado, and Courtney Brown for research assistant support.

**References**


Biol Psychiatry. Author manuscript; available in PMC 2009 November 25.


60. Lewis DA, Hashimoto T, Volk DW. Cortical inhibitory neurons and schizophrenia. Nat Rev Neurosci 2005;6:312–324. [PubMed: 15803162]


64. Leow, AD.; Thompson, PM.; Hayashi, KM.; Bearden, C.; Nicoletti, MA.; Monkul, SE., et al. Effects of lithium treatment on human brain structure mapped using longitudinal MRI. Presented at the American College of Neuropsychopharmacology meeting; Waikoloa, Hawaii. 2005.


Figure 1.
Expectation-maximization atlas tissue segmentation and regions of interest (ROI). Top, left: Examples of spoiled-gradient-recalled images and tissue segmentation. Tissue has been segmented and parcellated into ROI of neocortical gray matter (NCGM, green on right, blue on left hemisphere), cerebral white matter (CWM, yellow on right, beige on left hemisphere), sulcal cerebrospinal fluid (SCSF, red on right, brown on left side), and lateral ventricles (LV, dark blue on right and purple on left side). Note the exclusion of subcortical nuclei, the medial temporal region, and all infratentorial tissue. Top, right: Three-dimensional reconstructions of brain tissue ROI. Bottom: Upper portion shows lobar parcellation of NCGM into frontal lobe (green on right, blue on left hemisphere), temporal lobe (purple on right, pink on left hemisphere), and parietal lobe (red on right, pink on left hemisphere).
hemisphere), and parieto-occipital lobe (red on right, brown on left hemisphere). Lower part is a three-dimensional reconstruction. L, left; R, right; A, anterior; P, posterior.
Figure 2.
Cross-sectional relative volume comparisons. Each symbol represents a subject with first-episode schizophrenia (FESZ, red square), a subject with first-episode affective psychosis (FEAFF, blue circle), or a healthy control subject (HC, black inverted triangle). Horizontal lines are group means. Both patient groups showed smaller neocortical gray matter (NCGM; \( p = .001 \)) and enlarged SCSF \( (p = .004) \), compared with HC. Both patient groups showed a marginally significant enlargement of LV compared with HC \( (p = .050) \). There was no group difference for cerebral white matter \( (p = .70) \).
Figure 3.
Percentage longitudinal volume changes over 1.5 years. Color symbols are as described in Figure 2. The number of subjects in each group showing a volume increase and decrease are indicated. Horizontal lines are group means; the numeric value of the statistically significant mean percent changes is also indicated. Subjects with first-episode schizophrenia (FESZ) showed significant volume reduction in neocortical gray matter (NCGM) compared with both subjects with first-episode affective psychosis (FEAFF; $p = .0002$) and healthy control subjects (HC; $p = .014$); FESZ also showed significant enlargement of lateral ventricle (LV) compared with FEAFF ($p = .001$) and HC ($p < .0001$). Sulcal cerebrospinal fluid was enlarged in FESZ compared with HC ($p = .015$). The FEAFF subjects showed a volume increase in NCGM compared with FESZ ($p = .0002$) and HC ($p = .001$).
Figure 4.
Relative volumes and percent volume changes in the neocortical gray matter (NCGM) lobar parcellation. Because hemisphere by group interaction was not observed in any lobar region, bilateral volumes (left + right hemispheres) are used in this figure. In the scattergrams of relative volume, each subject’s volumes at the initial scan (T1) and the follow-up scan (T2) are connected with a straight line. Mean value of cross-sectional volume at T1 and T2 in each group is indicated by the open bar. Of particular note, the first-episode schizophrenia (FESZ) group shows NCGM volume reduction in frontal and temporal lobes but not in parieto-occipital lobe. In contrast, the first-episode affective psychosis (FEAFF) group shows a global NCGM volume increase from T1 to T2, with cross-sectional volumes at the follow-up scan (T2) similar to the HC group.
Figure 5.
Percentage volume changes of NCGM and medication in the interscan interval. Left panel: Purple symbols indicate patients who received only atypical antipsychotics (AP). Pink symbols indicate patients who took both typical and atypical AP. Note the large effect size of AP in reducing NCGM volume loss in schizophrenia. Right panel: Green symbols indicate patients on mood stabilizers (MS, lithium, and/or valproate). Black symbols indicate patients who were medication free for more than 3 months before the second scan due to noncompliance or nonprescription.
Figure 6.
Correlation between percentage volume change and symptom change in Brief Psychiatric Rating Scale (BPRS). In Y axes, plus values mean symptom deterioration over time and minus values mean symptom improvement. Note the association between the degree of reduction of neocortical gray matter (NCGM) in subjects with first-episode schizophrenia and poorer symptom outcomes. Note also the absence of correlations with increased NCGM in subjects with first-episode affective psychosis (FEAFF). Pearson’s Product–Moment Correlation Coefficient \( r \) and its \( p \) value are indicated within each diagram. LV, lateral ventricle.
Table 1

Demographic and Clinical Characteristics of Cross-Sectional Study Groups

<table>
<thead>
<tr>
<th>Variable</th>
<th>Schizophrenic Patients (n = 29)</th>
<th>Affective Psychosis Patients (n = 34)</th>
<th>Healthy Control Subjects (n = 36)</th>
<th>F or t Test or χ² Values</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (year)</td>
<td>24.3 (5.8) [18–40]</td>
<td>22.1 (3.1) [18–30]</td>
<td>22.9 (3.8) [18–34]</td>
<td>2.96</td>
<td>2.08</td>
</tr>
<tr>
<td>Gender (male/female)</td>
<td>24/5</td>
<td>26/8</td>
<td>31/5</td>
<td>2</td>
<td>1.11</td>
</tr>
<tr>
<td>Handedness misconduct</td>
<td>.78 (.22)</td>
<td>.72 (.29)</td>
<td>.76 (.20)</td>
<td>2.90</td>
<td>.49</td>
</tr>
<tr>
<td>Socioeconomic Status misconduct</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subject’s</td>
<td>3.3 (1.3)</td>
<td>2.7 (1.2)</td>
<td>2.3 (.9)</td>
<td>2.94</td>
<td>6.98</td>
</tr>
<tr>
<td>Parental</td>
<td>1.9 (1.0)</td>
<td>1.6 (.8)</td>
<td>1.5 (.7)</td>
<td>2.96</td>
<td>2.72</td>
</tr>
<tr>
<td>Education (school year)</td>
<td>13.4 (1.8)</td>
<td>13.9 (1.8)</td>
<td>15.1 (1.9)</td>
<td>2.91</td>
<td>6.92</td>
</tr>
<tr>
<td>WAIS-R: Information, Scaled</td>
<td>11.4 (2.9)</td>
<td>12.6 (2.9)</td>
<td>13.3 (2.1)</td>
<td>2.88</td>
<td>3.79</td>
</tr>
<tr>
<td>WAIS-R: Digits Span, Scaled</td>
<td>10.0 (2.5)</td>
<td>10.3 (2.7)</td>
<td>11.4 (2.9)</td>
<td>2.88</td>
<td>2.40</td>
</tr>
<tr>
<td>Antipsychotic Medication Dosage</td>
<td>266.3 (185.6)</td>
<td>243.8 (221.7)</td>
<td>NA</td>
<td>1.59</td>
<td>.18</td>
</tr>
<tr>
<td>Median Duration of Antipsychotic</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medication Before Scan (week)</td>
<td>3 [0–24]</td>
<td>1 [0–24]</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

WAIS-R, Wechsler Adult Intelligence Scale—Revised (70); NA, data not applicable.

* p < .05;

** p < .01.

a The dfs differ among variables owing to unavailability of data from some participants.

b Handedness was evaluated using the Edinburgh inventory (40), and right-handedness is positive.

c Higher scores indicate lower socioeconomic status (39).

d Chlorpromazine equivalent dosage (41) during the first hospitalization.
Table 2
Demographic and Clinical Characteristics of Longitudinal Study Groups

<table>
<thead>
<tr>
<th>Variable</th>
<th>Schizophrenia Patients (n = 17)</th>
<th>Affective Psychosis Patients (n = 21)</th>
<th>Healthy Control Subjects (n = 26)</th>
<th>df</th>
<th>F or t Test or χ² Values</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (year)</td>
<td>24.7 (7.0) [18–40]</td>
<td>22.4 (3.2) [18–30]</td>
<td>23.6 (4.1) [18–34]</td>
<td>2, 61</td>
<td>1.06</td>
<td>.35</td>
</tr>
<tr>
<td>Baseline Scan</td>
<td>24.7 (7.0) [18–40]</td>
<td>22.4 (3.2) [18–30]</td>
<td>23.6 (4.1) [18–34]</td>
<td>2, 61</td>
<td>1.06</td>
<td>.35</td>
</tr>
<tr>
<td>Second Scan</td>
<td>26.0 (6.8) [18–41]</td>
<td>23.7 (3.2) [19–32]</td>
<td>25.1 (4.0) [19–36]</td>
<td>2, 61</td>
<td>1.20</td>
<td>.31</td>
</tr>
<tr>
<td>Inter-scan Interval (month)</td>
<td>18.1 (11.6)</td>
<td>18.7 (9.6)</td>
<td>18.1 (8.2)</td>
<td>2, 61</td>
<td>0.03</td>
<td>.97</td>
</tr>
<tr>
<td>Gender (male/female)</td>
<td>14/3</td>
<td>17/4</td>
<td>22/4</td>
<td>2</td>
<td>.11</td>
<td>.95</td>
</tr>
<tr>
<td>Handedness</td>
<td>.77 (.25)</td>
<td>.69 (.31)</td>
<td>.75 (.22)</td>
<td>2, 59</td>
<td>.51</td>
<td>.60</td>
</tr>
<tr>
<td>Socioeconomic Status</td>
<td>3.2 (1.2)</td>
<td>2.8 (1.2)</td>
<td>2.3 (0.9)</td>
<td>2, 61</td>
<td>3.74</td>
<td>.03ᵇ</td>
</tr>
<tr>
<td>Subject's</td>
<td>1.9 (0.8)</td>
<td>1.7 (1.0)</td>
<td>1.4 (0.6)</td>
<td>2, 61</td>
<td>1.77</td>
<td>.18</td>
</tr>
<tr>
<td>Parental</td>
<td>14.0 (2.1)</td>
<td>14.7 (2.1)</td>
<td>15.9 (1.6)</td>
<td>2, 61</td>
<td>5.78</td>
<td>.005**</td>
</tr>
<tr>
<td>Education (school year)</td>
<td>11.7 (3.0)</td>
<td>13.4 (2.7)</td>
<td>13.3 (2.3)</td>
<td>2, 58</td>
<td>2.45</td>
<td>.10</td>
</tr>
<tr>
<td>WAIS-R: Information, Scaled (baseline)</td>
<td>9.6 (2.0)</td>
<td>11.0 (2.7)</td>
<td>11.4 (3.0)</td>
<td>2, 60</td>
<td>2.31</td>
<td>.11</td>
</tr>
<tr>
<td>WAIS-R: Digits span, Scaled (baseline)</td>
<td>27.5 (2.7)</td>
<td>28.8 (1.4)</td>
<td>28.9 (1.1)</td>
<td>2, 60</td>
<td>4.04</td>
<td>.02³</td>
</tr>
<tr>
<td>MMSE</td>
<td>27.8 (2.6)</td>
<td>29.5 (1.0)</td>
<td>29.2 (1.0)</td>
<td>2, 60</td>
<td>6.05</td>
<td>.004**</td>
</tr>
<tr>
<td>BPRS</td>
<td>41.5 (13.7)</td>
<td>33.7 (7.5)</td>
<td>NA</td>
<td>1, 36</td>
<td>5.09</td>
<td>.03ᵇ</td>
</tr>
<tr>
<td>Baseline Scan</td>
<td>27.8 (6.7)</td>
<td>27.5 (1.0)</td>
<td>NA</td>
<td>1, 36</td>
<td>1.71</td>
<td>.20</td>
</tr>
<tr>
<td>Second Scan</td>
<td>52.5 (12.6)</td>
<td>60.7 (18.3)</td>
<td>NA</td>
<td>1, 36</td>
<td>2.46</td>
<td>.13</td>
</tr>
<tr>
<td>Medication History Before the Second Scan</td>
<td>35.9 (7.8)</td>
<td>39.2 (7.9)</td>
<td>NA</td>
<td>1, 36</td>
<td>1.36</td>
<td>.20</td>
</tr>
<tr>
<td>AP-Treated/MS-Free (&gt;3 months)</td>
<td>52.5 (12.6)</td>
<td>60.7 (18.3)</td>
<td>NA</td>
<td>1, 36</td>
<td>2.46</td>
<td>.13</td>
</tr>
<tr>
<td>AP-Free/MS-Free (&gt;3 months)</td>
<td>35.9 (7.8)</td>
<td>39.2 (7.9)</td>
<td>NA</td>
<td>1, 36</td>
<td>1.36</td>
<td>.20</td>
</tr>
<tr>
<td>AP, Antipsychotics; BPRS, Brief Psychiatric Rating Scale (52); GAS, Global Assessment Scale (72); MMSE, Mini-Mental Examination (71); MS, Mood Stabilizers; NA, data not applicable; WAIS-R, Wechsler Adult Intelligence Scale-Revised (70).</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>*</td>
<td>p &lt; .05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>**</td>
<td>p &lt; .01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b</td>
<td>Handedness was evaluated using the Edinburgh inventory (40), and right-handedness is positive.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c</td>
<td>Higher scores indicate lower socioeconomic status (39).</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d</td>
<td>Antipsychotic-free and MS-free subgroups were not medicated for more than 3 months before the second scan because of noncompliance or nonprescription. For antipsychotic-treated subgroups, only one first-episode schizophrenia (FESZ) patient and two first-episode affective psychosis (FEAFF) patients were on both typical and atypical antipsychotics. All other antipsychotic-treated patients were only on atypical antipsychotics before the second scan. The antipsychotic-treated FESZ patients received olanzapine (75%), risperidone (17%), clozapine (17%), and quetiapine (8%). The antipsychotic-treated FEAFF patients received quetiapine (42%), olanzapine (33%), risperidone (25%), and clozapine (17%). The MS-treated FESZ patients received lithium (57%) and valproate (43%), and the MS-treated FEAFF patients received lithium (54%) and valproate (54%). Note: some patients received more than one antipsychotic and some more than one MS at a time.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The dfs differ among variables because data from some participants were unavailable.

Handedness was evaluated using the Edinburgh inventory (40), and right-handedness is positive.

Higher scores indicate lower socioeconomic status (39).