Published in final edited form as:

Biol Psychiatry. 2009 December 1; 66(11): 1055–1060. doi:10.1016/j.biopsych.2009.07.019.

Elucidating a Magnetic Resonance Imaging-Based Neuroanatomic Biomarker for Psychosis: Classification Analysis Using Probabilistic Brain Atlas and Machine Learning Algorithms

Daqiang Sun, Theo G.M. van Erp, Paul M. Thompson, Carrie E. Bearden, Melita Daley, Leila Kushan, Molly E. Hardt, Keith H. Nuechterlein, Arthur W. Toga, and Tyrone D. Cannon

Departments of Psychology (DS, TGMvE, CEB, LK, MEH, KHN, TDC) and of Psychiatry and Biobehavioral Sciences (CEB, MD, KHN, TDC), and the Semel Institute for Neuroscience and Human Behavior (CEB, MD, KHN, TDC), University of California at Los Angeles; and the Laboratory of Neuro Imaging (PMT, AWT), Department of Neurology, University of California at Los Angeles School of Medicine

Abstract

Background—No objective diagnostic biomarkers or laboratory tests have yet been developed for psychotic illness. Magnetic resonance imaging (MRI) studies consistently find significant abnormalities in multiple brain structures in psychotic patients relative to healthy control subjects, but these abnormalities show substantial overlap with anatomic variation that is in the normal range and therefore nondiagnostic. Recently, efforts have been made to discriminate psychotic patients from healthy individuals using machine-learning-based pattern classification methods on MRI data.

Methods—Three-dimensional cortical gray matter density (GMD) maps were generated for 36 patients with recent-onset psychosis and 36 sex- and age-matched control subjects using a cortical pattern matching method. Between-group differences in GMD were evaluated. Second, the sparse multinomial logistic regression classifier included in the Multivariate Pattern Analysis in Python machine-learning package was applied to the cortical GMD maps to discriminate psychotic patients from control subjects.

Results—Patients showed significantly lower GMD, particularly in prefrontal, cingulate, and lateral temporal brain regions. Pattern classification analysis achieved 86.1% accuracy in discriminating patients from controls using leave-one-out cross-validation.

Conclusions—These results suggest that even at the early stage of illness, psychotic patients present distinct patterns of regional cortical gray matter changes that can be discriminated from the normal pattern. These findings indicate that we can detect complex patterns of brain abnormality in early stages of psychotic illness, which has critical implications for early identification and intervention in individuals at ultra-high risk for developing psychosis/schizophrenia.

Keywords

Classification; cortical pattern matching; MRI; psychosis; PyMVPA; schizophrenia

© 2009 Society of Biological Psychiatry

Address correspondence to Tyrone D. Cannon, Ph.D., Department of Psychology, 1285 Franz Hall, University of California, Los Angeles, Los Angeles, CA 90095; cannon@psych.ucla.edu.

The authors report no biomedical financial interests or potential conflicts of interest.

Supplementary material cited in this article is available online.

Schizophrenia and other psychotic disorders are severe mental illnesses that cause tremendous human suffering and economic burden. These diseases are diagnosed on the basis of clinical evaluation of symptoms and functional impairment, and no objective diagnostic biomarkers or laboratory tests have yet been developed. Patients show significant group differences, relative to healthy control subjects, on a range of neurobiological and cognitive measures, but at the individual level, these measures show extensive overlap with the normal range and are therefore nondiagnostic. Recently, however, efforts have been made to discriminate patients from healthy individuals by applying multivariate analysis to various data sets of combined measures (1–3).

It is well established that schizophrenia and other psychoses are associated with brain structural abnormalities. Neuroimaging studies have found brain volume decreases in schizophrenia patients compared with healthy controls, with the most pronounced abnormalities in the hippocampus, superior temporal gyrus, prefrontal lobe, and cingulate gyrus (4,5). Thus far, most structural magnetic resonance imaging magnetic resonance imaging (MRI) studies of schizophrenia/psychosis have adopted a region-of-interest analysis or a mass-univariate method, in which statistics are performed on each region/voxel independently. These approaches are advantageous in identifying region-specific differences between groups (6) but can miss signals from different regions/voxels that are spatially correlated and tend to occur together. Further, because there is usually extensive group overlap even for the regions/voxels showing the most pronounced differences, univariate methods are unlikely to satisfactorily discriminate patients from healthy control subjects at the individual subject level.

Multivariate classification methods have been applied to neuroimaging data to address these limitations. Rather than treating each voxel independently from the others, multivariate analysis of image data seeks the best group classification by taking into account multiple voxels simultaneously. Machine-learning-based multivariate methods have been successfully used to differentiate task-specific brain activity in functional MRI studies (7,8), to segment brain structures (9), and to classify with high accuracy individuals in the prodromal phase of Alzheimer's disease (10).

Recently, these methods have been applied to brain structural images from schizophrenia patients and have resulted in a high accuracy of classification (81.1%) from images of healthy subjects (2). In a follow-up report by the same group, the pattern recognized in schizophrenia-control classification was applied to unaffected family members of schizophrenia patients, and a similar-to-patient pattern was found in family members (11). Instead of using voxel data as predictors, composite measures including principal components have also been used as features for machine learning, and reasonable classifications have been accomplished using linear discriminant analysis (12) and support vector machines (13).

In this study, we used the cortical pattern matching method (14) to assess cortical gray matter abnormalities in patients with recent-onset psychosis. We further applied the PyMVPA (Multivariate Pattern Analysis in Python, <http://www.pymvpa.org>) machine-learning package (15,16) to gray matter density (GMD) maps to determine whether group classification at the individual level could be achieved in the early course of psychosis. Cortical pattern matching is an advanced brain registration technique that can achieve accurate anatomical correspondence between brain surfaces. It has proved successful in detecting subtle cortical gray matter changes in normal development and aging and pathologic conditions (14). PyMVPA is a recently developed Python module that integrates and implements multiple machine-learning algorithms for pattern classification suitable for

neuroimaging data (15). We applied the sparse multinomial logistic regression (SMLR) classifier (17). It has several advantages—computational efficiency, good classification performance, and the intrinsic feature selection process that makes a separate signal filtering unnecessary.

We hypothesized that cortical gray matter deficits would be present in psychotic patients in the prefrontal, cingulate, and superior temporal regions, on the basis of the convergence of previous findings. Further, we expected that application of multivariate classification based on cortical gray matter patterns would be able to discriminate reliably between recent-onset psychosis patients and healthy control subjects.

Methods and Materials

Participants

Participants were recruited from the Aftercare Research Program and Adolescent Brain-Behavior Research Clinic (ABBRC) at the University of California—Los Angeles. Criteria for identification of subjects with recent-onset psychosis included onset of first-episode psychosis within the past 2 years and a DSM-IV diagnosis of a psychotic disorder. For patients, the baseline clinical assessment interview included a structured diagnostic interview (18) and a review of medical records and collateral information available from family and care providers that contributed to a consensus diagnosis of psychotic disorder. Detailed information regarding reliability and consensus diagnosis procedures are described elsewhere (19,20).

Healthy control subjects were recruited from the same communities as the patients through local advertising. They did not meet DSM-IV criteria for a psychiatric disorder, as determined by direct Structured Clinical Interview for DSM interview, and did not have a first-degree family history of a psychotic disorder or meet criteria for a prodromal state, as determined by Structured Interview for Prodromal Syndromes (SIPS) assessment with the patients and (for minors) their parents or legal guardians.

Additional exclusion criteria for all participants included the presence of a neurologic disorder, drug or alcohol abuse or dependence within the past 6 months, insufficient English fluency, and IQ below 70.

Clinical interviewers underwent a rigorous training protocol to ensure reliability on clinical measures. Certification requires that interviewers achieve “good” to “excellent” reliability (interclass correlations over .75 for all symptom ratings, and kappas of at least .80 for prodromal syndrome diagnoses) on SIPS training videotapes, compared with gold standard ratings developed at Yale University (21,22). Diagnostic reliability on the SCID-I/P (Patient Edition) requires independent rating of a minimum of six SCID-I/P training videos, followed by four live supervised assessments. Interrater reliability was calculated by comparing trainee ratings with gold standard ratings, as described in Ventura *et al.* (20); trainees are required to achieve a sensitivity kappa of .75 or greater and a specificity kappa of at least .75 regarding SCID-I/P symptoms, as well as 90% agreement on all diagnostic classifications across SCID-I/P assessments.

Thirty-six recent-onset psychosis patients (including four ultra-high-risk individuals who converted to psychosis within 12 months from the scan time) and 36 sex- and age-matched healthy control subjects were included in the analysis. Although most of the patients in the sample had a diagnosis of schizophrenia spectrum disorder (schizophrenia, schizophreniform, or schizoaffective disorder), to maximize the sample size, we also included a small number of patients with affective psychosis and atypical psychosis. There

were 13 females and 23 males in each group, and the mean ($\pm$ SD) ages were 19.0 ± 5.2 years and $19.0 \text{ years} \pm 4.8 \text{ years}$ for patients and control subjects, respectively. The IQ scores were 104.1 ± 17.1 and 109.2 ± 23.6 , and years of education were 11.2 ± 2.6 years and $11.8 \text{ years} \pm 3.0 \text{ years}$, for patients and control subjects, respectively. One patient and three control subjects were left-handed among all subjects. There were no significant differences in age, sex, handedness, or years of education (all $ps > .20$). The diagnostic breakdown of psychotic patients was schizophrenia ($n = 21$), schizophreniform disorder (6), schizoaffective disorder (5), bipolar disorder with psychotic features (2), major depression with mood-incongruent psychotic symptoms (1), and psychosis not otherwise specified (1). For 31 patients whose medication records were available, atypical anti-psychotic medications had been taken by 16 cases, mood stabilizers by 6 cases, and antidepressants by 9 cases, before MRI scans. Four patients had taken none of these medications before scans. The study was approved by Internal Review Board at the University of California—Los Angeles. All subjects signed informed consent/assent documents after the study procedures were fully explained.

Image Processing and Analysis

All participants were scanned on a 1.5-T Siemens Sonata MRI scanner. A three-dimensional (3D) magnetization prepared rapid gradient echo (MPRAGE) sequence generated 160 contiguous, 1.0-mm sagittal slices. Imaging parameters were echo time/repetition time = 4.38 msec/1900 msec; flip angle = 15° ; field of view = 256 mm^2 ; voxel dimension = 1 mm^3 .

For each scan, a radio frequency bias field correction was performed (23). The cerebrum was extracted from the remainder of the head in the image and was divided into left and right hemispheric images. These images were edited manually, keeping cerebral voxels and removing nonbrain tissue voxels. Automated tissue segmentation was performed on each scan to classify the image into gray matter, white matter, and cerebrospinal fluid (24). The gray matter image was retained for further analysis. The hemispheric images were registered to a standard 3D stereotaxic space (25) with nine degree-of-freedom linear transformations (26). Cortical pattern matching (14) was performed in the standard space, as described below.

A cortical surface extraction was performed to generate both hemispheric surface models for each brain (27). In this process, a spherical mesh surface was continuously deformed to fit the cortical surface that best differentiated brain tissue and cortical cerebrospinal fluid, and a high-resolution surface model representing 65,536 brain surface points was created. On each hemispheric surface, 29 anatomic landmark curves following major sulci and seven control curves that delineate the lateral surface and the medial surface were manually traced by image analysts blind to subject demographics and diagnosis. The tracing protocol is available on the Internet (<http://www.loni.ucla.edu/~esowell/edevl/MedialLinesProtocol.htm>). The reliability of tracing was tested on six standard brain surfaces, and the average distance between the sulcal curves tracing by the analysts and the standard sulcal curves was less than 2 mm in most regions.

The hemispheric surfaces and curves were then flattened to a two-dimensional plane, and average curves were generated by averaging the positions of the same curves across all subjects. The hemispheric surfaces were elastically warped to each other on the basis of matching individual curves to their corresponding average curves, and the coordinate positions of each surface point in their 3D space were preserved. The 3D surface models were reconstructed in the standard space and were transformed back to each individual scan's native space for GMD sampling.

In each scan's native space where the segmented gray matter image resided, local GMD was calculated and assigned to each point on the hemispheric surface models. This was done by creating a sphere of 15-mm radius around each surface point and calculating the proportion of gray matter volume within the sphere. The obtained GMD maps were transformed to the standard space for group comparison and classification.

Statistical Analysis

Student's *t* tests were performed to compute differences in mean GMD between patients and control subjects at each brain surface point, and uncorrected P-maps were generated for visualization. Following this, standard permutation tests were conducted to confirm the significance of the overall pattern of differences. We conducted 100,000 randomized permutations (randomly assigning subjects to either patient or control groups, while keeping the total number of subjects per group the same) for the whole hemispheric surfaces. In each permutation, the number of suprathreshold surface voxels for which $p < .01$ was computed, and the null distribution of this statistic was determined from all permutations. Between-group differences were considered significant if less than 5% ($p < .05$) of the results from all random permutations exceeded the actual result. This approach has been used in many prior studies and is comparable to an approach called set-level inference in functional brain imaging (28). In other words, the fraction of the surface that exceeds a predefined fixed threshold is estimated both on real data and in random simulations, and the fraction of the random simulations that beat the real effect is considered the *p* value or likelihood that the observed pattern of suprathreshold effects occurred by chance.

Classification

PyMVPA was applied to GMD maps for group classification (15). Basically, all coregistered GMD maps were converted to one-dimensional arrays and were combined to form a two-dimensional (GMD $\times$ subject) array. Binary group labels (0 for control subjects, and one for patients) were assigned to each subject. The SMLR classifier (17) was applied to the data set, with the lambda penalty set to the default value ($\text{lm} = .1$). We used the leave-one-out cross-validation method to determine the accuracy of classification. All subjects except one were chosen as the training data set, and a decision surface that best separated the two groups was computed by the SMLR algorithm. The decision surface was then applied to the left-out subject (test data set) to predict into which group he or she fell. Iteratively, the leave-one-out process was applied to each subject, and the accuracy of all predictions was calculated.

To evaluate the significance of the acquired accuracy value from the described cross-validation, 1000 permutations of leave-one-out cross-validations were conducted. In each permutation, the group labels (36 as "0," and 36 as "1") were randomly assigned to all subjects, and a leave-one-out process was performed to determine the prediction accuracy for this randomly permuted experimental data set. The accuracy values obtained from all permutations formed a null distribution, and the test was considered significant if less than 5% ($p < .05$) of the results from all permutations exceeded the actual result. Using this null distribution, it is possible to tell how likely it is to observe a certain classification accuracy purely by chance, and conversely it is possible to assign a significance level to the classification accuracy. This would not be possible without knowing the variance in the classification accuracy under the null hypothesis, which the randomization process simulates.

Sparse Multinomial Logistic Regression Classifier

Briefly, the SMLR classifier belongs to a set of machine-learning approaches that build a classification function from a weighted combination of basis functions, in which the weights

are tuned during the learning phase to produce an optimal classification of the training data (17). Sparse means that the weight estimates are encouraged during the learning process to be either high or exactly zero, to make the model more parsimonious, efficient to run, to avoid overfitting, and to improve generalization capacity. To do this, a regularization or penalty function is included during the learning process to promote a sparse model. Multinomial logistic regression is one approach to the multiclass classification problem, in which the features in the images are used as predictor variables, and the output is a categorization (here there are two classes). Basically, for each category, a multiple regression is run to predict the odds ratio that the image falls into that category. The odds ratios are then converted into probabilities, and the most likely class is chosen.

SMLR uses a general linear model (multiple regression) to predict the logarithm of the odds ratio of an image belonging to a specific class. Then, the odds ratio is converted into a probability using a nonlinear transfer function (sum of exponentials), which ensures that all the classification probabilities sum to one across all classes. To see this, we can assume that the training set consists of n examples of data falling into k categories with p explanatory variables. Then, we let π_j be the multinomial probability of an observation falling into the j th category. Then, we run a multiple regression on the p features to predict the log of the odds ratio that the training example belongs to the j th versus the k th category (29):

$$\log \left(\frac{\pi_j(x_i)}{\pi_k(x_i)} \right) = \beta_{0j} + \beta_{1j} x_{1i} + \beta_{2j} x_{2i} + \cdots + \beta_{pj} x_{pi}; \quad \begin{matrix} j=1, 2, \dots, (k-1) \\ i=1, 2, \dots, n. \end{matrix}$$

Since all the π 's add to unity, this reduce to

$$\log(\pi_j(x_i)) = \frac{\exp(\beta_{0j} + \beta_{1j} x_{1i} + \beta_{2j} x_{2i} + \cdots + \beta_{pj} x_{pi})}{1 + \sum_{j=1}^{k-1} \exp(\beta_{0j} + \beta_{1j} x_{1i} + \beta_{2j} x_{2i} + \cdots + \beta_{pj} x_{pi})}$$

Results

Recent-onset psychosis patients showed significantly lower cortical GMD compared with sex- and age-matched healthy control subjects ($p < .0002$ in 100,000 permutations). As shown in Figure 1, differences were most pronounced in the lateral surface of the prefrontal and temporal lobes, limbic regions along the cingulate sulci, and areas along the parieto-occipital fissures. The primary sensorimotor cortex was relatively spared, as well as the primary visual cortex on the medial surface.

The PyMVPA classification analysis, using leave-one-out cross-validation, accurately classified 86.1% of the patients and control subjects. That is, in both groups, 31 of 36 subjects were correctly assigned to their actual groups, when applying the learned patterns from all other subjects to each individual subject, which gave an accuracy of 86.1% for both groups. Among 1000 permutations of leave-one-out cross-validations, none gave accuracy values higher than 86.1%, ensuring that the classification was statistically significant ($p < .001$). The accuracy of classification was 84.4% when the four ultra-high-risk converter subjects and four sex- and age-matched control cases were excluded from the analysis (78.1% for patients, and 90.6% for control subjects). When patients with affective psychosis and psychosis not otherwise specified ($n = 4$) and four sex- and age-matched control subjects were excluded, the accuracy of classification was 87.5% (90.6% for patients and 84.4% for control subjects).

One hundred twenty-nine feature surface voxels were selected by the SMLR process and were linearly combined for the classification (30). The 25% with the highest weights, which constituted 18 clusters and accounted for 59.9% of the total weights, were plotted on the 3D surface to identify the regions that contributed the most to the classification (Figure 2). The regions containing surface points with highest weights included the frontal pole, superior and middle temporal regions on the left hemisphere, and the superior temporal, somatomotor, and subgenual regions on the right hemisphere.

Permutation tests by keeping the original sex ratios for both groups did not change the result for either between-group comparison or classification.

Discussion

Here we applied a unique combination of imaging analysis methods to a sample of recent-onset psychosis patients and well-matched healthy subjects, revealing significant local gray matter deficits that were most pronounced in prefrontal, cingulate, and lateral temporal regions in patients. Using a machine-learning algorithm (PyMVPA), we achieved highly accurate (86.1%) brain-image-based group classification.

Cortical pattern matching, a surface-based registration method using reliable, manually selected anatomic landmarks and a sophisticated warping technique, ensured that interindividual brain anatomic variability was accounted for as far as possible. SMLR, a novel classifier provided by the PyMVPA package, was used to assign individuals into patient or control groups. Because the SMLR algorithm has feature selection as an intrinsic process, the whole gray matter density maps can be used as the input data, making feature reduction such as principal component analysis or analysis of variance–based filtering unnecessary and avoiding the loss of potentially relevant information.

The prefrontal, cingulate, and lateral temporal regions showed significant gray matter differences between patients and control subjects, consistent with previous findings on schizophrenia/psychosis (4,5). These cortical regions play important roles in working memory, executive function, and auditory sensation and language processing, all of which are impaired in schizophrenic and psychotic patients. These gray matter deficits may therefore underlie such functional abnormalities. The primary sensorimotor cortices and primary visual cortex were relatively spared, probably indicating that these primary cortices are less severely involved in the underlying pathophysiology of psychosis compared with higher order association cortices.

The 86.1% classification accuracy is comparable to the best-performing classifications in prior studies of schizophrenia. This result suggests that even at the early stage of illness, psychotic patients demonstrate some distinct patterns of cortical gray matter differences that distinguish them from healthy individuals. Each individual feature only accounts for a small portion of the classification, and not all selected feature surface points were within the regions of significant difference. The latter pattern is consistent with previous reports (10) and indicates that some discriminative signals could be missed if only features showing significant between-group differences were searched. It is also widely acknowledged in the machine-learning literature that “weak learners,” or classifiers whose performance alone is only slightly better than chance, can be combined to produce a powerful classifier (9). The functional implications of and the interplay among these “discriminative” regions can be the targets of further investigations.

It is interesting to consider how cortical pattern matching deals with sulcal variability and, more specifically, whether greater sulcal variability among patients could contribute to the classification results. It has been noted before that patients may have slightly higher

geometric variation in their sulcal landmarks, when data from multiple subjects are mapped to a standard stereotaxic space. This may be true in schizophrenia and psychosis, although the evidence for greater cortical pattern variation in disease is stronger in other disorders such as Alzheimer's disease (31,32).

If patients are more variable in cortical structure than control subjects, and if a standard registration approach were used without modeling the cortex, then the cross-subject registration errors in cortical anatomy can seriously reduce the power to detect group differences and disease effects. To alleviate this loss of power due to data misregistration across subjects at the cortex, cortical pattern matching matches as far as possible major sulcal landmarks traced by trained image analysts. This process also matches the entire surface model of the cortex across subjects and also performs a higher order matching of intervening surface areas between the major sulci, using advanced mathematical models based on covariant partial differential equations and continuum mechanics. Because cortical anatomic regions are generally defined by these major sulci, the cortical regions are therefore accurately matched. It may be worth noting that volume-based nonlinear registration algorithms cannot achieve highly accurate brain surface matching, as suggested by a recent large-scale comparison study (33).

In situations in which cortical patterns between brains are topologically different, a one-to-one match may not exist at the gross anatomic level. A well-studied example is that the paracingulate sulcus is less frequently present in schizophrenic patients. In such cases, the best conceivable solution is to match accurately what can be matched, and treat that which cannot be matched as unmodeled residual variation. The lack of a sulcus would be likely to be associated with fewer gray matter in that area, which can be perfectly detected by cortical pattern matching using the GMD measure.

In terms of classification, greater sulcal variability, if not controlled for by cortical pattern matching, would reduce the power of a classifier. The reason for this is that the unmodeled cortical variance would tend to reduce the anatomic homology of the gray matter features fed into the classifier as training data, making classification more difficult. Cortical matching was performed explicitly to reduce this source of error.

The limitations of this study include a relatively modest sample size and the heterogeneous psychotic disorder diagnoses in the patient group. Images were visually inspected for artifacts, and those with observable artifacts were excluded from the analyses. We cannot fully exclude the possibility that greater motion in the patient group not resulting in observable artifacts may have resulted in artifactual appearance of thinner cortex (34), but in simulations in which potential motion-related cortical thinning was assumed to be global, we showed that classification accuracies were not severely affected (Supplement 1). The accuracy value was derived using a leave-one-out cross-validation, thereby avoiding inflated accuracy due to data overfitting, but a new data set might be preferable in some respects as a test data set. Antipsychotics and other medications had been administered to most patients before MRI scans, and therefore we cannot rule out the possibility that medication effects may contribute to the observed differences. It can be speculated, however, that the medication effect is less of a confounding factor for the purposes of classification, because the types varied considerably across patients, and there is unlikely to be a common brain change pattern related to such a diversity of medications.

Although we focused on cortical gray matter here, classification analysis is not restricted to this particular measure. Measures from other brain structures, other imaging modalities including diffusion tensor imaging and functional MRI, electrophysiology measures, and cognitive and clinical variables can all be integrated into a classification analysis, and

improved accuracy could potentially be achieved. Classifiers other than SMLR or combinations of different classifiers can be used to improve classification accuracy as well. Boosting in particular offers one such approach, because it combines many features whose classification performance on their own is only marginally better than chance (i.e., weak learners). Even so, the advantage of boosting is greatest when all the available features are weak, and not much improvement is achieved if strong classifier is already available (35). At the same time, the current methods can be applied in future studies to individuals at ultra-high risk for developing psychosis or schizophrenia. Patterns derived from recent-onset psychotic patients can be applied to ultra-high-risk cases for outcome prediction, and ultimately, patterns from individuals who later converted to psychosis can be used to predict conversion for future cases. An image-based tool is promising for diagnostic use in standard clinical settings and for prediction of outcome among clinically at risk individuals.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was supported by the following grants: National Institute of Mental Health (NIMH) Grant No. MH65079 (to TDC) and NIMH Grant No. P50 MH066286 (to KN), NIMH Grant No. MH037705 (to KN), National Alliance for Research on Schizophrenia and Depression Young Investigator Award (to CEB), Grant No. K23MH079028 (to MD), Grant No. EB008432 (to PMT), Grant No. EB007813 (to PMT), Grant No. EB008281 (to PMT), Grant No. HD050735 (to PMT), Grant No. AG020098 (to PMT), National Institutes of Health (NIH)/National Center for Research Resources Grant No. P-41 (to AWT), and NIH Grant No. U54 RR021813 (UCLA Center for Computational Biology), as well as by donations from the Rutherford Charitable Foundation and Staglin Music Festival for Mental Health to the University of California-Los Angeles Foundation.

References

1. Caprihan A, Pearlson GD, Calhoun VD. Application of principal component analysis to distinguish patients with schizophrenia from healthy controls based on fractional anisotropy measurements. *Neuroimage*. 2008; 42:675–682. [PubMed: 18571937]
2. Davatzikos C, Shen D, Gur RC, Wu X, Liu D, Fan Y, et al. Whole-brain morphometric study of schizophrenia revealing a spatially complex set of focal abnormalities. *Arch Gen Psychiatry*. 2005; 62:1218–1227. [PubMed: 16275809]
3. Michael AM, Calhoun VD, Andreasen NC, Baum SA. A method to classify schizophrenia using inter-task spatial correlations of functional brain images. *Conf Prog IEEE Eng Med Biol Soc*. 2008; 2008:5510–5513.
4. Shenton ME, Dickey CC, Frumin M, McCarley RW. A review of MRI findings in schizophrenia. *Schizophr Res*. 2001; 49:1–52. [PubMed: 11343862]
5. Pantelis C, Yucel M, Wood SJ, Velakoulis D, Sun D, Berger G, et al. Structural brain imaging evidence for multiple pathological processes at different stages of brain development in schizophrenia. *Schizophr Bull*. 2005; 31:672–696. [PubMed: 16020551]
6. Friston KJ, Ashburner J. Generative and recognition models for neuroanatomy. *Neuroimage*. 2004; 23:21–24. [PubMed: 15325348]
7. Haxby JV, Gobbini MI, Furey ML, Ishai A, Schouten JL, Pietrini P. Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*. 2001; 293:2425–2430. [PubMed: 11577229]
8. Kay KN, Naselaris T, Prenger RJ, Gallant JL. Identifying natural images from human brain activity. *Nature*. 2008; 452:352–355. [PubMed: 18322462]
9. Morra JH, Tu Z, Apostolova LG, Green AE, Toga AW, Thompson PM. Automatic subcortical segmentation using a contextual model. *Med Image Comput Comput Assist Interv Int Conference*. 2008; 11:194–201.

10. Fan Y, Resnick SM, Wu X, Davatzikos C. Structural and functional biomarkers of prodromal Alzheimer's disease: A high-dimensional pattern classification study. *Neuroimage*. 2008; 41:277–285. [PubMed: 18400519]
11. Fan Y, Gur RE, Gur RC, Wu X, Shen D, Calkins ME, et al. Unaffected family members and schizophrenia patients share brain structure patterns: A high-dimensional pattern classification study. *Biol Psychiatry*. 2008; 63:118–124. [PubMed: 17555720]
12. Kawasaki Y, Suzuki M, Kherif F, Takahashi T, Zhou SY, Nakamura K, et al. Multivariate voxel-based morphometry successfully differentiates schizophrenia patients from healthy controls. *Neuroimage*. 2007; 34:235–242. [PubMed: 17045492]
13. Yoon U, Lee JM, Im K, Shin YW, Cho BH, Kim IY, et al. Pattern classification using principal components of cortical thickness and its discriminative pattern in schizophrenia. *Neuroimage*. 2007; 34:1405–1415. [PubMed: 17188902]
14. Thompson PM, Hayashi KM, Sowell ER, Gogtay N, Giedd JN, Rapoport JL, et al. Mapping cortical change in Alzheimer's disease, brain development, and schizophrenia. *Neuroimage*. 2004; 23(suppl 1):S2–18. [PubMed: 15501091]
15. Hanke M, Halchenko YO, Sederberg PB, Olivetti E, Frund I, Rieger JW, et al. PyMVPA: A unifying approach to the analysis of neuroscientific data. *Front Neuroinformatics*. 2009; 3:3.
16. Hanke M, Halchenko YO, Sederberg PB, Hanson SJ, Haxby JV, Pollmann S. PyMVPA: A python toolbox for multivariate pattern analysis of fMRI data. *Neuroinformatics*. 2009; 7:37–53. [PubMed: 19184561]
17. Krishnapuram B, Carin L, Figueiredo MA, Hartemink AJ. Sparse Multinomial Logistic Regression: Fast Algorithms and Generalization Bounds. *Trans IEEE Pattern Anal Mach Intell*. 2005; 27:957–968.
18. First, MB.; Spitzer, RL.; Gibbon, M.; Williams, JB. Structured Clinical Interview for DSM-IV Axis I Disorders. Washington, DC: American Psychiatric Publishing; 1997.
19. Meyer SE, Bearden CE, Lux SR, Gordon JL, Johnson JK, O'Brien MP, et al. The psychosis prodrome in adolescent patients viewed through the lens of DSM-IV. *J Child Adolesc Psychopharmacol*. 2005; 15:434–451. [PubMed: 16092909]
20. Ventura J, Liberman RP, Green MF, Shaner A, Mintz J. Training and quality assurance with the structured clinical interview for DSM-IV (SCID-I/P). *Psychiatry Res*. 1998; 79:163–173. [PubMed: 9705054]
21. McGlashan TH, Miller TJ, Woods SW. Pre-onset detection and intervention research in schizophrenia psychoses: Current estimates of benefit and risk. *Schizophr Bull*. 2001; 27:563–570. [PubMed: 11824483]
22. Miller TJ, McGlashan TH, Rosen JL, Somjee L, Markovich PJ, Stein K, et al. Prospective diagnosis of the initial prodrome for schizophrenia based on the structured interview for prodromal syndromes: Preliminary evidence of interrater reliability and predictive validity. *Am J Psychiatry*. 2002; 159:863–865. [PubMed: 11986145]
23. Sled JG, Zijdenbos AP, Evans AC. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Trans Med Imaging*. 1998; 17:87–97. [PubMed: 9617910]
24. Zhang Y, Brady M, Smith S. Segmentation of brain MR images through a hidden markov random field model and the expectation maximization algorithm. *IEEE Trans Med Imaging*. 2001; 20:45–57. [PubMed: 11293691]
25. Mazziotta JC, Toga AW, Evans A, Fox P, Lancaster J. A probabilistic atlas of the human brain: Theory and rationale for its development. The International Consortium for Brain Mapping (ICBM). *Neuroimage*. 1995; 2:89–101. [PubMed: 9343592]
26. Collins DL, Neelin P, Peters TM, Evans AC. Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. *J Comput Assist Tomogr*. 1994; 18:192–205. [PubMed: 8126267]
27. MacDonald JD, Kabani N, Avis D, Evans AC. Automated 3-D extraction of inner and outer surfaces of cerebral cortex from MRI. *Neuroimage*. 2000; 12:340–356. [PubMed: 10944416]
28. Friston KJ, Holmes A, Poline JB, Price CJ, Frith CD. Detecting activations in PET and fMRI: Levels of inference and power. *Neuroimage*. 1996; 4:223–235. [PubMed: 9345513]

29. Chatterjee, S.; Hadi, AS. Regression Analysis by Example. 4. Hoboken, NJ: Wiley; 2006. p. 317-340.
30. Ferrarini L, Palm WM, Olofsen H, van der Landen R, van Buchem MA, Reiber JH, et al. Ventricular shape biomarkers for Alzheimer's disease in clinical MR images. *Magn Reson Med*. 2008; 59:260–267. [PubMed: 18228600]
31. Narr K, Thompson P, Sharma T, Moussai J, Zoumalan C, Rayman J, et al. Three-dimensional mapping of gyral shape and cortical surface asymmetries in schizophrenia: Gender effects. *Am J Psychiatry*. 2001; 158:244–255. [PubMed: 11156807]
32. Thompson PM, Moussai J, Zohoori S, Goldkorn A, Khan AA, Mega MS, et al. Cortical variability and asymmetry in normal aging and Alzheimer's disease. *Cereb Cortex*. 1998; 8:492–509. [PubMed: 9758213]
33. Klein A, Andersson J, Ardekani BA, Ashburner J, Avants B, Chiang MC, et al. Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. *Neuroimage*. 2009; 46:786–802. [PubMed: 19195496]
34. Blumenthal JD, Zijdenbos A, Molloy E, Giedd JN. Motion artifact in magnetic resonance imaging: Implications for automated analysis. *Neuroimage*. 2002; 16:89–92. [PubMed: 11969320]
35. Martinez-Ramon M, Koltchinskii V, Heileman GL, Posse S. fMRI pattern classification using neuroanatomically constrained boosting. *Neuroimage*. 2006; 31:1129–1141. [PubMed: 16529955]

Figure 1.

Maps of average gray matter density (GMD) and P-maps of recent-onset psychosis patients versus healthy control subjects comparison. Patients showed significantly lower GMD, particularly in the lateral surface of the prefrontal and temporal lobes, limbic regions along the cingulate sulci, and areas along the parieto-occipital fissures. The primary sensorimotor cortex and primary visual cortex on the medial surface was relatively spared.

Figure 2. Sparse multinomial logistic regression weights for patient-control classification. The 25% of voxels with highest weights, within 3-mm kernels, were plotted, forming 18 separate clusters. The regions containing surface points with highest weights for group classification included the frontal pole; superior and middle temporal regions on the left hemisphere; and the superior temporal, somatomotor, and subgenual regions on the right hemisphere.