A multi-scanner study of subcortical brain volume abnormalities in schizophrenia

Theo G.M. van Erp[a,*], **Douglas N. Greve**[b], **Jerod Rasmussen**[a], **Jessica Turner**[c,d], **Vince D. Calhoun**[c,e], **Sarah Young**[a], **Bryon Mueller**[f], **Gregory G. Brown**[g], **Gregory McCarthy**[h], **Gary H. Glover**[i], **Kelvin O. Lim**[f], **Juan R. Bustillo**[d], **Aysenil Belger**[j], **Sarah McEwen**[k], **James Voyvodic**[l], **Daniel H. Mathalon**[m], **David Keator**[a], **Adrian Preda**[a], **Dana Nguyen**[a], **Judith M. Ford**[m], **Steven G. Potkin**[a], and **FBIRN**[n]

[a]Department of Psychiatry and Human Behavior, University of California Irvine, Irvine, CA 92617, United States

[b]Department of Radiology, Massachusetts General Hospital, Boston, MA, 02115, United States

[c]Mind Research Network, Albuquerque, NM 87106, United States

[d]Departments of Psychiatry & Neuroscience, University of New Mexico, Albuquerque, NM 87131, United States

[e]Department of Electrical and Computer Engineering, University of New Mexico, Albuquerque, NM, United States

[f]Department of Psychiatry, University of Minnesota, Minneapolis, MN 55454, United States

[g]VA San Diego Healthcare System and Department of Psychiatry, University of California San Diego, CA 92161, United States

[h]Department of Psychology, Yale University, New Haven, CT 06250, United States

[i]Department of Radiology, Stanford University, Stanford, CA 94305, United States

[j]Departments of Psychiatry and Psychology, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, United States

[k]Department of Psychology, University of California, Los Angeles, Los Angeles, CA 90095, United States

[l]Brain Imaging and Analysis Center, Duke University Medical Center, Durham, NC 27710, United States

[m]Department of Psychiatry, University of California, San Francisco, San Francisco, CA 94143, United States

*Corresponding Author: Theo G.M. van Erp, Department of Psychiatry and Human Behavior, School of Medicine, University of California Irvine, 5251 California Avenue, Suite 240, Irvine, CA 92617, USA, Tel. +1 (949) 824-3331, fax: +1 (949) 924-3324, tvanerp@uci.edu.

[n]http://www.birncommunity.org

Abstract

Schizophrenia patients show significant subcortical brain abnormalities. We examined these abnormalities using automated image analysis software and provide effect size estimates for prospective multi-scanner schizophrenia studies. Subcortical and intracranial volumes were obtained using FreeSurfer 5.0.0 from high-resolution structural imaging scans from 186 schizophrenia patients (mean age±SD=38.9±11.6, 78% males) and 176 demographically similar controls (mean age±SD=37.5±11.2, 72% males). Scans were acquired from seven 3-Tesla scanners. Univariate mixed model regression analyses compared between-group volume differences. Weighted mean effect sizes (and number of subjects needed for 80% power at $\alpha=0.05$) were computed based on the individual single site studies as well as on the overall multi-site study. Schizophrenia patients have significantly smaller intracranial, amygdala, and hippocampus volumes and larger lateral ventricle, putamen and pallidum volumes compared with healthy volunteers. Weighted mean effect sizes based on single site studies were generally larger than effect sizes computed based on analysis of the overall multi-site sample. Prospectively collected structural imaging data can be combined across sites to increase statistical power for meaningful group comparisons. Even when using similar scan protocols at each scanner, some between-site variance remains. The multi-scanner effect sizes provided by this study should help in the design of future multi-scanner schizophrenia imaging studies.

Keywords

Psychosis; Magnetic resonance imaging (MRI); Effect size; Imaging

1. Introduction

Schizophrenia patients show significant structural brain abnormalities when studied with magnetic resonance imaging (MRI). In vivo study of these abnormalities may aid in our understanding of etiology, pathogenesis, and treatment effects. In this study we examine whether subcortical volume alterations can be observed in prospective multi-center imaging studies despite additional between-scanner variance. We provide effect size estimates for single center (based on meta-analysis of single site effects) versus multi-center (based on mega-analysis correcting for site effects) structural imaging studies in schizophrenia.

Effect size estimates for structural brain alterations in schizophrenia are predominantly based on single center studies (Haijma et al., 2013; Shepherd et al., 2012); but for a simulation of multi-center study effect sizes, see Suckling et al. (Suckling et al., 2010). The increase in multi-scanner imaging studies, as well as increased efforts towards data sharing, emphasizes the need for effect-size estimates for multi-scanner data acquisitions. The ability to detect statistically significant differences between conditions depends on the effect size, sample size, α-level, and power of the test (Cohen, 1992). In power analyses, the researcher sets the desired α-level and power of the test. The effect size is preferably gleaned from the literature or otherwise estimated, and the sample size that will be required to observe a statistically significant effect is estimated.

The effect size for mean comparisons can be computed as the mean difference between two conditions divided by the pooled standard deviation of the measurements (Cohen, 1992) and thus depends on measurement variability. In single scanner studies, such variability depends on subject variability, between-acquisition scanner variability, and measurement-method reliability. In multi-scanner studies it also depends on between-scanner and other between-site (e.g., sample demographics) variance. Subject variability depends on the relative homogeneity or heterogeneity of the sample(s). Between-acquisition scanner variability depends on the stability over time of the MRI scanner. Brain-measurement reliabilities are estimated from multiple measurements on the same cases and include inter- and intra-scanner reliability (Jovicich et al., 2009), rater reliability (van Erp et al., 2004), and measurement-method reliability (Dewey et al., 2010; Tae et al., 2008; Wonderlick et al., 2009).

Measurements should not only be reliable but also valid. A measure is considered valid when the inferences made from it are appropriate, meaningful, and useful. The calculation of inter-method reliability in which a new method is compared to a GOLD standard, or a method that has been shown to produce valid measurements, provides one way to validate a measurement method. The more similar the measurements are (the higher the intra-class correlation), the more valid the measurements based on the new method. Nevertheless, validation should also be established by confirming that meaningful variability can be observed with the measurements.

Given between-scanner variability, the question remains as to how many additional data sets need to be collected in multi-scanner versus single scanner studies to observe differences between schizophrenia patients and controls? In this study, we compare subcortical volumes between chronic schizophrenia patients and healthy volunteers, and we report the weighted mean effect sizes as well as multi-center-based (*n*=7) effect sizes for subcortical volumes. Based on the effect sizes reported in meta-analyses (Haijma et al., 2012; Shepherd et al., 2012) (see Table 4), we hypothesized that we would find smaller amygdala, hippocampus, and intracranial volume and larger lateral ventricle and pallidum volumes in patients with schizophrenia compared with healthy volunteers.

2. Methods

2.1. Participants

The participants comprised 186 schizophrenia patients (mean age±SD=38.9±11.6, 145 males) and 176 healthy volunteers (mean age±SD=37.5±11.2, 126 males) with similar mean age, sex, handedness, and race distributions from seven sites (Table 1; see Supplement 1, Table 1S, for demographic data by site). Patient inclusion criteria were schizophrenia diagnosis based on the Structured Clinical Interview for DSM-IV-TR Axis I Disorders (First et al., 2002b). All patients were clinically stable outpatients whose antipsychotic medications and doses had not changed within the last 2 months. Current neuroleptic medication data were available for 171 of the 186 patients (antipsychotics: 136 atypical, 20 typical, 10 both; mood stabilizers: 2, and anxiolytics: 3). Chlorpromazine-equivalent dosages could be computed for 151 patients (Woods et al., 2005). Schizophrenia patients and healthy volunteers with a history of major medical illness, drug dependence in the last 5

years, current substance abuse disorder, MRI contraindications, or eyesight not correctable to normal acuity with MRI-compatible lenses were excluded. Patients with significant extrapyramidal symptoms and healthy volunteers with a current or past history of major neurological or psychiatric illness (First et al., 2002a) or with a first degree relative with an Axis-I psychotic disorder diagnosis were also excluded. Patient's clinical assessments included the Positive and Negative Syndrome Scale (Kay et al., 1989). All participants were assessed for socioeconomic status (Hollingstead, 1975), handedness (Oldfield, 1971), basic demographics, and premorbid IQ (Uttl, 2002). The sample includes 137 paranoid, 7 disorganized, 30 undifferentiated, and 12 residual patients. Before data collection, experienced clinicians were jointly trained on the clinical assessment rating scales with patient interviews. The raters' assessments were compared with expert ratings. Additional training was provided when raters deviated by more than 1 point for each item from the expert ratings.

Written informed consent, including permission to share de-identified data between the centers and with the wider research community, approved by the University of California Irvine, Los Angeles, and San Francisco, Duke University, University of North Carolina, New Mexico, Iowa, and Minnesota Institutional Review Boards, was obtained from all study participants.

2.2. Image acquisition

High-resolution structural brain scans were acquired on six 3T Siemens Tim® Trio System and one 3T General Electric (GE) Discovery MR750 scanner using standardized sequences. Siemens MP-RAGE scan parameters were: TR/TE/TI=2300/2.94/1100 ms, flip angle=9°, resolution=256×256×160. GE IR-SPGR scan parameters were: TR/TE/TI=5.95/1.99/450 ms, flip angle=12°, resolution=256×256×166. All scans covered the entire brain with field of view (FOV)=220 mm^2, voxel size=0.86×0.86×1.2mm, sagittal scan plane, GRAPPA/ ASSET acceleration factor=2, and number of excitations (NEX)=1. The 3T MRI scanners at each site were calibrated to meet FBIRN Phantom quality assurance (Friedman and Glover, 2006a, 2006b; Greve et al., 2011), and scanner quality was monitored during the entire study. Before the study began, a traveling engineer visited each site to review MRI study protocol adherence based on FBIRN's multi-center imaging study recommendations (Glover et al., 2012).

2.3. Image processing

Left and right lateral ventricle, thalamus, caudate, putamen, pallidum, hippocampus, and amygdala volumes as well as intracranial volumes (Table 2; see Supplement 1, Table 2S for absolute volumes by site) were obtained using Freesurfer Version 5.0.0 (http:// surfer.nmr.mgh.harvard.edu; (Fischl, 2012; Fischl et al., 2002). All regions of interest with a volume larger or smaller than 1.5 times the inter-quartile range (IQR) were identified and visually inspected by overlaying them on the subject's anatomical images. Based on these inspections, data from one patient were removed due to poor image quality and 0–2 additional data points for each of the regions of interest were eliminated from the final analyses.

2.4. Statistical analyses

Group differences for each region were examined using univariate mixed model regression analyses (Proc Mixed, SAS v9.2, SAS Institute Inc.) predicting subcortical volumes with group, site, sex, age, group × site, group × hemisphere, site × hemisphere, and group × site × hemisphere interactions. Hemisphere and intracranial volume entered the model as repeated measures and covariate variables, respectively. For the regions with a-priori-defined directional hypotheses based on the literature (amygdala, hippocampus, intracranial, lateral ventricle, and pallidum), the significance threshold was set at $p<0.05$ (one-tailed). We also indicate which of the findings pass the more conservative Bonferroni multiple comparison-corrected threshold of $p<0.00625$ (two-tailed). Based on evidence for the effects of antipsychotic medication on basal ganglia volumes (Gur et al., 1998; Hulshoff Pol and Kahn, 2008; Boonstra et al., 2011; Li et al., 2011), we examined the effect of current antipsychotic medication type (typical/atypical) and dose (chlorpromazine-equivalent dose) on basal ganglia volumes, using mixed model regression and correlation analysis, respectively. Cohen's *d* weighted mean effect sizes based on each of the seven single site samples and also based on the full sample were computed. Sample size estimates were based on computed effect sizes using one-tailed *t*-tests in G*Power Version 3.2.1 (Erdefelder et al., 1996).

3. Results

3.1. Group effects

There were significant effects of group (schizophrenia patient, healthy volunteer) on hippocampus, pallidum, putamen, lateral ventricle, amygdala and intracranial volumes (Figs. 1A–1F), with hippocampus, amygdala, and intracranial volumes smaller and pallidum, putamen, and lateral ventricle volumes larger in schizophrenia patients compared with controls (Table 3, Fig. 1; see Supplement 1, Table 3S*p* for statistical results by site). The effects for hippocampus and pallidum also passed the conservative Bonferroni two-tailed -value threshold of $p<0.00625$.

3.2. Site and group by site interaction effects

There were significant effects of site on hippocampus, putamen, thalamus, and intracranial volume, and a marginal effect of site ($p=0.05$) on caudate volume. Group × site interactions were confined to the amygdala. Decomposition of the interaction effect showed that at one site amygdala volume in patients was larger than in controls ($t_{349}=2.97$, $p=0.003$), at two sites amygdala volume was smaller in patients compared with controls ($t_{349}=-2.57$, $p=0.01$; $t_{349}=-2.0$, $p<0.05$), and at four sites amygdala volumes did not differ between the groups ($t_{349}=0.05$, $p=0.96$; $t_{349}=-1.36$, $p=0.17$; $t_{349}=0.19$, $p=0.85$, $t_{349}=-0.07$, $p=0.94$).

3.3. Other effects

There were significant hemisphere effects on all subcortical volumes (Table 3) with left hippocampus, amygdala, caudate, and thalamus volumes smaller than right, and right putamen, pallidum, and ventricular volumes smaller than left (Table 2). There were significant hemisphere × site interactions on amygdala, pallidum, and thalamus volumes. Decomposition of the hemisphere × site interactions showed that left amygdala, pallidum,

and thalamus volumes were smaller than right for all the sites, but that the strength of the differences varied by site. There were significant effects of sex on hippocampus, amygdala, putamen, pallidum, and intracranial volumes, with women showing smaller volumes than men. Age showed a significant negative association with subcortical and intracranial volumes and a significant positive association with lateral ventricle volumes. Intracranial volume was positively associated with all measured volumes (Table 3).

3.4. Effect-size estimates

The effect sizes based on the full multisite sample were on average 13% smaller than the effect-size estimates based on the weighted mean effect sizes from each individual site (Table 4). The sample sizes required for 80% power at α-level=0.05 for the five regions that showed significant group differences were between 1.3 and 3.4 times larger for multi-scanner compared with single scanner data acquisitions.

3.5. Antipsychotic medications

Current medication dose in chlorpromazine equivalents was negatively associated with left caudate ($r_{150}=-0.16$, $p<0.05$), and left and right putamen ($r_{150}=-0.20$, $p=0.01$; $r_{150}=-0.19$, $p=0.02$)volumes. There were no effects of typical versus atypical antipsychotics on any of the regions.

4. Discussion

The principal findings of this study are as follows: (1) that we confirm smaller hippocampus, amygdala, and intracranial volumes and larger lateral ventricle, putamen, and pallidum volumes in patients with schizophrenia compared with healthy volunteers based on a prospective seven-site imaging study; (2) that significant site effects are present for hippocampus, putamen, thalamus, and intracranial volumes; and (3) that effect sizes for regional volume differences based on the multisite sample analysis were on average 13% smaller than those based on the weighted single site sample means.

The findings of smaller hippocampus, amygdala, and intracranial volumes and larger lateral ventricle, putamen, and pallidum volumes in patients with schizophrenia compared with healthy volunteers are consistent with and validate those observed in single site studies (Shepherd et al., 2012). While ventricle size showed the largest percentage between-group difference in volume, consistent with the fact that this is among the most robust structural brain abnormalities observed in patients with schizophrenia, surprisingly its effect size was relatively small. The effect size for amygdala volume deficit in schizophrenia patients compared with controls was also relatively small (see Table 4) (Shepherd et al., 2012). The relatively small effect sizes for these regions are unlikely due to multi-scanner acquisition given that for these regions (a) site effects are not significant and (b) the weighted single-scanner and multi-site effect sizes are quite similar. The small effect size for ventricle volume is likely due to large variability in this sample (see SD, Table 2). With regard to amygdala volumes, FreeSurfer reliability estimates are high (Jovicich et al., 2009; Morey et al., 2010), but validity may be questionable given reports of relatively low correlations with volumes derived from manual amygdala tracings (amygdala: $r=0.56$; for comparison

hippocampus: r=0.82) (Morey et al., 2009), though see Dewey et al. (Dewey et al., 2010). Lower amygdala volumes are generally not observed in first episode schizophrenia, but there is high-moderate quality evidence for lower amygdala volumes in chronic schizophrenia (Shepherd et al., 2012). Given that our sample is older than in those reported in the meta-analyses of Shepherd et al. (2012) and Haijama et al. (2012), demographic differences other than age (e.g., medication treatment, diagnoses) may play a role. For instance, significantly lower amygdala volumes are also observed in patients with mood disorders (Hamilton et al., 2008; Sacher et al., 2012), and future work should investigate whether lower amygdala volumes are more robust in patients with schizoaffective disorder than schizophrenia. Schizoaffective disorder patients are often included in schizophrenia studies but are excluded from this study.

It is not surprising that we did not find group differences for caudate and thalamus volumes as the ranges of effect sizes for these structures in chronic schizophrenia patients include 0, though caudate and thalamus volumes appear lower in antipsychotic-naïve patients and may normalize after medication treatment (Haijma et al., 2012; Shepherd et al., 2012). In contrast to our expectations, we found that current medication doses in chlorpromazine equivalents were negatively associated with left caudate and bilateral putamen volumes. These findings are in contrast to longitudinal studies that have reported that antipsychotic treatments normalize caudate volume deficits observed in medication-naïve patients (Haijma et al., 2012) and increase putamen and pallidum volumes (Boonstra et al., 2011; Gur et al., 1998; Hulshoff Pol and Kahn, 2008; Li et al., 2011), though additional, systematic, long-term (> 1 year) longitudinal studies are likely required to more fully understand the effects of antipsychotic treatment on brain morphology.

Right-left asymmetries for hippocampus, amygdala, caudate, and lateral ventricle, as well as left-right asymmetries for putamen and pallidum volumes, have been reported in studies using gold-standard manual tracing methods (Aas et al., 2012; Erdogan et al., 2004; Peterson et al., 1993; Van Erp et al., 2002), and similar asymmetries observed in this study lend credibility to the quality of FreeSurfer's automated segmentations.

In the analyses on the full data set, we found significant effects of site on hippocampus putamen, thalamus, and intracranial volumes. These site effects may be due to differences in sample demographics between sites or differences between scanners. Upon further investigation, the observed effect of site was mainly due to higher volumes produced by site 1 (GE scanner, sample with lowest mean age). When data from this center were excluded from the analysis, all effects of site, except for those on the hippocampus, were rendered non-significant. Closer investigation showed that the remaining site effect was due to lower hippocampus volumes for site 3 compared with all the other sites and that patients in this site had significantly older mean age than those of the other sites. While determination of scanner effects requires a traveling subject study, in which the same subjects are scanned at different sites, the main take-home message from our analysis is that careful consideration of site effects may provide clues about factors that influence sample heterogeneity in observed associations.

To our knowledge, this is the first multi-site study of subcortical brain volumes in chronic schizophrenia that provides estimates of the costs associated with multi-site versus single site data collection/analysis by providing effect sizes for mega- vs. meta-analysis of multi-center data. When analyzing data from seven scanners, we observed the need for on average a 1.26 to 3.4 times increase in sample size to detect the same volume differences observed when conducting single scanner studies. In contrast to an earlier report (Jovicich et al., 2009), we found that multi-site acquisition of hippocampus volume data required a 1.26 times larger data sample to obtain similar statistical power to that of data acquisition at a single scanner.

Strengths of the study are as follows: (1) the results are based on a largely fully automated analysis with few exclusions of subjects based on quality-assurance procedures; (2) the study includes a large cohort of chronic schizophrenia patients and controls such that variance estimates and therefore effect size estimates are likely to be robust; (3) the study includes data collected from seven scanners – albeit six Siemens 3T and one General Electric 3T– providing ample opportunity for between-scanner variability; and (4) the study was carefully planned and used the FBIRN phantom for scanner calibration, a traveling engineer to review protocol compliance before study initiation, and standardized scan sequences across scanners (Friedman and Glover, 2006a, 2006b; Greve et al., 2011; Glover et al., 2012)

Some study weaknesses must be noted as follow: (1) the focus of FBIRN was not on improving brain morphometric measures across scanners and included a standard high-resolution T1-weighted scan for registration purposes in functional imaging studies. Hence, higher effect sizes may be obtained for sequences optimized for morphometric studies (Jovicich et al., 2009); (2) the study included only one GE and no Philips scanners; (3) the study did not measure lifetime medication exposure or employ a longitudinal design to study medication effects; and (4) demographics were balanced within site, but variance in demographics (mean age, race distribution, and sex distribution) between sites was allowed. Such variance can improve generalizability of findings, but it also contributes to sample heterogeneity. Whether to allow for such variance should be taken into account when planning multi-center studies.

In conclusion, we were able to confirm the observations of smaller hippocampus, amygdala, and intracranial volumes and larger lateral ventricle, putamen, and pallidum volumes in patients with schizophrenia compared with healthy volunteers in a carefully planned prospective, multi-site imaging study. The effect sizes provided should help in the design of future multi-scanner schizophrenia imaging studies. Although some extra data collection is required in multi-scanner studies, benefits of multi-center data collection (Glover et al., 2012) may in many cases outweigh the cost.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

Aas M, Navari S, Gibbs A, Mondelli V, Fisher HL, Morgan C, Morgan K, MacCabe J, Reichenberg A, Zanelli J, Fearon P, Jones PB, Murray RM, Pariante CM, Dazzan P. Is there a link between childhood trauma, cognition, and amygdala and hippocampus volume in first-episode psychosis? Schizophrenia Research. 2012; 137:73–79. [PubMed: 22353995]

Boonstra G, van Haren NE, Schnack HG, Cahn W, Burger H, Boersma M, de Kroon B, Grobbee DE, Hulshoff Pol HE, Kahn RS. Brain volume changes after withdrawal of atypical antipsychotics in patients with first-episode schizophrenia. Journal of Clinical Psychopharmacology. 2011; 31:146–153. [PubMed: 21346618]

Cohen J. A power primer. Psychological Bulletin. 1992; 112:155–159. [PubMed: 19565683]

Dewey J, Hana G, Russell T, Price J, McCaffrey D, Harezlak J, Sem E, Anyanwu JC, Guttmann CR, Navia B, Cohen R, Tate DF. Reliability and validity of MRI-based automated volumetry software relative to auto-assisted manual measurement of subcortical structures in HIV-infected patients from a multisite study. Neuroimage. 2010; 51:1334–1344. [PubMed: 20338250]

Erdefelder E, Faul F, Buchner A. GPOWER: A general power analysis porgram. Behavior Research Methods, Instruments, & Computers. 1996; 28:1–11.

Erdogan AR, Dane S, Aydin MD, Ozdikici M, Diyarbakirli S. Sex and handedness differences in size of cerebral ventricles of normal subjects. International Journal of Neuroscience. 2004; 114:67–73. [PubMed: 14660068]

First, MB.; Spitzer, RL.; Gibbon, MG.; Williams, JBW. Structured Clinical Interview for DSM-IV-TR Axis I Disorders - Non-Patient Edition (SCID-I/NP, 11/2002 revision). New York State Psychiatric Institute; New York, NY: 2002a.

First, MB.; Spitzer, RL.; Gibbon, MG.; Williams, JBW. Structured Clinical Interview for DSM-IV-TR Axis I Disorders - Patient Edition (SCID-I/P, 11/2002 revision). New York State Psychiatric Institute; New York, NY: 2002b.

Fischl B. FreeSurfer. Neuroimage. 2012; 62:774–781. [PubMed: 22248573]

Fischl B, Salat DH, Busa E, Albert M, Dieterich M, Haselgrove C, van der Kouwe A, Killiany R, Kennedy D, Klaveness S, Montillo A, Makris N, Rosen B, Dale AM. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. Neuron. 2002; 33:341–355. [PubMed: 11832223]

Friedman L, Glover GH. Reducing interscanner variability of activation in a multicenter fMRI study: controlling for signal-to-fluctuation-noise-ratio (SFNR) differences. Neuroimage. 2006a; 33:471–481. [PubMed: 16952468]

Friedman L, Glover GH. Report on a multicenter fMRI quality assurance protocol. Journal of Magnetic Resonance Imaging. 2006b; 23:827–839. [PubMed: 16649196]

Glover GH, Mueller BA, Turner JA, van Erp TG, Liu TT, Greve DN, Voyvodic JT, Rasmussen J, Brown GG, Keator DB, Calhoun VD, Lee HJ, Ford JM, Mathalon DH, Diaz M, O'Leary DS, Gadde S, Preda A, Lim KO, Wible CG, Stern HS, Belger A, McCarthy G, Ozyurt B, Potkin SG. Function biomedical informatics research network recommendations for prospective multicenter functional MRI studies. Journal of Magnetic Resonance Imaging. 2012; 36:39–54. [PubMed: 22314879]

Greve DN, Mueller BA, Liu T, Turner JA, Voyvodic J, Yetter E, Diaz M, McCarthy G, Wallace S, Roach BJ, Ford JM, Mathalon DH, Calhoun VD, Wible CG, Brown GG, Potkin SG, Glover G. A

novel method for quantifying scanner instability in fMRI. Magnetic Resonance in Medicine. 2011; 65:1053–1061. [PubMed: 21413069]

Gur RE, Maany V, Mozley PD, Swanson C, Bilker W, Gur RC. Subcortical MRI volumes in neuroleptic-naive and treated patients with schizophrenia. American Journal of Psychiatry. 1998; 155:1711–1717. [PubMed: 9842780]

Haijma SV, Van Haren N, Cahn W, Koolschijn PC, Hulshoff Pol HE, Kahn RS. Brain volumes in schizophrenia: a meta-analysis in over 18 000 subjects. Schizophrenia Bulletin. 2013; 39 (5): 1129–1138. [PubMed: 23042112]

Hamilton JP, Siemer M, Gotlib IH. Amygdala volume in major depressive disorder: a meta-analysis of magnetic resonance imaging studies. Molecular Psychiatry. 2008; 13:993–1000. [PubMed: 18504424]

Hollingstead, AB. Four-factor index of social status. Yale University; New Haven, CT: 1975.

Hulshoff Pol HE, Kahn RS. What happens after the first episode? A review of progressive brain changes in chronically ill patients with schizophrenia. Schizophrenia Bulletin. 2008; 34:354–366. [PubMed: 18283048]

Jovicich J, Czanner S, Han X, Salat D, van der Kouwe A, Quinn B, Pacheco J, Albert M, Killiany R, Blacker D, Maguire P, Rosas D, Makris N, Gollub R, Dale A, Dickerson BC, Fischl B. MRI-derived measurements of human subcortical, ventricular and intracranial brain volumes: Reliability effects of scan sessions, acquisition sequences, data analyses, scanner upgrade, scanner vendors and field strengths. Neuroimage. 2009; 46:177–192. [PubMed: 19233293]

Kay SR, Opler LA, Lindenmayer JP. The Positive and Negative Syndrome Scale (PANSS): rationale and standardisation. British Journal of Psychiatry Suppl. 1989; 7:59–67.

Li M, Chen Z, Deng W, He Z, Wang Q, Jiang L, Ma X, Wang Y, Chua SE, Cheung C, McAlonan GM, Sham PC, Collier DA, Gong Q, Li T. Volume increases in putamen associated with positive symptom reduction in previously drug-naive schizophrenia after 6 weeks antipsychotic treatment. Psychological Medicine. 2011; 42:1475–1483. [PubMed: 22030695]

Morey RA, Petty CM, Xu Y, Hayes JP, Wagner HR 2nd, Lewis DV, LaBar KS, Styner M, McCarthy G. A comparison of automated segmentation and manual tracing for quantifying hippocampal and amygdala volumes. Neuroimage. 2009; 45:855–866. [PubMed: 19162198]

Morey RA, Selgrade ES, Wagner HR 2nd, Huettel SA, Wang L, McCarthy G. Scan-rescan reliability of subcortical brain volumes derived from automated segmentation. Human Brain Mapping. 2010; 31:1751–1762. [PubMed: 20162602]

Oldfield RC. The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia. 1971; 9:97–113. [PubMed: 5146491]

Peterson BS, Riddle MA, Cohen DJ, Katz LD, Smith JC, Leckman JF. Human basal ganglia volume asymmetries on magnetic resonance images. Magnetic Resonance Imaging. 1993; 11:493–498. [PubMed: 8316062]

Sacher J, Neumann J, Funfstuck T, Soliman A, Villringer A, Schroeter ML. Mapping the depressed brain: a meta-analysis of structural and functional alterations in major depressive disorder. Journal of Affective Disorders. 2012; 140:142–148. [PubMed: 21890211]

Shepherd AM, Laurens KR, Matheson SL, Carr VJ, Green MJ. Systematic meta-review and quality assessment of the structural brain alterations in schizophrenia. Neuroscience & Biobehavioral Reviews. 2012; 36:1342–1356. [PubMed: 22244985]

Suckling J, Barnes A, Job D, Brenan D, Lymer K, Dazzan P, Marques TR, MacKay C, McKie S, Williams SR, Williams SC, Lawrie S, Deakin B. Power calculations for multicenter imaging studies controlled by the false discovery rate. Human Brain Mapping. 2010; 31:1183–1195. [PubMed: 20063303]

Tae WS, Kim SS, Lee KU, Nam EC, Kim KW. Validation of hippocampal volumes measured using a manual method and two automated methods (FreeSurfer and IBASPM) in chronic major depressive disorder. Neuroradiology. 2008; 50:569–581. [PubMed: 18414838]

Uttl B. North American Adult Reading Test: age norms, reliability, and validity. Journal of Clinical Experimental Neuropsychology. 2002; 24:1123–1137. [PubMed: 12650237]

van Erp TG, Saleh PA, Huttunen M, Lonnqvist J, Kaprio J, Salonen O, Valanne L, Poutanen VP, Standertskjold-Nordenstam CG, Cannon TD. Hippocampal volumes in schizophrenic twins. Archives of General Psychiatry. 2004; 61:346–353. [PubMed: 15066892]

Van Erp TG, Saleh PA, Rosso IM, Huttunen M, Lonnqvist J, Pirkola T, Salonen O, Valanne L, Poutanen VP, Standertskjold-Nordenstam CG, Cannon TD. Contributions of genetic risk and fetal hypoxia to hippocampal volume in patients with schizophrenia or schizoaffective disorder, their unaffected siblings, and healthy unrelated volunteers. American Journal of Psychiatry. 2002; 159:1514–1520. [PubMed: 12202271]

Wonderlick JS, Ziegler DA, Hosseini-Varnamkhasti P, Locascio JJ, Bakkour A, van der Kouwe A, Triantafyllou C, Corkin S, Dickerson BC. Reliability of MRI-derived cortical and subcortical morphometric measures: effects of pulse sequence, voxel geometry, and parallel imaging. Neuroimage. 2009; 44:1324–1333. [PubMed: 19038349]
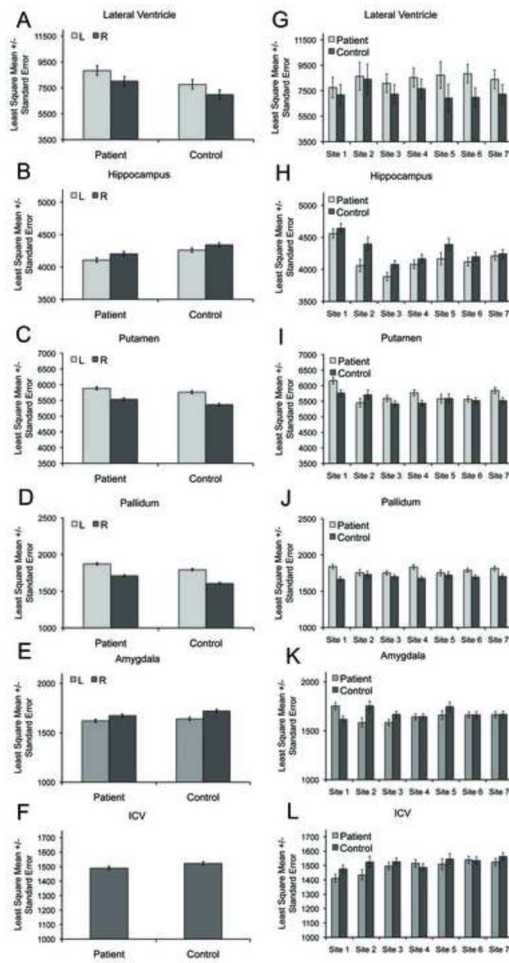
Fig. 1.
Regional volume differences between patients and controls across and within site. 1A-F:
Regional volume differences based on multi-site analyses; 1G-L: Regional volume
differences based on within-site analyses (due to lack of group by hemisphere interactions,
data are collapsed across hemisphere in the within-site figures).

Table 1

Sample Demographics

	Schizophrenia Patients (n=186)	Healthy Volunteers (n=176)	Statistic	p- value
Mean Age (SD)	38.9 (11.6)	37.5 (11.2)	$t_{360}=1.12$	0.26
Subject Education[b] (SD)	3.4(0.9)	2.3(0.9)	$t_{360}=11.47$	<0.0001
Parental Education[b] (SD)	2.4(1.8)	2.1(1.5)	$t_{360}=1.49$	0.14
NAART	29.4(12.4)	39.7(11.4)	$t_{357}=-8.22$	<0.0001
Age at Onset	21.8 (7.6)			
Duration of Illness	17.1 (11.5)			
PANSS positive	15.5(5.1)			
PANSS negative	14.5(5.6)			
PANSS general	28.6(7.5)			
PANSS composite	0.9(6.3)			
Gender (M/F)	145/41	126/50	$\chi^2_1=0.83$	0.36
Handedness[a] (bilateral/left/right)	4/12/170	2/7/167	FET	0.46
Race			FET	0.10
American Indian or Alaskan Native	4	3		
Asian	22	16		
Black or African American	38	20		
Native Hawaiian or Pacific Islander	2	2		
White	120	135		

FET=Fisher's Exact Test

[a] based on Edinburgh Handedness Inventory (Oldfield, 1971)

[b] based on the Hollingstead Socioeconomic Status Scale (Hollingstead, 1975)

NAART = North American Adult Reading Test (Uttl, 2002)

PANSS = Positive and Negative Syndrome Scale (Kay et al., 1989)

Table 2

Absolute Volumes

Region	Schizophrenia Patients (n=186)	Healthy Volunteers (n=176)
Hippocampus		
Left	4106 (516)	4279 (483)
Right	4201 (531)	4348 (462)
Amygdala		
Left	1629 (237)	1655 (213)
Right	1686 (233)	1728 (241)
Caudate		
Left	3702 (540)	3721 (502)
Right	3729 (539)	3756 (545)
Putamen		
Left	5935 (832)	5837 (669)
Right	5596 (810)	5435 (633)
Pallidum		
Left	1883 (250)	1813 (229)
Right	1720 (237)	1625 (206)
Thalamus		
Left	7154 (905)	7272 (876)
Right	7334 (906)	7457 (831)
Lateral Ventricle		
Left	8705 (5539)	7605 (4091)
Right	7920 (4784)	6958 (4034)
Intracranial Volume	1539 (177)	1562 (156)

The absolute volumes are presented as mean mm^3 (SD) except for the Intracranial volume which is presented as mean cc^3 (SD).

Table 3

Univariate mixed-model regression analysis results.

Variable (nDF, dDF)	Hippocampus		Amygdala		Caudate		Putamen		Pallidum		Thalamus		Lateral Ventricle		Intracranial Volume	
	F	p-value	F	p-value	F	p-value	F	p-value	F	p-value	F	p-value	F	p-value	F	p-value
Diagnosis (1,343)	**14.68**	**0.0002**	3.09	0.08	0.53	0.47	**5.98**	**0.02**	**23.85**	**<0.0001**	0.39	0.53	**5.22**	**0.02**	**4.08**	**0.04**
Hemisphere (1,345)	**36.50**	**<0.0001**	**63.08**	**<0.0001**	**10.27**	**0.002**	**460.11**	**<0.0001**	**440.46**	**<0.0001**	**83.84**	**<0.0001**	**41.30**	**<0.0001**		
Site (6,343)	**15.40**	**<0.0001**	1.08	0.37	2.46	0.02	**4.46**	**0.0002**	0.25	0.96	**12.28**	**<0.0001**	0.29	0.87	**3.29**	**0.004**
Diagnosis × Hemisphere (6,345)	0.27	0.60	2.90	0.09	0.15	0.70	1.93	0.17	2.55	0.11	0.01	0.93	0.03	0.87		
Diagnosis × Site (6,343)	0.73	0.63	**2.94**	**0.008**	1.41	0.21	1.95	0.07	1.48	0.18	0.25	0.96	0.23	0.97	0.97	0.45
Hemisphere × Site (6,345)	1.85	0.08	**6.83**	**<0.0001**	1.75	0.11	1.59	0.15	**2.80**	**0.01**	**9.03**	**<0.0001**	1.29	0.26		
Diagnosis × Site × Hemisphere (6,345)	1.18	0.32	0.95	0.46	1.00	0.42	1.93	0.07	0.43	0.83	1.06	0.39	2.09	0.05		
Sex (1,343)	**4.29**	**0.04**	**8.50**	**0.004**	3.39	0.07	**14.18**	**0.0002**	**6.59**	**0.01**	0.01	0.91	0.20	0.65	**119.34**	**<0.0001**
Age (1,343)	**13.16**	**0.0003**	**14.68**	**0.0002**	**18.31**	**<0.0001**	**102.00**	**<0.0001**	**46.87**	**<0.0001**	**41.80**	**<0.0001**	**47.97**	**<0.0001**	**3.84**	**0.05**
Intracranial Volume (1,343)	**93.81**	**<0.0001**	**87.18**	**<0.0001**	**176.24**	**<0.0001**	**69.00**	**<0.0001**	**115.49**	**<0.0001**	**240.40**	**<0.0001**	**37.20**	**<0.0001**		

nDF=nominator degrees of freedom, dDF=denominator degrees of freedom

Quality control procedures resulted in the removal of one patient due to poor scan quality and between 0–2 additional data points for each of the regions of interest.

Effects of diagnosis in **bold** pass the one-tailed (p<0.05) threshold set for the a priori hypotheses. The effects for hippocampus and pallidum also pass the conservative Bonferroni two-tailed p-value threshold of p<0.00625.

Table 4

Least Square Means, Percent Difference, and Effect Sizes

Region	Schizophrenia Patients (n=186)	Healthy Volunteers (n=176)	Percent Difference	Full Sample Effect Size Cohen's d/n80	Weighted Mean Effect Size Cohen's d/n80	Meta-analysis Cohen's d Haijma medicated/ antipsychotic naive	Meta-analysis*** Cohen's d/n80 Shepherd
Pallidum	1792 (193)	1700 (184)	5.44	0.490/53	0.732/24	0.26	1.06[LH] to 1.34[RH]/10–12
Hippocampus	4156 (428)	4302 (409)	−3.41	−0.350/102	−0.391/81	−0.52/−0.43	−0.38 to −0.58/48–110
Putamen	57101 (620)	5566 (592)	2.60	0.239/218	0.443/64	0.10	0.21[LH] to 0.24[RH]/274–357
Lat. Ventricle	8429 (4832)	7378 (4614)	14.25	0.222/252	0.312/127	0.45	0.39[RH] to 0.51[LH]/62–105
ICV	1488 (165)	1523 (163)	−2.26	−0.209/284	−0.182/370	−0.17/−0.14	
Amygdala	1648 (204)	1681 (195)	−1.95	−0.165/455	−0.090/1528	−0.31	−0.38 to −0.72/32–110
Caudate	3755 (483)	3714 (461)	1.08	0.085/1713	0.102/1123	−0.03/−0.38	−0.06[RH] to 0.06[LH]/4362
Thalamus	7287 (675)	7311 (643)	−0.33	−0.037/9033	−0.098/1263	−0.31/−0.68	−0.18 to 0.33/486–146

n80 = number of subjects per group needed, assuming an equal number of subjects per group, for 80% power to detect a difference at p<0.05.

*
= Cohen's *d* of volume difference between medicated schizophrenia patients and healthy controls based on the Haijma et al. (2012) meta-analysis of structural brain abnormalities in schizophrenia

**
= Percent difference and Cohen's *d* of volume difference between chronic schizophrenia patients and healthy controls based on the Shepherd et al. (2012) meta-review of structural brain abnormalities in schizophrenia; LH = for left hemisphere; RH = for right hemisphere Lat. Ventricle = Lateral Ventricle; ICV = Intracranial Volume.