ELSEVIER

Psychiatry Research: Neuroimaging Section 67 (1996) 203–214

PSYCHIATRY
RESEARCH
NEUROIMAGING

Stereological MRI volumetry of the frontal lobe

Yvette I. Sheline*[a,b,c,e], Kevin J. Black[a,b], Daniel Y. Lin[a], Gary E. Christensen[c,d,e], Mokhtar H. Gado[c,e], Barry S. Brunsden[c,e], Michael W. Vannier[c,e]

[a] *Department of Psychiatry, Washington University School of Medicine, 4940 Childrens Place, St. Louis, MO 63110, USA*
[b] *Department of Neurology, Washington University School of Medicine, 4940 Childrens Place, St. Louis, MO 63110, USA*
[c] *Department of Radiology, Washington University School of Medicine, 4940 Childrens Place, St. Louis, MO 63110, USA*
[d] *Department of Surgery, Washington University School of Medicine, 4940 Childrens Place, St. Louis, MO 63110, USA*
[e] *Mallinckrodt Institute of Radiology, Washington University School of Medicine, 4940 Childrens Place, St. Louis, MO 63110, USA*

Received 2 May 1995; revised 15 February 1996; accepted 18 February 1996

Abstract

Stereology was used to measure frontal lobe volume on magnetic resonance imaging (MRI) scans in a multi-observer repeated-measures trial in 17 adults. Prior to measurement, MR image volumes were reoriented into coronal sections perpendicular to the bicommissural plane. Three observers blinded to subject identity repeatedly used fixed grid stereology to estimate frontal lobe volumes, defined as all sections of the frontal lobe anterior to the anterior commissure. The lateral ventricles were excluded. Stereological measurement yielded high repeatability and precision, and was time efficient for the raters. The coefficient of error was 0.03. The inter-rater correlation coefficient = 0.95 for three raters; intra-rater correlation coefficients = 0.95–0.98. A comparison was made between stereological and traditional edge tracing measurement of the frontal lobe volumes. The overall correlation between the two methods was 0.95. The use of internal landmarks to define orientation and 3-D orthogonal views to define frontal lobe boundaries on 3-D images was critical to obtaining repeatable measurements. Frontal lobe volumetry by brain MR used to estimate small differences postulated to occur in certain psychiatric and neurologic disorders requires high precision and repeatability. Stereology, a semi-automated method, can reliably estimate frontal lobe volumes. This method may distinguish small frontal lobe volume differences within individuals and between groups.

Keywords: 3-D MRI; Stereology; Volumetry; Frontal lobe; Segmentation; Depression; Schizophrenia

*Corresponding author, Box 8134, Washington University School of Medicine, 4940 Childrens Place, St. Louis, MO 63110, USA. Tel: +1 314 362 2586; fax: +1 314 362 2099.

PII S0925-4927(96)02831-4

1. Introduction

Stereological techniques, based on the Cavalieri principle, have often been applied in histological and pathological studies (Weibel, 1979). More recently, the techniques have been adapted for 3-D MRI volumetry. The volume of the structure is estimated by multiplying the cross-sectional area times the slice thickness, or in the case of non-contiguous slices, times the slice thickness plus the interslice gap. When applied to brain MR data, stereological methods for MR volumetry provide an accurate, unbiased method for determining brain structural volumes (Gunderson, 1992). Point counting methods have been shown to have accuracy and reproducibility in various brain volume studies in which fixed grids were applied in repeatability trials (MacFall et al., 1994; Arndt et al., 1994). Stereological techniques have been applied to measure many different brain structures including pituitary (Axelson et al., 1992), hippocampus (Axelson et al., 1993; Haller et al., 1994; Sheline et al., 1996) caudate (Krishnan et al., 1992) and putamen (Husain et al., 1991). Parameters which affect the volume assessment include the number of slices intersecting the structure, the orientation, boundary definitions, grid density and slice thickness. In the studies enumerated above, slice thickness, orientation and boundaries were described, and two studies (MacFall et al., 1994; Haller et al., 1994) systematically tested grid sizes. In the current work, we used a computer generated point counting technique which automatically summed up the points counted and calculated an estimated volume. We explicitly defined (1) frontal lobe boundaries, (2) orientation using internal landmarks to improve repeatability, (3) grid size and number of slices, and (4) a slice thickness of 0.5 mm, which minimizes partial volume effect.

In applying these methods to frontal lobe volumetry, accuracy is important since differences expected between scans of patients with psychiatric disorders are small compared with normal controls. Coffey et al. (1993) found differences in frontal lobe volumes between depressed subjects and normal subjects of only 7%. Studies in schizophrenia have reported differences of approximately 9% for total prefrontal cortex volume (Breier et al., 1992). Using a standardized approach to define 'frontal lobe' boundaries, we measured frontal lobe volumes by stereological techniques in 17 subjects. We determined both intra- and inter-rater reliability using multiple raters and repeated trials. In addition, we compared measurement of frontal lobe volumes by stereological and traditional edge tracing techniques.

2. Methods

2.1. Subjects

Seventeen subjects were recruited from the Memory and Aging Project of the Alzheimer's Disease Research Center (ADRC) at Washington University School of Medicine (WUSM) and from the WUSM outpatient psychiatry service. Subjects ranged in age from 51 to 86. All subjects were in good health, and were excluded if they had any medical problem which could potentially affect the central nervous system, such as a history of neurological disorder, myocardial infarction, hypertension, chronic steroid use, drug or alcohol abuse or head trauma. Informed consent was obtained for all subjects.

2.2. MRI scan parameters

Studies were performed on a Magnetom SP-4000 1.5 Tesla imaging system (Siemens Medical Systems, Inc., Iselin, New Jersey) using a 30-cm circularly polarized head coil. Subjects were positioned with their canthomeatal line perpendicular to the table, and head position was secured with adjustable cushioned head supports. A fast gradient echo MR acquisition, magnetization prepared rapid gradient echo (MPRAGE), was utilized to generate anatomic images. One hundred twenty-eight contiguous 1.25 mm thick sagittal slices were acquired. No sedation was used during scanning. Specific MPRAGE scanning parameters were: TR = 10 ms; TE = 4 ms; inversion time = 300 ms; flip angle = 8; 1 slab = 60 mm; 3-D partitions = 128; slice thickness = 1.25 mm; pixel

size = 1 × 1 mm; FOV = 256 mm; matrix = 256 × 256; voxel size = 1 × 1 × 1.25; scan time = 11.04 min.

2.3. Preprocessing of images

All image processing was done using *ANALYZE*™ (Biomedical Imaging Resource, Mayo Foundation) (Robb, 1990) on a UNIX workstation (Sun Microsystems Inc., Mountain View, CA). Prior to any stereological measurements of the frontal lobe volume, the images underwent three processes. The first process was interpolation to 0.5-mm sections. The second process was to reorient the images to the anterior commissure-posterior commissure (AC-PC) plane (Fig. 1). Upon reorientation, the images were processed for gray scale normalization.

Reorienting the images to the AC-PC plane allowed us to obtain coronal sections in a standard orientation for more repeatable determination of volumes. For consistency and reliability, internal landmarks were chosen. The AC and PC were both identified and images were reoriented so that coronal sections were perpendicular to the AC-PC plane as follows: After locating the AC and PC in the mid-sagittal plane, the image was rotated so that a line passing through the most superior point of the AC and the most inferior point of the PC was parallel to the base of the screen. Next, in the transverse orientation, the image was rotated so that the central fissure was oriented 90° to the base of the screen. Then in the coronal view, the image was rotated so that the central fissure was oriented 90° to the base of the screen. Finally, the matrix describing the

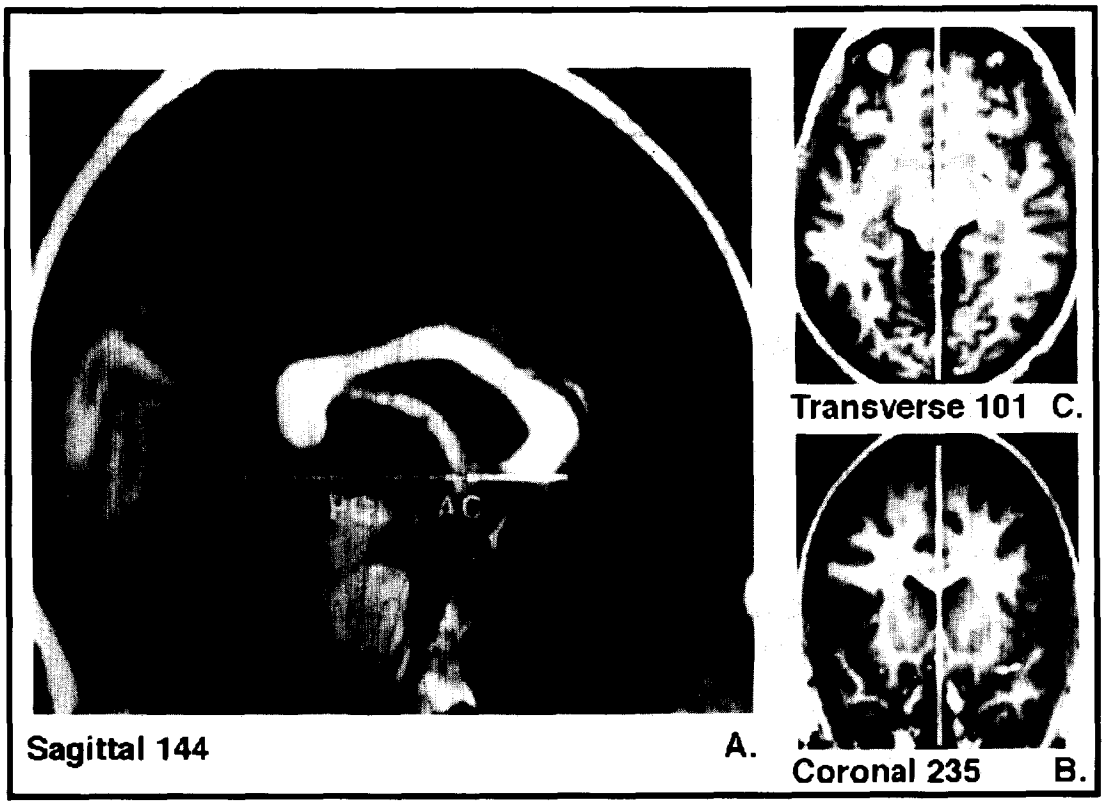

Fig. 1. AC-PC reorientation. The images were oriented using the anterior commissure-posterior commissure (AC-PC) plane as the reference plane. The interrupted line passes through the anterior and posterior commissures. This plane is orthogonal to the mid-transverse and mid-coronal planes.

oblique bicommissural plane was stored and applied to the original data set.

2.4. Gray scale normalization

The third process involved gray scale normalization (Fig. 2). Normalizing the gray scale range is essential for repeatable valid measurements because scans between patients yielded wide variations of gray scale range. Normalization was done by drawing a circle on the mid-hippocampal transverse slice with the center on the cerebral aqueduct. A cylindrical sub-volume was formed containing the hippocampus, adjacent white matter, and ventricular cerebral spinal fluid (CSF) but excluding virtually all fat or other tissue. The gray scale histogram of this cylindrical sub-volume was then used as an aid for tissue classification. The cylindrical 16-bit data sets were analyzed (*PeakFit*, Jandel Scientific, San Rafael, CA), by the Marquardt-Levenburg algorithm for non-linear curve fitting (Donald and Marquardt, 1963; Levenberg, 1944). Five Gaussian curves were used to fit each histogram where different peaks represented the different tissues (white matter, gray matter, CSF, and gray-white and gray-CSF partial volumes). This technique produced gray scale histograms with well-defined peaks which also showed consistent shapes between different subjects. We arbitrarily defined a cutoff for each tissue type as $\pm$ 2 standard deviations (S.D.) from its peak. We then scaled each image from CSF peak minus 2 S.D. to white matter peak plus 2 S.D. Finally, the data were scaled down to 8-bit resolution using these thresholds. This process produced working data sets with consistent contrast and appearance of all brain matter.

2.5. Stereological method and reliability

Prior to measurement, scans were numerically coded so that raters remained blind to subject identity. Three experienced raters measured right,

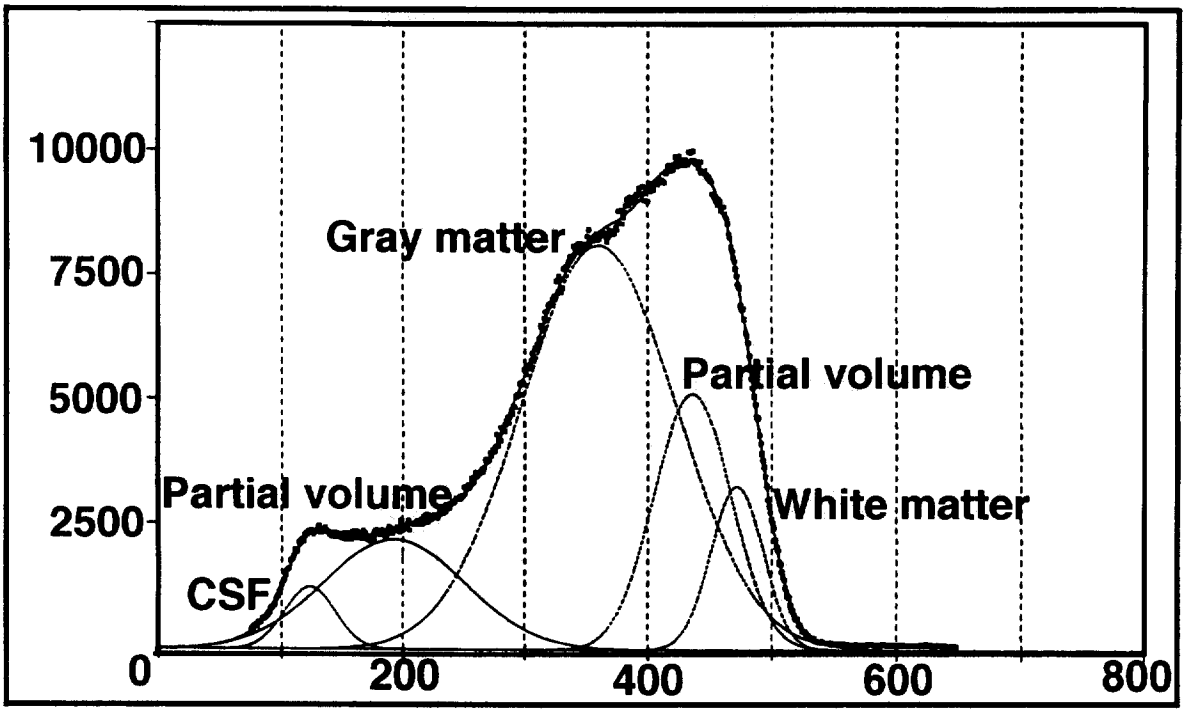

Fig. 2. Gray scale normalization. Histograms of these consistent shapes were obtained from each MR scan. Five curves were used to generate a line that matches the data points: gray matter, white matter, CSF, gray-white partial volume, CSF-gray partial volume.

left and total frontal lobe volumes. Volumes were measured by the stereologic method long used in microscopy and more recently applied to MR volumetry (Gundersen et al., 1988; Mayhew, 1992). From the 3-D MRI volume composed of (0.5 mm)^3 voxels, true coronal slices (Fig. 3) were sampled every 4.5 mm. A start slice 0–8 slices anterior to the anterior commissure was randomly chosen for each trial. The frontal area on each of these slices was measured as follows: A 14 × 14 mm^2 rigid grid of points was superimposed on the image, with starting position and angle of deviation from horizontal chosen randomly for each trial. Those grid points falling within the frontal lobes (see definition of 'frontal lobe' volume below) were marked and counted. Points falling within this volume were further labelled as belonging to the right or left hemisphere. Points falling on the edge were viewed using magnification (see Fig. 3). If the center of the cross-hair fell on a gray pixel (any non-black pixel), it was counted 'in.' An important feature of the *ANALYZE*™ implementation of this stereologic method is the ability to interactively view any grid point marked on transverse and sagittal, as well as coronal planes, providing clarification of the anatomy before selecting a point. From the number of 'hits,' or grid points falling within the volume of interest, an unbiased volume estimate is easily obtained. Raters 2 and 3 measured frontal volume on two independent trials for each subject. We also tested the hypothesis

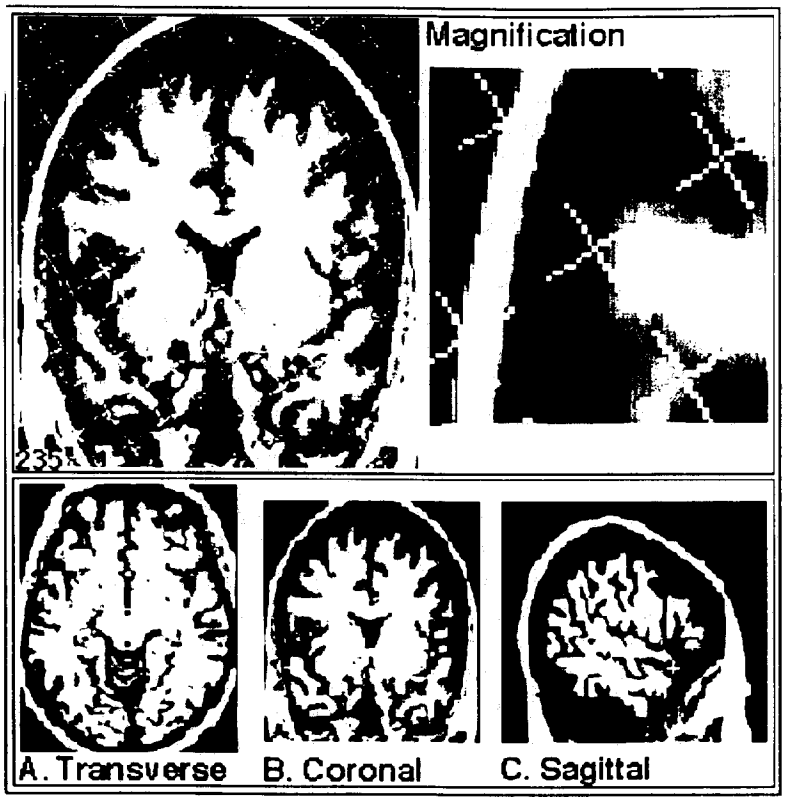

Fig. 3. Measuring frontal volume — stereology. A fixed grid of 14 × 14 mm^2 was used to measure the volume; a magnification tool allows each point to be magnified. Individual points can be viewed concurrently in all three planes: transverse (A), coronal (B) and sagittal (C).

that an average of two measurements would be more accurate and rater 1 measured frontal lobe volume four times for each subject.

2.6. Stereological parameter selection

The coronal orientation was chosen for three reasons. First, the distinction of frontal from temporal lobes is most apparent in this orientation. Second, this provides the most direct comparison with the prior studies of Coffey et al. (1993) and Breier et al. (1992). Most importantly, the shape of the frontal lobes varies least from slice to slice along this axis. Grid size and slice spacing were chosen as follows. Gundersen and Jensen (1986) demonstrate that for a desired coefficient of error (CE) of 0.03, and for complex shapes such as human cortex, approximately 175 'hits' will suffice. The CE measures the slice to slice variability of measurement and was calculated automatically by ANALYZE™. Mayhew and Olsen (1991) compared CEs obtained with different numbers of slices. They demonstrated a CE of 0.027 for a 1:3 sampling and 0.049 for 1:5 sampling. However, their slice thickness was 5 mm so that a 1:3 sampling would occur every 15 mm. Using conservative estimates of the shape complexity of the cortex, one can estimate the number of slices needed for reliable measurement with 150 points (Mayhew, 1992; Gundersen and Jensen, 1986), yielding an estimate of 16 slices in our case. In fact, it has previously been demonstrated that the volume of the entire cerebral cortex can be accurately estimated by this stereologic method when these points are spread over only 13 slices (Regeur and Pakkenberg, 1989). Our preliminary studies, using images from subjects not measured in this study, showed that the grid size and slice distance we use here gives approximately 300 frontal 'hits' on about 16 slices, or well over the necessary number of points and slices estimated by mathematical principles. In fact, the approximately 150 points predicted to fall in each hemisphere are close enough to the number required for accurate measurement that we tested accuracy of measurement of these subvolumes.

2.7. Definition of frontal lobe volume

As described above, we first reoriented the MR scans in alignment with the AC-PC plane. This made it easy to apply our definition of 'frontal lobe' volume: namely all brain parenchyma anterior to the anterior commissure excluding the temporal lobes (see Fig. 4). Volume was determined stereologically using only coronal slices anterior to the slice on which the AC was best visualized.

2.8. Comparison with edge tracing

In order to relate the values obtained by our method to other means of determining MRI volumetry, a direct comparison was made between stereological measurement techniques and edge tracing for all subjects. The definition of 'frontal lobe' volume was the same as for the stereological measurements. A manual tracing was performed on 4-mm thick, nonoverlapping, contiguous coronal slices beginning at the AC, to define an *ANALYZE*™ object. Manual touch-up was allowed on thinner slices if needed. Manual segmentation for each subject was added by masking pixels with CSF intensity. Specifically, a brain:CSF threshold for each subject was set as the average of the top of the CSF peak and bottom of the gray matter peak, as described above in the section on gray scale normalization (see Fig. 2). Although technically this manual segmentation was not performed blind to the prior stereologic volume measurements, it was done 5–9 months later without reference to the earlier volume measurements, on numerically coded scans.

2.9. Statistical analysis

Source of error for stereological measurements was evaluated by ANOVA. Nested ANOVAs for both inter-rater and intra-rater measurement error were performed factoring in rater, repeated measure and the residual term (measurement session nested within rater). The Spearman-Brown prediction formula was used to determine inter-rater and intra-rater reliability:

$$r_s = \frac{k\theta}{1 + k\theta} \qquad \text{where } \theta = \frac{\sigma_\pi^2}{\sigma_\eta^2} \tag{1}$$

where r_s is the Spearman-Brown correlation. In

equation (1), k is the number of raters and theta is the variance of the instrument divided by the variance of the measurement error (Winer et al., 1991).

The intra-rater correlation coefficient was calculated for within rater measures from 34 measurements (2 measures of 17 subjects), with a minimum goal of 0.95. Inter-rater correlation coefficients were calculated between the 3 raters using equation (1) above with $k = 3$. The goal for minimum between-rater measures was set at 0.90. A comparison between stereological and edge

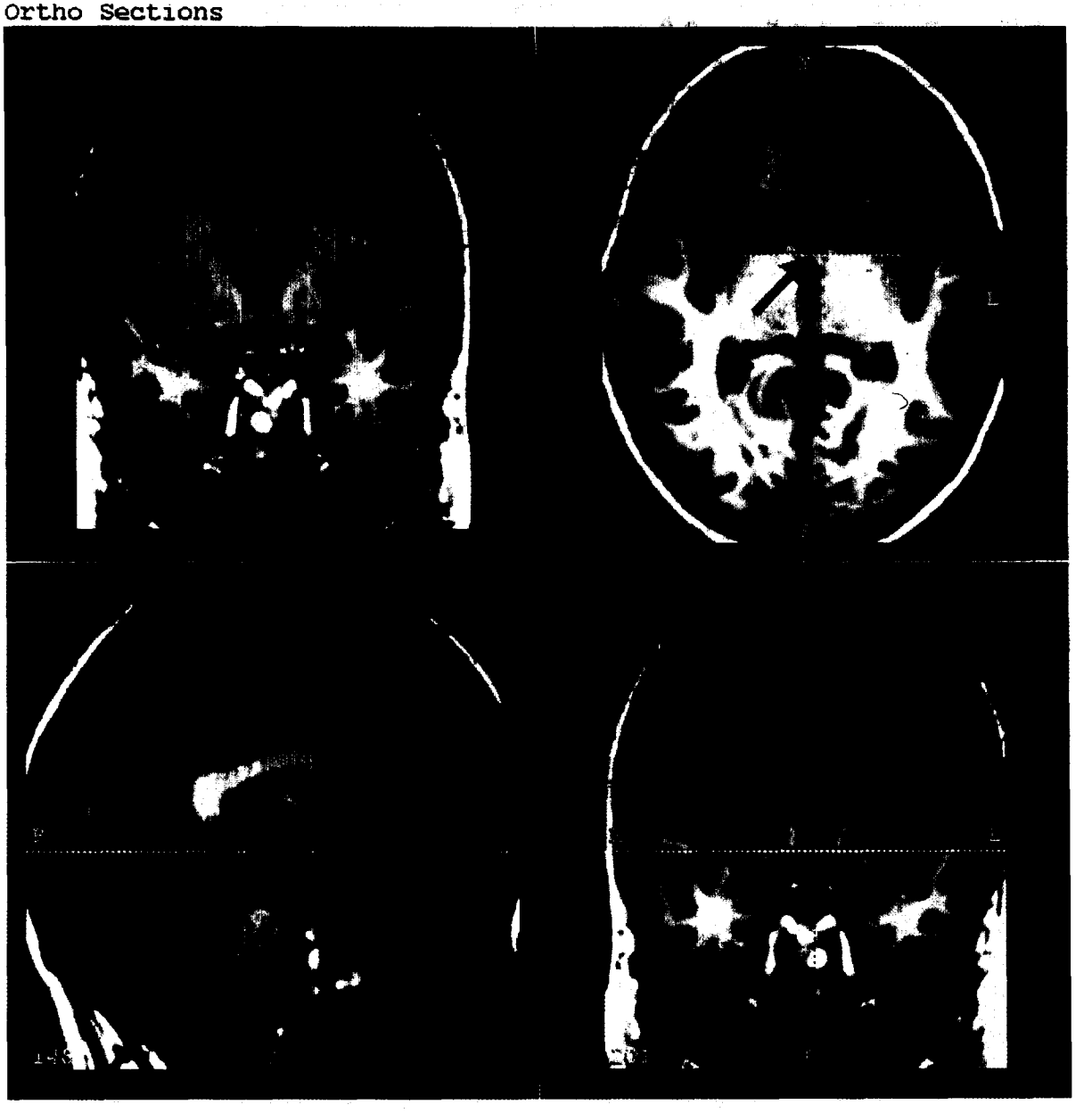

Fig. 4. Measuring frontal volume — edge tracing. The figure shows orthogonal views of a frontal lobe volume determined by edge-tracing. Yellow area corresponds to the traced frontal lobe volume. Any point can be selected and visualized in three orthogonal sections. In this figure, the cursor points to the anterior commissure.

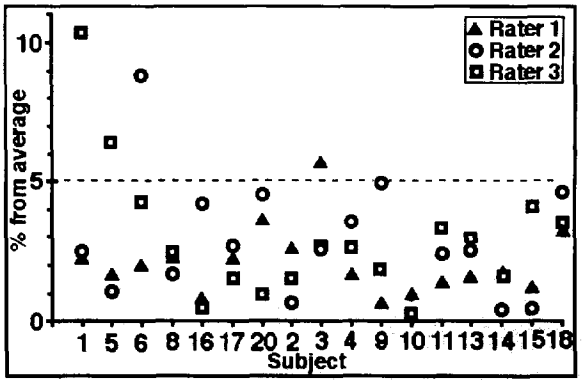

Fig. 5. Intra-rater repeatability. The figure shows deviation from perfect correlation; most measurements (47 out of 51) fell within 5% of the average.

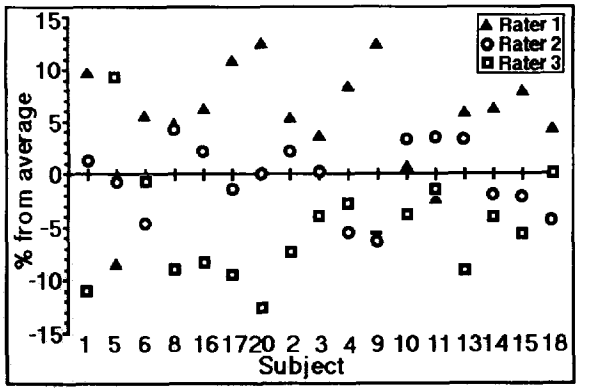

Fig. 6. Inter-rater repeatability. The figure shows deviation from perfect correlation; most measurements (46 out of 51) fell within 10% of the average.

tracing measurements of frontal lobe volumes was also made using the Spearman-Brown prediction formula.

3. Results

All three raters achieved an intra-rater correlation coefficient ≥ 0.95 for the total frontal lobe volume measurements and intra-rater correlation coefficients of ≥ 0.93 for all measurements (Table 1). Table 1 also displays the CE associated with each rater's measurements, ranging from 0.03 to 0.05. In addition, most measurements (47 out of 51) were within 5% of each rater's own mean; all were within 10.3% (Fig. 5). Inter-rater correlation coefficients were 0.95 for total frontal lobe volume, 0.94 for left frontal lobe and 0.96 for right frontal lobe. No measurements differed from the mean of the group by more than 12% (Fig. 6).

Measurements using stereology compared with edge tracing are shown for all subjects in Table 2. The methods were highly correlated, with an inter-class correlation coefficient of 0.95. The mean and standard deviation for frontal lobe volume measurement using stereology were 359.0 and 51.7 cm^3, respectively, and for edge tracing the respective mean and standard deviation were 370.5 and 54.6 cm^3. The differences ranged from 0.2% to 10.2%, with a mean difference of 4.5% between the two methods.

4. Discussion

The goal of this work, to achieve high precision and repeatability in estimating 'frontal lobe'

Table 1
Intra-rater correlation coefficients

	Right frontal correlation	CE	Left frontal correlation	CE	Total frontal correlation	CE
Rater 1*	0.97	0.03	0.96	0.03	0.98	0.03
Rater 2	0.93	0.03	0.95	0.03	0.95	0.03
Rater 3	0.93	0.05	0.95	0.03	0.95	0.04

The intra-rater correlation coefficients and coefficients of errors are shown. Rater 1 made four measurements whereas raters 2 and 3 made two measurements each.

Table 2
Comparison of stereological and edge-tracing measurement techniques

Subject	Edge tracing volume (mm^3)	Stereology volume (mm^3)	Volume difference (mm^3)	Percent difference (%)
S-01	467545	439016	28529	6.10
S-05	367380	345818	21562	5.87
S-06	327146	327736	−590	−0.18
S-08	351140	352286	−1146	−0.33
S-16	367519	335160	32359	8.80
S-17	310235	333764	−23529	−7.58
S-20	455070	447466	7604	1.67
S-02	369522	337953	31569	8.54
S-03	484245	461433	22812	4.71
S-04	308856	277279	31577	10.22
S-09	412152	399908	12244	2.97
S-10	348979	353682	−4703	−1.35
S-11	322134	310244	11890	3.69
S-13	316356	312448	3908	1.24
S-14	376970	381465	−4495	−1.19
S-15	373974	368456	5518	1.48
S-18	338629	318843	19786	5.84
Mean	370462	358997	11464	2.97
Standard deviation	54631	51658	15791	4.49

A comparison of edge tracing and stereological techniques for measuring frontal lobe volume is displayed.

volumes, was attained: The overall CE was 0.03; the inter-rater reliability was 0.95, and the intra-rater reliability ranged from 0.94 to 0.96. Stereological estimates of 3-D MRI volumes provided an efficient and reliable means of determining 'frontal lobe' volumes. While this volume does not correspond exactly to the true frontal lobe, the method described provides a means for measuring a 'frontal lobe' volume defined by internal landmarks with a specific posterior boundary (the anterior commissure) which allows highly repeatable and accurate measurements.

To date, most studies have used lower resolution MR data, usually with 5-mm thick sections and manual segmentation, which is labor intensive and has variable repeatability in structures with ambiguous boundaries. MacFall et al. (1994) concluded from both theoretical and experimental evidence that systematic point counting provides better results than tracing, although to date there has been no systematic comparison. In a direct comparison of stereology with more traditional edge tracing methods in the current study, the methods yielded very similar results. This further supports the validity of stereological measurements. One difference in our comparison was that stereological techniques were more time efficient, requiring an average of 25 min for measuring both right and left frontal lobes separately, whereas edge tracing required 40 min on average to trace the total frontal lobe. An advantage of edge tracing, as currently configured on *ANALYZE*™ software, on the other hand, is that the object being traced is saved and therefore can be directly compared between raters, whereas the points selected in stereology are not saved. Mayhew and Olsen (1991) have demonstrated the validity of point counting methods in imaging studies of a post-mortem brain specimen. As discussed in the Introduction, findings from stereology are generalizable and previous studies have used stereological techniques to measure a vari-

ety of structures ranging from very small structures such as the pituitary (Axelson et al., 1992) and hippocampus (Axelson et al., 1993; Haller et al., 1994; Sheline et al., 1996) to the whole brain. Theoretical considerations involve the increase in the measured surface to volume ratio as the total volume decreases because the relative number of edge pixels traced manually and, therefore, the error increases in smaller structures. Volume measurement error caused by error in segmentation labeling is related to the ratio of the structure size and the voxel dimension. The smaller the structure is relative to the voxel dimension, the larger the error is in the volume measurement due to mis-labeled voxels. Table 3 demonstrates this by tabulating the effect of a one voxel erosion and dilation of a 3-D hippocampus segmentation on the volume measurement at four different resolutions. This table shows that a systematic error of plus or minus one layer of surface voxels has approximately an order of magnitude greater effect when using 1 mm resolution isotropic voxels than when using 0.1 mm resolution isotropic voxels. Therefore, a small error in the segmentation of a small structure relative to the voxel dimension causes a large error in the volume measurement, whereas a small error in the segmentation of a large structure only causes a small error in the volume measurement. Stereology, with its lesser dependence upon measurement of surfaces, is a particularly attractive method relative to edge tracing for small, complex structures such as the hippocampus.

The frontal lobes are difficult to distinguish from the parietal lobes. Although the central sulcus provides a dividing boundary on the surface, there is no means to extend the demarcation internally. Most studies have used an arbitrary coronal plane to separate the brain, with all the tissue anterior to the plane consisting of the 'frontal' lobe. As most studies have relied on patient positioning in the scanner by external landmarks to define the coronal orientation, this introduces a potentially significant source of error. In this study, we tested a method which utilizes internal landmarks (anterior and posterior commissures) to define a specific plane of orientation, and is thus consistent between patients. We defined the subvolume of frontal lobe anterior to the anterior commissure as 'frontal lobe' volume.

The anterior commissure was selected as the internal landmark and boundary for several reasons. First, since we have thin sections (0.5 mm) we can use a small landmark for a more precise boundary. Second, we could use a landmark which was part of the brain proper, and which achieved a division of a subvolume closely corresponding to the true frontal lobe. The Talairach Atlas, Talairach and Tournony, 1988 displays the localization of the central sulcus in 20 subjects in relation to the anterior commissure. In all cases, the most anterior and lateral extent of the central sulcus fell within 4.5 mm of the coronal section through the AC. In the midline, the central sulcus falls approximately at the location of the PC, and thus a portion of frontal lobe is omitted in our definition of 'frontal lobe.' The portion of frontal lobe volume omitted in our measurements is motor cortex, which is unlikely to be specifically involved

Table 3
Theoretical effects of changing the edge in segmentation

Resolution (mm)	Original volume (mm^3)	Volume after erode (mm^3)	Volume after dilate (mm^3)	Percent change due to erosion	Percent change due to dilation
0.1	3451	3210	3699	7.0	7.2
0.2	3452	2977	3955	13.8	14.6
0.5	3447	2316	4752	32.8	37.9
1.0	3490	1424	6237	59.2	78.7

Effect of a one voxel erosion and dilation of a 3-D hippocampus segmentation on volume measurements at different resolutions.

in disorders such as depression and schizophrenia. The AC relationship with central sulcus is known, whereas the relationships between central sulcus and other points of demarcation, such as the optic chiasm, have not been defined. Others have used the genu of the corpus callosum (Breier et al., 1992), which has the disadvantage that the corpus callosum itself has been implicated in schizophrenia. Another landmark which has been used is the optic chiasm (Coffey et al., 1993), which is larger than the AC, introducing more variability in measurement. Wible et al. (1995) used the temporal stem as the posterior boundary, which has the advantage that it controlled for differences in lateral asymmetries, but may be influenced by temporal lobe anomalies. Finally, an additional reason to use the anterior commissure is that it is a Talairach Atlas fixed point, and thus can be used in conjunction with other methods which use the Talairach Atlas for anatomical definition, and which is standard in a large neuroscience literature. One reason for the high correlation in our study between frontal lobe volumes measured by stereology or by edge tracing was that both methods used the same highly repeatable boundary for determining frontal lobe demarcation. In summation, we believe our method of estimating frontal lobe volume to be both reliable anatomically and efficient methodologically.

Additional methodological issues not addressed by this article include sources of variability introduced by many of the parameters chosen for this investigation which were not directly tested. Areas for future investigation include systematically testing the effect of varying the numbers of points and number of slices on the reliability of frontal lobe volume determination, and directly comparing the variability obtained using different posterior margins for the frontal lobes, such as the corpus callosum, optic chiasm, lateral ventricles or striatum. Additionally, since in some studies volume reductions have been reported in prefrontal white matter (Breier et al., 1992) and correlations with temporal lobe areas have been demonstrated for right prefrontal white (Breier et al., 1992) and left prefrontal gray matter (Wible et al., 1995), it will be important in future investigations to apply stereological techniques to distinguishing white and gray matter during volumetric determinations.

Frontal lobe volume reductions have been reported in patients with schizophrenia and depression compared with normal controls. In schizophrenia, MRI studies have found evidence for reductions in the volume (Andreasen et al., 1986) of the gray matter in the frontal lobe (Zipursky et al., 1992) and white matter of the frontal lobe (Breier et al., 1992), although others have not found differences between schizophrenic patients and controls in either prefrontal cortex volumes or prefrontal white matter volumes (Wible et al., 1995). In primary affective disorder, only one controlled study of frontal lobe volume has been published to date (Coffey et al., 1993), finding that depressed patients had smaller frontal lobe volumes than matched controls. Using our data, we made power calculations from which we were able to determine the sample sizes necessary to detect group differences. Assuming a 7% difference in average frontal lobe volume for a two-tailed test at the 5% level and a power of 90%, the sample sizes needed for demonstrating left, right and total frontal lobe differences would be 60, 46 and 52, respectively. Assuming a 10% difference, the respective sample sizes required would be 29, 23 and 25 (Kraemer and Thiemann, 1987). These sample sizes are reasonable and agree with other work reported in the literature. Obviously, these sample sizes only obtain for our particular subject population, and the ability to extrapolate to another population would depend critically upon the amount of frontal lobe volume variability in the population being measured. Given the small differences in frontal lobe volumes reported in studies to date, it is important to have a method with high repeatability. Stereology satisfies this criterion and can improve frontal lobe volume measurement reliability.

Acknowledgements

This research was performed at the Mallinckrodt Institute of Radiology at Washington University Medical Center, St. Louis. The Biomedical Imaging Resource at the Mayo Clinic, Rochester,

MN, provided the *ANALYZE*™ software courtesy of Richard A. Robb, Ph.D. This work was supported in part by NIH grant ROI NS35368 to MWV and the Clinical Mental Health Academic Award 1 K07 MH01370 01A1 to Y15. Additional support was provided by an NIAAA fellowship to KJB.

References

Andreasen, N.C., Nasrallah, H.A., Dunn, V., Olson, S.C., Grove, W.M., Ehrhardt, J.C., Coffman, J.A. and Crossett, J.H.W. (1986) Structural abnormalities in the frontal system in schizophrenia: A magnetic resonance imaging study. *Arch Gen Psychiatry* 43, 136–144.

Arndt, S., Swayze, V.W., Cizadlo, T., O'Leary, D., Cohen, G., Yuh, W.T.C., Ehrhardt, J.C. and Andreasen, N.C. (1994) Evaluating and validating two methods for estimating brain structure volumes: Tessalation and simple pixel counting. *NeuroImage* 1(3), 191–198.

Axelson, D.A., Doraiswamy, P.M., Boyko, O.B., Escalona, P.R., McDonald, W.M., Ritchie, J.C., Patterson, L.J., Ellinwood, E.H., Nemeroff, C.B. and Krishnan, K.R.R. (1992) In vivo assessment of pituitary volume with MRI and systematic stereology. *Psychiatry Res* 44, 63–70.

Axelson, D.A., Doraiswamy, P.M., Boyko, O.B., Escalona, P.R., McDonald, W.M., Ritchie, J.C., Patterson, L.J., Nemeroff, C.B., Ellinwood, E.H., Jr. and Krishnan, K.R.R. (1993) Hypercortisolemia and hippocampal changes in depression. *Psychiatry Res* 47, 163–173.

Breier A., Buchanan, R.W., Elkashef, A., Munson, R.C., Kirkpatrick, B. and Gellad, F. (1992) Brain morphology and schizophrenia: magnetic resonance imaging study of limbic, prefrontal cortex, and caudate structures. *Arch Gen Psychiatry* 49, 921–926.

Coffey, C.E., Wilkinson, W.E., Weiner, R.D., Parashos, L.A., Djang, W.T., Webb, M.C., Figiel, G.S. and Spritzer, C.E. (1993) Quantitative cerebral anatomy in depression: A controlled magnetic resonance imaging study. *Arch Gen Psychiatry* 50, 7–16.

Donald, J. and Marquardt, W. (1963) An algorithm for least-square estimation of nonlinear parameters. *J Soc Indust Appl Math* 11(2), 431–441.

Gundersen, H.J.G. (1992) Stereology: The fast lane between neuroanatomy and brain fuction or still only a tightrope. *Acta Neurol Scand Suppl* 137, 8–13.

Gundersen, H.J.G., Bendtsen, T.F. and Korbo, N. (1988) Some new, simple and efficient stereological methods and their use in pathological research and diagnosis. *Acta Pathol Microbiol Scand* 96, 379–394.

Gundersen, H.J.G. and Jensen, E.B. (1986) The efficiency of systematic sampling in stereology and its prediction. *J Microsc* 143(3), 229–263.

Haller, J., Botteron, K., Brunsden, B., Sheline, Y., Walkup, R., Black, K., Gado, M. and Vannier, M. (1994) Hippocampal MR volumetry. *Int Soc Opt Eng Proc* 2359, 660–671.

Husain, M., McDonald, W.M., Doraiswamy, P.M., Figiel, G.S., Na, C., Escalona, P.E., Boyko, O.B., Nemeroff, C.B. and Krishnan, K.R.R. (1991) A magnetic resonance imaging study of putamen nuclei in major depression. *Psychiatry Res Neuroimaging* 40, 95–99.

Kraemer, H.C. and Thiemann, S. (1987) *How Many Subjects?* Sage Publications, Newbury Park, CA, p. 110.

Krishnan, K.R.R., McDonald, W.M., Escalona, P.R., Doraiswamy, P.M., Na, C., Husain, M.M., Figiel, G.S., Boyko, O.B., Ellinwood, E.H. Jr. and Nemeroff, C.B. (1992) MRI of the caudate nuclei in depression. *Arch Gen Psychiatry* 49, 553–557.

Levenberg, K. (1944) A method for the solution of certain non-linear problems in least square. *Quantitative Appl Math* 2, 164–168.

MacFall, J.R., Bynum, C.E., Parashos, I., Early, B., Charles, H.C., Chittilla, V., Boyko, O.B., Upchurch, L. and Krishnan, K.R.R. (1994) Relative accuracy and reproducibility of regional MRI brain volumes for point-counting methods. *Psychiatry Res Neuroimaging* 55, 167–177.

Mayhew, T.M. and Olsen, D.R. (1991) Magnetic resonance imaging (MRI) and model-free estimates of brain volume determined using the Cavalieri principle. *J Anat* 178, 133–144.

Mayhew, T.M. (1992) A review of recent advances in stereology for quantifying neural structure. *J Neurocytol* 21, 313–328.

Robb, R.A. (1990) A software system for interactive and quantitative analysis of biomedical images. In: Hohne, K.H., Fuchs, M. and Pizer, S.M. (Eds.), *3-D Imaging in Medicine*. Nato ISI Series, Vol. F, pp. 333–361.

Regeur, L. and Pakkenberg, B. (1989) Optimizing sampling designs for volume measurements of components of human brain using a stereological method. *J Microsc* 155, 113–121.

Sheline, Y.I., Wang, P.W., Gado, M.H., Csernansky, J.G. and Vannier, M.W. (1996) Hippocampal atrophy in recurrent major depression. *Proc Natl Acad Sci USA* 93, 3908–3913.

Talairach, J. and Tournoux, P. (1988) *Co-planar Stereotaxic Atlas of the Human Brain: Three-dimensional Proportional System: An Approach to Cerebral Imaging*. Thieme Medical Publishers, New York.

Weibel, E.R. (Ed.) (1979) *Stereological Methods. Practical Methods for Biological Morphometry*. Academic Press, New York, Vol. 1, pp. 1–413.

Wible, C.G., Shenton, M.E., Hokama, H., Kikinis, R., Jolesz, F.A., Metcalf, D. and McCarley, R.W. (1995) Prefrontal cortex and schizophrenia: A quantitative magnetic resonance imaging study. *Arch Gen Psychiatry* 52, 279–288.

Winer, B.J., Brown, D.R. and Michels K.M. (1991) *Statistical Principles in Experimental Design*, 3rd Edition, McGraw Hill, New York, pp. 1014–1021.

Zipursky, R.B., Lim, K.O., Sullivan, E.V., Brown, B.W. and Pfefferbaum, A.A. (1992) Widespread cerebral gray volume deficits in schizophrenia. *Arch Gen Psychiatry* 49, 195–205.