A Voxel-Based Method for the Statistical Analysis of Gray and White Matter Density Applied to Schizophrenia

I. C. Wright,* 1 P. K. McGuire,* 1 J.-B. Poline,* 1 J. M. Travers,* 1 R. M. Murray,* C. D. Frith,†$$ R. S. J. Frackowiak,* 1 and K. J. Friston†

*Department of Psychological Medicine, Institute of Psychiatry, London, SE5 8AF, United Kingdom; †Wellcome Department of Cognitive Neurology, c/o MRC Cyclotron Unit, Hammersmith Hospital, London, United Kingdom; $$Cyceron PET Research Center, Caen, France; and $Department of Psychology, University College London, London, United Kingdom

Received April 17, 1995

We describe a novel technique for characterizing regional cerebral gray and white matter differences in structural magnetic resonance images by the application of methods derived from functional imaging. The technique involves automatic scalp-editing of images followed by segmentation, smoothing, and spatial normalization to a symmetrical template brain in stereotactic Talairach space. The basic idea is (i) to convert structural magnetic resonance image data into spatially normalized images of gray (or white) matter density, effected by segmenting the images and smoothing, and then (ii) to use Statistical Parametric Mapping to make inferences about the relationship between gray (or white) matter density and symptoms (or other pathophysiological measures) in a regionally specific fashion. Because the whole brain sum of gray (or white) matter indices is treated as a confound, the analysis reduces to a characterization of relative gray (or white) matter density on a voxel by voxel basis. We suggest that this is a powerful approach to voxel-based statistical anatomy. Using the technique, we constructed maps of the regional cerebral gray and white matter density correlates of syndrome scores (distinct psychotic symptoms) in a group of 15 schizophrenic patients. There was a negative correlation between the score for the reality distortion syndrome and regional gray matter density in the left superior temporal lobe ($P = 0.01$) and regional white matter density in the corpus callosum ($P < 0.001$). These abnormalities may be associated with functional changes predisposing to auditory hallucinations and delusions. This method permits the detection of structural differences within the entire brain (as opposed to selected regions of interest) and may be of value in the investigation of structural gray and white matter abnormalities in a variety of brain diseases.

INTRODUCTION

Structural studies of the human brain must overcome the problems posed by individual variability in brain size, shape, and composition (such as relative gray and white matter density). Correction for interindividual differences in brain size and shape is one technique that may be used to improve detection of regional and global differences in cerebral structure. Various methods have been described which include linear transformation of brain images into a “bounding box” of constant size (Andreasen et al., 1994) or registration of images with a template image in stereotactic space (Collins et al., 1994). Techniques which operate on the whole brain can generally be automated and are faster than more traditional methods of analysis (which involve manual measurements of regions of interest). This consideration is particularly important for the analysis of more modern magnetic resonance image (MRI) scans with thin slices for which the data sets are very large.

It has been suggested that schizophrenia is associated with larger ventricles (Van Horn and McManus, 1992) and with both widespread (Zipursky et al., 1992; Harvey et al., 1993) and localized changes (Shenton et al., 1992) in cerebral structure. Therefore methods which correct for interindividual variability in brain size and shape and which can examine both gray and white matter would be of particular value in the investigation of schizophrenia.

In functional imaging studies, methods have been developed to spatially transform images into a standard stereotactic space prior to intersubject averaging with Statistical Parametric Mapping (SPM; Friston et al., 1995a). The spatial normalization utilizes nonlinear spatial deformations which may be employed to match a subject image with a template image. The spatial transformation is carried out simultaneously with transformation of voxel (intensity) values using a least-squares solution and a series of linearizing devices.
The technique is noninteractive (automatic), non-linear, and noniterative. This paper describes the application of these methods to structural MRIs, permitting the structural mapping of regional cerebral gray and white matter within a standard stereotactic space (Talairach and Tournoux, 1988). As an example of the technique, we have examined the structural correlates of regional gray and white matter density with syndrome scores in a group of patients with schizophrenia.

METHODS

Subjects

Fifteen right-handed (as assessed by the scale of Annett, 1970) male patients with DSM-IV schizophrenia were recruited from the Maudsley Hospital. All were in remission at the time of the study. The mean age of the subjects was 32 years (range 24 to 44 years), the mean duration of illness was 9 years, the mean premorbid IQ was 107 (as assessed by the NART; Nelson and O’Connell, 1978), and the mean current dose of neuroleptic medication was 564 chlorpromazine equivalents. Subjects gave written informed consent to participate in the project, which was approved by the local hospital ethical committee.

Clinical Assessment

Schizophrenic symptoms present during the 3 months prior to scanning were assessed by psychiatric interview using the Scale for Assessment of Positive Symptoms (SAPS) and the Scale for Assessment of Negative Symptoms (SANS; Andreasen, 1987). Liddle’s three schizophrenic syndrome scores (Liddle et al., 1992) were calculated from SAPS and SANS scores using a formula (Liddle, personal communication; Liddle, 1987); the reality distortion score was calculated as the sum of hallucinations (SAPS subscale global score) and delusions (SAPS subscale global score); the disorganization score was calculated as the sum of positive formal thought disorder (SAPS subscale global score), poverty of content of speech (SANS item score), and inappropriate affect (SANS item score); and the psychomotor poverty score was calculated as the sum of poverty of speech (SANS item score), decreased spontaneous movement (SANS item score), and flattening of affect (which was obtained by calculating the mean

FIG. 1. Processing of structural MRI scans. The brain is reoriented to the intercommissural (AC–PC) plane (A) and then scalp-edited by the automatic scalp-editing program ATOMIA with final completion performed manually (B). The gray matter image is produced by segmentation between the gray–CSF and gray–white thresholds and setting pixels within this range to an arbitrary intensity value (C). The gray matter image is smoothed using an 8-mm FWHM (full width half maximum) isotropic Gaussian kernel (D). The white matter image is produced by segmentation above the gray–white threshold and setting pixels within this range to an arbitrary intensity value (E). The white matter image is smoothed using an 8-mm FWHM (full width half maximum) isotropic Gaussian kernel (F).
of the SANS item scores for affective nonresponse, unchanging facial expression, paucity of expressive gesture, and lack of vocal inflection).

Data Acquisition

Brain MRI scans were obtained with a 1.0-T Picker HPQ Vista system using a radiofrequency spoiled volume acquisition relatively T1-weighted to give good gray/white contrast and anatomical resolution (TR, 21 ms; TE, 6 ms; flip angle, 35°; field of view, 25 cm; 192 × 256 matrix × 140 slices of 1.3 mm thickness).

Image Processing and Spatial Normalization

Image analysis was performed on a SPARC X workstation (Sun Microsystems Europe Inc., Surrey, UK) using ANALYZE software (BRU, Mayo Foundation, Rochester, MN; Robb, 1990), ATOMIA software for automatic scalp-editing (Allain et al., 1995), and SPM94 software (spatial normalization and Statistical Parametric Mapping software developed in the Wellcome Department of Cognitive Neurology, UK; Friston et al., 1995a,b) running in MATLAB (Mathworks Inc., Sherborn, MA).

MR images were interpolated to yield a cubic 1-mm voxel size (using “matrix operations” in ANALYZE) and the data set was reduced from 16 to 8 bits (using “load images” in ANALYZE). Images were flipped (reversing left and right to allow compatibility with PET data acquired in our unit). They were randomized with scans from control subjects so processing was carried out blind to subject identity and diagnosis. Images were reoriented, by eye, parallel to the intercommissural (AC–PC) plane (Fig. 1A) using the “oblique display fly” option in ANALYZE. Most tissues exterior to the brain were removed by ATOMIA, an automatic scalp-editing program. Scalp-editing was completed manually (Fig. 1B), in the “image edit” option in ANALYZE, utilizing autotrace set to the gray–CSF threshold pixel intensity value and manual limits to remove remaining tissues exterior to the brain, sequentially processing sagittal slices of the image. A transverse plane 2 mm below the level of the most inferior part of the cerebellum was taken as the inferior limit of the brain.

The separation of gray matter from CSF was achieved by a visual interactive method (Harvey et al., 1993). For each scan (before scalp-editing) seven transverse sections at 20-mm intervals were displayed at the top of a screen and the same sections were displayed as binary images at the bottom of the screen with a variable pixel threshold (using the interactive threshold in the image edit option of ANALYZE). Adjustment of the pixel threshold to a point where the binary images appeared to match the brain slices displayed above at the gray–CSF boundary gave the point of rarity for the gray–CSF threshold for the scan. All estimations were performed by a single rater and test–retest reliability was high (Pearson’s r = 0.81). The frequency distribution of the gray and white compartment pixel intensities was bimodal, so the gray–white threshold was obtained by fitting two Gaussian curves to the distribution and the point of rarity was taken as the minimum between the summated curves. This gave good agreement with the point of rarity when determined by the visual interactive method (Pearson’s r = 0.92). Using the gray–CSF and gray–white thresholds the gray matter in each scan was segmented and set to an arbitrary intensity value of 128 (using the “image algebra” option in ANALYZE with the formula output = (input > (gray–CSF threshold – 1)) × (input < (gray–white threshold + 1)) × 128). This produced the gray matter image (Fig. 1C). A white matter image was constructed in a similar manner (Fig. 1E) using the gray–white threshold (with the formula output = (input > gray–white threshold) × 128).

Using SPM94 software, the gray matter image was smoothed (Fig. 1D) with a 8-mm FWHM (full width half maximum) isotropic Gaussian kernel (using the smooth option with the filter specified as 8 8 8). This uses the partial volume effect to create a spectrum of gray matter intensities. This facilitates the detection of differences between scans, after spatial normalization, in which the gray matter has a spatial extent or thickness close to the size of the filter (such as the cortical gray matter in the brain). The white matter image was also smoothed with a 8-mm FWHM Gaussian kernel (Fig. 1F). The smoothed segmented images can be thought of as images of gray (or white) matter density.

The smoothed gray matter image was then transformed using SPM94 into a standard space (Talairach and Tournoux, 1988). This normalizing spatial transformation matches each scan (in a least-squares sense) to a template image that already conforms to standard space. The procedure starts with a 12-parameter affine transformation and a 6-parameter three-dimensional quadratic (or second order) deformation (Fig. 2) followed by nonlinear (plastic) deformations on a slice by slice basis using Fourier-like basis functions (Fig. 3); the parameters are estimated using a least-squares approach after linearizing the problem (Friston et al., 1995a). The normalization was carried out to produce an image in a bounding box (in Talairach space) with x, y, z dimensions of +64 mm to −64 mm, −104 mm to +68 mm, and −56 mm to +72 mm, respectively (using the normalization option in SPM94); the voxel size was set to 2, 2, 4 mm (and the pitch estimate to 0, because the images and templates were oriented at 0° to the AC–PC plane). The template image was constructed from MRI scans of six normal right-handed control males from the MRC Cyclotron Unit MRI database. These scans had undergone gray segmentation and smoothing (as described above) and affine spatial normalization with the MRC Cyclotron Unit PET regional cere-
bral blood flow template; they have been averaged by summation (in ANALYZE) and a symmetrical template was constructed by summing the averaged image with the same image flipped in the interhemispheric plane. The smoothed white matter image was subjected to the same transformation as the corresponding smoothed gray matter image. The spatially normalized gray and white matter images were then smoothed again using a 12-mm FWHM isotropic Gaussian kernel. These processing steps produced a smoothed gray and white matter image for each subject in standard Talairach space.

**Statistical Analysis**

The processed images were analyzed using Statistical Parametric Mapping in SPM94, which implements the General Linear Model. The syndrome scores (reality distortion, disorganization, and psychomotor poverty) were treated as the covariates of interest. The global gray or white matter in the processed scans was treated as a confounding covariate (global normalization) and these analyses can therefore be regarded as ANCOVAs (Friston et al., 1990); this process equalizes total gray (or white) matter in each scan and highlights regional gray (or white) matter density differences between scans irrespective of their total gray (or white) matter. Age was also treated as a confounding covariate. After specifying the appropriate design matrix (in SPM94: studies, 1; subjects, 15; scans, 1; covariates of interest, 1 (syndrome scores); covariates of no interest, 1 (subject ages); contrasts, 2 (1 for positive correlations and −1 for negative correlations)), the subject and covariate effects were estimated at each and every voxel (Friston et al., 1991, 1995b). To test hypotheses about regionally specific covariate effects, the estimates were compared using linear contrasts. The resulting set of voxel values for each contrast constitutes a Statistical parametric Map of the t statistic (SPM(t)). The SPM(t) were transformed to the unit normal distribution (SPM(Z)) and thresholded at 2.33 (or P = 0.01 uncorrected). The significant foci in the SPM(Z) (at a threshold of P = 0.01) were then characterized in terms of spatial extent (k) and peak height (u). The significance of each region was estimated using distributional approximations from the theory of Gaussian Fields. This characterization is in terms of the probability that a region of the observed number of voxels (or bigger) could have occurred by chance (P(n_max > k)), or that the peak height observed (or higher) could have occurred by chance (P(Z_max > u)) over the entire volume analyzed (i.e., a corrected P value).

**RESULTS**

**Image Processing**

The spatial normalization of the MR images with the MRI template was successful in most transverse planes except those at the very top and bottom of the brain where paucity of information about structural topography in the subject image slice led to a distorted result; randomly distributed artifacts in these slices should not generally lead to significant foci after Statistical Parametric Mapping.

**Correlation between Regional Cerebral Gray and White Matter and Syndrome Scores**

**Reality distortion.** The Statistical Parametric Map (SPM(Z)) of negative correlations between regional cerebral gray matter density and reality distortion score is shown in Fig. 4. There was a negative correlation with regional gray matter density in a region which spanned the left transverse temporal (Heschl’s) gyrus (maximum at −38, −30, +12 mm according to the atlas of Talairach, 1988), the left superior temporal gyrus, and the left inferior parietal lobule. The Z scores for these foci exceeded 3.16 (P < 0.001 uncorrected) and the regional difference was significant on the basis of its spatial extent (P(n_max > k)) within the volume analyzed (P = 0.01 corrected). The SPM of positive correlations between regional cerebral gray matter density and reality distortion score did not show any foci with corrected P < 0.01.

The Statistical Parametric Map of negative correlations between regional white matter density and reality distortion score is shown in Fig. 5. There was a negative correlation with regional white matter along the undersurface of the corpus callosum which was significant within the volume analyzed on the basis of its spatial extent (P(n_max > k) less than 0.001) with a maximum in the splenium of the corpus callosum (Z = 4.12 at −18, −40, +8 mm). The SPM of positive correlations between regional white matter density and reality distortion score did not show any foci with corrected P < 0.01.

**Disorganization.** The Statistical Parametric Maps of negative and positive correlations between regional gray and white matter densities and disorganization score did not show any significant foci with corrected P < 0.01.

**Psychomotor poverty.** The Statistical Parametric Map of negative correlations between regional gray matter density and psychomotor poverty score did not show any significant foci with corrected P < 0.01. The SPM of positive correlations between regional gray matter density and psychomotor poverty score (not shown) showed a significant focus in the left temporal lobe (Z = 4.87 at −48, −56, +12 mm) with corrected P = 0.01 (P(Z_max > u) = 0.005). The SPM of negative correlations between regional white matter density and psychomotor poverty score (not shown) showed a significant focus in the left frontal lobe white matter (Z = 3.88 at −8, −12, +36 mm) with corrected P = 0.01 (P(n_max > k) = 0.005). The SPM of positive correlations between
FIG. 2. Sagittal, coronal, and transverse views showing the linear (affine) and nonlinear (quadratic intraslice) spatial normalizing transformation of the smoothed gray matter subject image with the gray matter template.

FIG. 3. Transverse views at two levels (z = 0 and 20 mm) showing the smoothed gray matter image after affine and nonlinear (quadratic) transformation (left), the template image (right), and the smoothed gray matter image after the final nonlinear spatial (plastic) transformation.
regional white matter density and psychomotor poverty score (not shown) showed significant foci in the deep white matter of the left and right temporoparietal lobes with corrected \( P \leq 0.01 \) \((P(Z_{\text{max}} > u) = 0.01)\).

**DISCUSSION**

The objectives of this study were to apply established techniques of functional image analysis to structural MRI scans and to present, as an exemplar, the structural gray and white matter density correlates of syndrome scores in a group of patients with schizophrenia.

**Methodological Issues**

The basic idea was (i) to convert structural MRI data into spatially normalized images of gray (or white) matter density, effected by segmenting the images and smoothing, and then (ii) to use SPM to make inferences about the relationship between gray (or white) matter density and symptoms (or other pathophysiological measures) in a regionally specific fashion. Because the whole brain sum of gray (or white) matter indices was treated as a confound, the analysis reduced to a characterization of relative gray (or white) matter density on a voxel by voxel basis.

The application of spatial normalization to structural MRI images corrects for differences in brain size and shape and facilitates intersubject averaging. Segmentation of images into gray and white matter permits separate analyses of these closely related structures; analysis of the ventricular system (not described here) is possible by a similar technique. Smoothing of the gray and white matter images using an isotropic Gaussian kernel facilitates detection of differences between images where the gray (or white) matter has a spatial extent or thickness comparable with the size of the filter. Global normalization of gray (or white) matter permits partitioning of differences between groups into global and regional effects, for example, to address the issue of regional versus global structural changes in schizophrenia. We have employed a symmetrical MRI template which has undergone affine transformation into Talairach space to perm the interpretation of findings with reference to the Talairach atlas and to ensure compatibility with other studies performed in this stereotactic space. Use of a symmetrical template preserves asymmetries in gray matter density while removing asymmetries of gross anatomy. We will present an analysis of asymmetry indices in a subsequent paper.

The spatial normalization process results in some artifact near the top and bottom of the brain which is liable to reduce detection of significant results in these areas. We have used volumetric MRI scans with a 1.3-mm-slice thickness to give good anatomical resolution. Although the relatively T1-weighted sequence gives good gray/white contrast, segmentation of images has involved a visual interactive process. The construction of gray and white matter images using this technique must be considered an approximation of the true brain structures because of errors resulting from field inhomogeneity in the MRI magnet, variation in the true segmentation threshold within the image, and partial volume effects. However, global normalization of gray and white matter between images will tend to reduce detection of regional differences resulting from errors in selection of segmentation thresholds. Furthermore, regionally specific errors in identifying the threshold are unlikely to be correlated with the effects (covariates) of interest and will therefore be unlikely to produce artifacts in the SPMs. An alternative approach would be to employ multisequence multislice image acquisition and an automated segmentation technique such as cluster analysis, although slice thickness might be greater with inferior anatomical resolution. Further development of the technique would include a method for automatic reorientation of images into the intercommisural plane, which should soon be available (Traver, personal communication).

**Structural Gray and White Matter Correlates of Syndrome Scores**

In this group of patients with schizophrenia we will discuss the findings relating to the reality distortion score, which illustrates the application of this technique. There were negative correlations between reality distortion score and normalized gray matter density
in the region of the left superior temporal cortex and with normalized white matter density in the region of the undersurface of the corpus callosum. In this sample the reality distortion score reflected propensity to experience hallucinations and delusions during a period of remission from acute episodes of illness. It has been postulated that these positive symptoms of schizophrenia result from a failure of integration between intrinsically generated action and perception which, at the neural level, has a basis in abnormal frontotemporal connectivity (Friston, 1995c; McGuire, 1995). In this model, auditory hallucinations may result when neural activity in premotor speech areas (e.g., Broca's area) activates semantic representations in superior temporal regions and there is a concomitant failure to attribute these representations to intrinsically rather than extrinsically generated speech.

McGuire et al. (1993) have found in a SPECT (single photon emission computed tomography) study that during auditory hallucinations, schizophrenic patients showed significantly increased regional cerebral blood flow in Broca's area and trends for increased blood flow in left temporal areas. The authors concluded that these language regions were implicated in the hallucinatory process. Other studies have found an association between auditory hallucinations and increased left superior temporal blood flow (Matsuda et al., 1988; Suzuki et al., 1993). McGuire et al. have performed a PET (positron emission tomography) study comparing schizophrenic hallucinators with schizophrenic nonhallucinators and control subjects carrying out inner speech and auditory verbal imagery tasks (McGuire, 1995). The schizophrenic hallucinators showed significant regional cerebral blood flow decreases in the left temporal cortex and supplementary motor area during the tasks (relative to a control task) whereas the other two groups showed increases in these areas. McGuire et al. have suggested that these regions may have a role in the attribution of speech as intrinsic rather than extrinsic.

Structural abnormalities in the left superior temporal region and corpus callosum may contribute to the development of auditory hallucinations. Abnormalities in the left superior temporal region may be associated with functional changes which reduce the threshold for spontaneous intrinsic generation of semantic representations in language areas. An association has been found between the propensity to suffer auditory hallucinations in schizophrenia and reduced right ear (and presumably left temporal lobe) advantage in dichotic listening tasks (Matthew et al., 1993; Green et al., 1994). Functional abnormalities in auditory cortex in schizophrenia have also been suggested by studies of event-related potentials; reduced amplitude of the auditory P300 event-related potential was associated with gray matter volume reductions in the left superior temporal gyrus in schizophrenics but not in controls (McCarley et al., 1993). Structural studies of left superior temporal gyrus volume have found a negative correlation with scores for auditory hallucinations in one study (Barta et al., 1990) although not in others (Shenton et al., 1992; DeLisi et al., 1994).

The corpus callosum may have a role in inhibitory interhemispheric modulation (Cook, 1986) and structural disturbance may be associated with impairment in this function, either reducing the threshold for intrinsic activity in language areas or interfering with the correct attribution of this activity to internal processes. Woodruff et al. (1993) found reduced area of the corpus callosum in male schizophrenics compared to controls and the area difference was particularly marked in the midsection connecting the two temporal lobes; there was a negative correlation between the corpus callosum area and a score for delusions. Reduced size of the posterior corpus callosum has been found to correlate with reduced right ear advantage in the auditory comprehension test in schizophrenics (Colombo et al., 1993). During auditory verbal imagery in a PET study, schizophrenic hallucinators showed significantly greater ($P < 0.01$) right superior temporal gyrus activation than nonhallucinators and controls (McGuire, personal communication); this might reflect an imbalance in activation between the superior temporal gyri, reflecting left superior temporal abnormality or impaired regulation of coactivation mediated by loss of corpus callosum connections.

An alternative conjecture to explain these results is that in this sample of patients the reality distortion score is associated with underlying disease etiology or severity. Schizophrenia is believed to be a neurodevelopmental disorder (Murray and Lewis, 1987) and various studies have suggested that the disease process is associated with enlarged cerebral ventricles (Van Horn and McManus, 1992), reduced area of the corpus callosum (Woodruff et al., 1993), and possibly reduced asymmetry in the planum temporale area (supported by the study of DeLisi et al., 1994, although not the studies of Rossi et al., 1994 or Kleinschmidt et al., 1994). The potentially confounding influence of these factors is therefore worth further investigation.

The application of these imaging techniques to the study of structural data permits separate analyses of gray and white matter and the partitioning of global and regional differences. The method corrects for interindividual variation in brain size and shape, thus potentially enhancing detection of differences associated with the variable of interest (e.g., disease pathology, clinical characteristics). This preliminary study of the structural correlates of syndrome scores in a small group of patients with schizophrenia needs further replication. The method may be valuable in the investigation of other brain diseases in which structural abnormalities are suspected.
ACKNOWLEDGMENTS

We gratefully acknowledge the help of Dr. M. Juptner in preparation of the gray matter template image for the MRC Cyclotron Unit, the cooperation of the patients who took part in this study, and the help and advice of colleagues in the Institute of Psychiatry and the MRC Cyclotron Unit.

REFERENCES


Andreasen, N. C. 1987. Comprehensive Assessment of Symptoms and History. Univ. of Iowa College of Medicine, Iowa City, IA.


