Eur Arch Psychiatry Clin Neurosci (2001) 251 : 255–261

ORIGINAL PAPER

Ingrid Agartz · Gaku Okuguwa · Mikael Nordström · Dan Greitz · Vincent Magnotta · Göran Sedvall

Reliability and reproducibility of brain tissue volumetry from segmented MR scans

Received: 24 April 2001 / Accepted: 24 September 2001

Abstract Reliable measurement of different tissue volumes in the living brain is of great importance for human brain research. In this article, we report on the inter- and intraoperator reliability and scan-rescan reproducibility of segmented intracranial tissue volumes from MR images using the image analysis software suite BRAINS. The absolute data of tissue volume measurements are also presented. The reliability and consistency of the measurements of the segmented volumes were excellent. The segmentation is robust and rapid and the volume measurements are plausible and suitable for quantitative studies in clinical brain research.

Key words magnetic resonance imaging · brain anatomy · tissue segmentation · image analysis · schizophrenia

I. Agartz M.D., Ph. D. (✉) · G. Okuguwa M.D., Ph. D. · M. Nordström M.Sc. · G. Sedvall M.D., Ph. D.
Dept. Clinical Neuroscience, Human Brain Informatics, Psychiatry Section,
Karolinska Institutet and Hospital,
17176 Stockholm, Sweden
Tel.: +46-8/51 77 17 58
Fax: +46-8/51 77 17 17
E-Mail: Ingrid.Agartz@cns.ki.se

G. Okuguwa M.D., Ph. D.
Department of Neuropsychiatry
Kansai Medical University
10–15 Fumizonocho, Moriguchi
Osaka, 570–8506 Japan

D. Greitz M.D., Ph. D.
Dept. Clinical Neuroscience
MR Research Center
Karolinska Institutet and Hospital
17176 Stockholm, Sweden

V. Magnotta Ph.D.
Mental Health Clinic and Research Center
University of Iowa Hospitals and Clinics
College of Medicine
Iowa City, IA, USA

Introduction

At the Human Brain Informatics Center (HUBIN), Karolinska Institutet, a relational database on human brain data in schizophrenia is under development. One important part of the database is to characterize brain morphology using magnetic resonance imaging (MRI) and computer image analysis. Morphological changes in schizophrenia are subtle and not obvious from visual inspection of magnetic resonance (MR) images. Therefore, sensitive and precise methods are required for correct quantification of alterations in morphology. To study the relationship between brain morphology, gene expression and psychomotor functions, exact measures of brain structures in individual brains are necessary. In vivo segmentation of grey levels of MRI images can be used to evaluate brain tissue composition. To increase accuracy and precision of the measurements of brain tissue segmentation from MR images, it is important to have reliable and valid tissue classification methods. The measurement error must be considerably smaller than the between-group differences and the variability between subjects. Excellence with respect to reliability across different research centers that use the same method makes it possible to pool data. In the future, many groups will build databases on human morphological brain data. This kind of general methodological research and method evaluation is necessary for the future development of research in brain morphology.

In the current study, we examined the inter- and intraoperator reliability and scan-rescan reproducibility of segmented intracranial tissue class volumes from high resolution MR images. For tissue segmentation we used BRAINS, a software suite for automatic characterization of brain structures (e. g., Andreasen et al. 1992b, 1993a, Cohen et al. 1992, Harris et al. 1999, Magnotta et al. 1999a) and surface anatomy (Magnotta et al. 1999b) from MR data. Andreasen's group has previously reported excellent reliability and validity using the same program (Harris et al. 1999). However, it is not evident

that other research centers will obtain the same results. Therefore it is important for each center that uses the same program to test reliability and reproducibility.

In volumetric studies, it is also of great importance to report original quantitative data. In the current study, we examined and report on not only measures of reliability and reproducibility but also the absolute values of the segmented intracranial tissue class volumes. This was not the case with the reliability study by Andreasen's group. In some respects, our methods also differ from those of Harris et al. 1999. We use thinner MR slices and a different definition of the inferior border of the intracranial volume (ICV). We also report the absolute values of the segmented tissue volumes, in the right and the left hemisphere.

Methods

Tissue segmentation

To segment brain tissue means to separate the tissue into different segments of homogenous tissue content. This can be accomplished by classification of tissue properties on MR images. Measures in BRAINS, generated by the automatic segmentation method, have been carefully validated against manual tracings considered to be the gold standard (Harris et al. 1999). The segmentation is based on a 3D MR acquisition. The tissue classification program can process unimodal (typically T1 volumes) or multimodal (T1 and T2, or T1, T2 and PD) data sets. The first step in the tissue classification process is the automatic identification of training classes that are entered into a discriminant analysis function that classifies the tissue into grey matter (GM), white matter (WM) cerebrospinal fluid (CSF), and venous blood. The generated image is a continuous (fuzzy) representation of the tissue types. The *continuous classification* identifies the predominant (containing more than 50 % of the tissue type) and next most likely tissue component (class) within each voxel.

The program also provides a *discrete classification*. The discrete classification uses the same discriminant analysis as the continuous method but the computer is forced to make a choice regarding the class in which each voxel belongs. Each tissue class is then assigned a distinct grey scale level, whereas in the continuous classification each class is assigned an intensity range.

Subjects

All subjects were recruited at the Section of Psychiatry, Dept. of Clinical Neuroscience, Karolinska Hospital. They had either taken part in previous studies as healthy subjects or they were staff members of the Psychiatry Section. One subject was a schizophrenic patient. Age range for all subjects was 18–56 years.

The same criteria were adopted as for all subjects who participate in our MRI studies. Inclusion criteria were informed consent. Exclusion criteria were metal depositions in the body, pregnancy or current or past treatment for a psychiatric disorder according to DSM-IV criteria (American Psychiatric Association, 1994) (except for schizophrenia spectrum disorder for the one patient that was included). Alcohol or narcotics dependence/abuse (DSM-IV), a history of cancer, organic brain disease or brain trauma or any other significant somatic disease that may affect brain function were also exclusion criteria.

The study was approved by the IRB at the Karolinska Institute and conducted according to the 1964 Declaration of Helsinki. All subjects provided written informed consent.

MR data acquisition

The subjects were investigated in a 1.5 T GE Signa Echospeed MR scanner (Milwaukee, Wis. USA) system at the MR Research Center, Karolinska Hospital, Stockholm, Sweden. Selected pulse sequences were used in one scanning session with the total duration of approximately 45 minutes. A fast spin echo (T2-weighted) axial oblique sequence was used for clinical evaluation. For classification, T1-weighted images, using a spoiled GRASS sequence, were acquired with the following parameters: 1.5 mm coronal slices, 35 degrees flip angle, TR 24 ms, TE 6.0 ms, 2 NEX, 24 cm FOV, and 256 x 192 acquisition matrix. The T2-weighted images were acquired with the following parameters: 2.0 mm coronal slices, TR 6000 ms, TE 84 ms, 2 NEX, 24 cm FOV and acquisition matrix 256 x 192. From visual inspection, all scans were judged to be excellent without any obvious motion artifact. The scan data were stored on magnetic tape and on CD-ROM.

Post acquisition image processing

The scans were transferred from CD-ROM for processing on Silicon Graphics O_2 workstations. First the T1 scan was loaded and resampled such that the interhemispheric fissure was aligned vertically in the axial and coronal views and the line between the anterior commissure (AC) and posterior commissure (PC) was aligned horizontally in the sagittal view. The alignment of the interhemispheric fissure defined the left/right hemisphere division. The bounding box for the cerebrum and AC and PC points are defined and are used to warp the Talairach grid (Talairach and Tournoux 1988) to the current brain. The T2 weighted images were fit to the resampled T1 weighted images using the AIR program (Woods et al. 1998). Once the quality of the fit was verified by visual inspection, the T2 data set was resampled to provide two data sets which are spatially registered with the voxel size 1.0156 x 1.0156 x 1.0156 mm^3.

The program searches the brain for pure samples from each tissue class (GM/WM/CSF) for use in the discriminant analysis. The search is based on finding plugs which have a low variance, and assuming that pure tissue samples will have a minimal variance in all imaging modalities used for classification. The search for pure samples of tissue is done in a conservative estimate of the portion of the 3D volume that corresponds to brain tissue (i. e., GM, WM and CSF), (Andreasen et al. 1996) using the Talairach Atlas. Since the amount of venous blood is very small, the operator is required to pick samples of venous blood. The sample of venous blood is typically chosen from the sigmoid sinus on several slices from both the left and right hemispheres. Once the venous blood samples are generated, the traces, the tissue plugs and the image are used as input into a discriminant analysis which includes spatial information that can correct for variation in signal intensity using a second order polynomial fit. After the discriminant function is obtained for the tissue samples, it is used to classify the entire image. The generated image is a *continuous* (fuzzy) representation of the tissue types. An 8 bit grey scale image is generated ranging from 0 to 255 with a voxel signal intensity of 10 representing pure CSF, 130 pure grey matter and 250 pure white matter. For example 190 is a voxel containing 50 % grey matter and 50 % white matter. Venous blood is coded as 1 and other material such as air is coded as 0. In the *discrete classification*, the tissue classes are assigned one distinct grey scale level (0, 1, 10, 130 or 250).

The intracranial tissue class volumes in the left and right hemisphere were measured separately. The ICV measure included the cerebellum and the upper part of the brainstem. To define the ICV, the program has to cut out ("deskull") the brain tissue and sulcal CSF from the dura and extradural tissue. For this purpose, the BRAINS program uses an artificial neural network (ANN) (Magnotta et al. 1999a). ANNs are massively parallel arrays of simple processing units that can be used for computationally complex tasks such as image processing, machine and computer vision. ANNs are well suited for making decisions such as voxel classification (i. e., whether a voxel is or is not part of the structure in question) in a rapid and robust manner. In BRAINS, the continuous tissue classified image is used as input into the neural net structure identification module. The neural net has been trained based on a human operator's definition of a structure and was taught to identify the brain (Fig. 1). Andreasen's

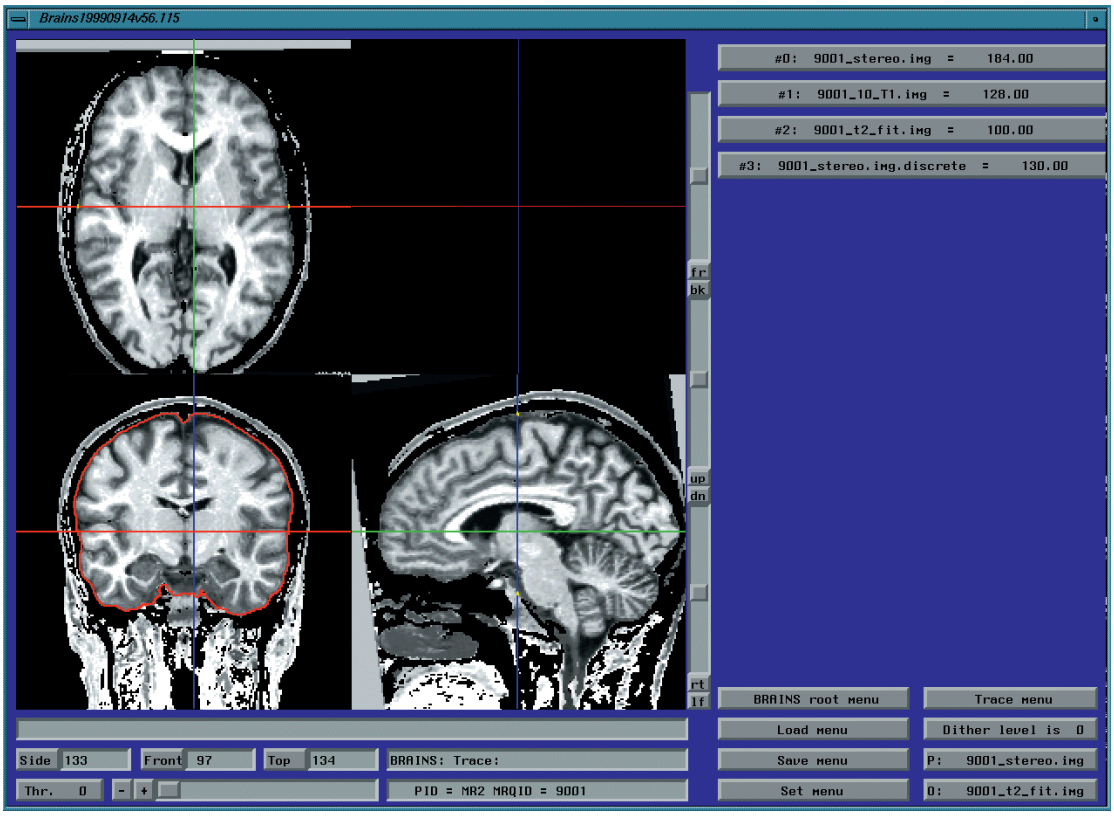

Fig. 1 The figure demonstrates a segmented human brain in the axial (upper left), coronal (lower left) and sagittal (lower right) projections. The red outline is an artificial neural network trace separating the intracranial volume from the extracranial tissue. This segmentation is continuous and visually resembles the grey scale found on T1-weighted MR images.

group follows the traces from the vertebral arteries to define the lower limit of the ICV. Since we found it difficult to reliably identify the arteries, we chose the following approach to define the inferior limit of the ICV. The mid-sagittal section was identified on the continuous image. A straight line was drawn between the lowest tip of the clivus and the lowest point of the occipital bone. The cut off of the brainstem was made on coronal sections using the guidelines (intersections) from the traces from the mid-sagittal section. This level of the inferior border corresponds to the foramen magnum.

Interoperator reliability

MR scans from ten subjects in the age range of 20–46 years (mean age LSD was 30.3 ± 8.9 years, median age was 28.5 years) were independently segmented by two operators IA and GO who were also blind to subject identity and diagnosis. Four subjects were women. One subject was a schizophrenic male patient and the other nine were healthy controls. All scans were obtained within a 4 week period (between Oct. 19 and Nov. 19, 1999).

Intraoperator reliability

The same ten MR scans were segmented on a second occasion by GO, who was blind with respect to subject identity and diagnosis, using the same procedure. The time interval between the first and second segmentation by the same operator were in the range of 52–74 days (mean number of days ± SD was 62.3 ± 7.0, median number of days was 59.5). All first and second segmentations were made between Nov. 8, 1999, and Jan. 17, 2000.

Scan-rescan reproducibility

Eleven healthy control subjects different from the ones that were used for the inter- and intraoperator reliability studies were scanned twice. The operator was blind to scanning order, subject identity and diagnosis. Five of the subjects were women. The age range of these subjects was 22–56 years, mean age ± SD was 37.5 ± 11.7 years, median age was 33 years. The time interval between investigations were 6–77 days, mean interval ± SD was 35.3 ± 25.3 days, median interval was 36 days. All scans were obtained between Nov. 26, 1999 and June 9, 2000.

Statistical analysis

The intraclass correlation coefficient (r) (Shrout et al. 1979) was used as a reproducibility index.

Results

All MR scans were evaluated by a clinical neuroradiologist. No subject was found to have a major organic brain pathology on the MRI investigation.

Interoperator reliability

Table 1a presents the mean volumes and range of the intracranial tissues classified either as GM, WM or CSF, left and right side, continuous and discrete classification,

Table 1a Interoperator reliability. Mean tissue volume and range in cubic centimeters (cc) for intracranial grey matter, white matter and CSF, left and right side, continuous and discrete classification for the first (1) and second (2) operator

Tissue class	Continuous			Discrete		
	average	min	max	average	min	max
Grey left						
1	384.7	315.7	463.7	400.7	333.9	487.5
2	386.8	315.1	468.8	405.8	333.2	496.0
Grey right						
1	382.5	324.9	455.5	402.7	347.9	486.4
2	382.9	324.4	451.5	406.3	349.0	482.6
White left						
1	221.9	183.0	277.6	252.4	202.9	311.0
2	221.5	183.9	278.9	249.7	203.4	309.2
White right						
1	234.9	189.1	299.6	261.3	205.5	326.3
2	233.9	189.6	299.5	258.3	207.0	323.5
CSF left						
1	92.2	68.7	137.1	45.8	21.6	95.6
2	92.0	66.5	138.9	44.8	21.8	88.4
CSF right						
1	91.4	76.6	141.3	44.9	20.9	97.4
2	91.3	75.3	141.4	43.6	20.6	88.1

Table 1b Interoperator reliability. Intraclass correlations for intracranial grey matter, white matter and CSF tissue volumes, left and right side, continuous and discrete classification, between the first and second operator

Tissue class	Continuous	Discrete
Grey		
left	0.993	0.974
right	0.992	0.969
White		
left	0.997	0.991
right	0.999	0.992
CFS		
left	0.996	0.995
right	0.998	0.992

first and second operator. Table 1b presents the intraclass correlations for the different tissue class volumes obtained by the two independent operators. Intraclass correlations were above 0.99 ($r^2 > 0.99$) for the continuous and discrete measures except for left and right discrete measures, which were slightly lower ($r^2 > 0.96$).

Intraoperator reliability

Table 2a presents the mean volumes and range of the intracranial tissues classified either as GM, WM or CSF, left and right side, continuous and discrete classification, as obtained by a single operator on two different occasions. Table 2b presents the intraclass correlations for the different tissue class volumes from the two operations, same operator. The results are presented with and without the exclusion of outliers. Intraclass correlations for both the continuous and discrete were excellent ($r^2 > 0.99$).

Table 2a Intraoperator reliability. Mean tissue volume and range in cubic centimeters (cc) for intracranial grey matter, white matter and CSF, left and right side, continuous and discrete classification, obtained by a single operator who classified the same scan on a first (1) and a second (2) occasion

Tissue class	Continuous			Discrete		
	average	min	max	average	min	max
Grey left						
1	384.7	315.7	463.7	400.7	333.9	487.5
2	384.5	318.7	457.9	401.7	337.9	482.4
Grey right						
1	382.5	324.9	455.5	402.7	347.9	486.4
2	382.4	321.9	455.3	403.8	344.3	487.0
White left						
1	221.9	183.0	277.6	252.4	202.9	311.0
2	221.1	183.7	276.5	250.7	203.5	308.3
White right						
1	234.9	189.1	299.6	261.3	205.5	326.3
2	234.1	188.7	299.9	259.8	205.3	325.5
CSF left						
1	92.2	68.7	137.1	45.8	21.6	95.6
2	92.7	70.0	138.0	45.9	20.9	93.7
CSF right						
1	91.4	76.6	141.3	44.9	20.9	97.4
2	92.3	76.9	141.6	45.1	21.5	96.4

Table 2b Intraoperator reliability. Intraclass correlations for intracranial grey matter, white matter and CSF tissue volumes left and right side, continuous and discrete classification for one operator who classified the same scans on two occasions

Tissue class	Continuous		Discrete	
		without outliers		without outliers
Grey				
left	0.996	0.993	0.994	0.991
right	0.997	0.993	0.996	0.993
White				
left	0.998	0.996	0.997	
right	0.999	0.997	0.998	
CFS				
left	0.998		0.999	
right	0.997		0.999	

Scan-rescan reproducibility

Table 3a presents the mean tissue volumes and range of the intracranial tissues classified either as GM, WM or CSF, left and right side, continuous and discrete classification, at the first (1) and second (2) MR scan occasion. Table 3b presents the intraclass correlations for the different tissue class volumes obtained at the first and the second scan occasion. Scan-rescan reproducibility for both the continuous and discrete classifications were excellent ($r^2 > 0.97$).

Proportions of relative tissue class volumes

For the first set of scans, values (in percent of the intracranial volume) were 86.95 % total brain tissue volume (54.5 % GM, 32.5 % WM) and 13 % CSF (continu-

Table 3 a Scan-rescan reproducibility. Mean tissue volume and range in cubic centimeters (cc) of intracranial grey matter, white matter and CSF tissue volumes, left and right sides, continuous and discrete classification, at the first scan (1) and second scan (2) occasion

Tissue class	Continuous			Discrete		
	average	min	max	average	min	max
Grey left						
Scan 1	381.3	322.4	426.0	401.2	340.9	456.6
Scan 2	379.8	320.6	428.0	399.3	341.2	455.4
Grey right						
Scan 1	373.7	317.1	417.7	396.7	337.6	448.3
Scan 2	374.9	321.1	418.6	397.8	344.5	455.9
White left						
Scan 1	213.6	190.2	242.9	240.7	211.3	278.0
Scan 2	213.5	188.8	242.2	240.7	206.6	277.4
White right						
Scan 1	227.8	194.7	258.3	250.5	213.0	287.8
Scan 2	227.6	193.0	258.0	250.6	208.7	287.3
CSF left						
Scan 1	95.6	77.3	116.9	48.3	33.2	66.6
Scan 2	93.3	72.5	114.1	46.8	29.7	63.6
CSF right						
Scan 1	94.0	71.0	110.7	48.3	32.1	65.9
Scan 2	93.9	75.0	116.2	47.7	30.6	68.0

Table 3 b Scan-rescan reproducibility. Intraclass correlations for intracranial grey matter, white matter and CSF tissue volumes within the intracranial volume, left and right side, continuous and discrete classification between the first and second scan occasion

Tissue class	Continuous	Discrete
Grey		
left	0.993	0.991
right	0.981	0.976
White		
left	0.968	0.988
right	0.989	0.992
CSF		
left	0.984	0.982
right	0.975	0.983

ous) and 93.5% total brain tissue volume (57.1% GM, 36.5% WM) and 6.4% CSF (discrete). For the second set of scans values were 86.3% total brain tissue volume (54.5% GM, 31.8% WM) and 13.7% CSF (continuous) and 93.1% total brain tissue volume (57.6% GM, 35.5 WM) and 7% CSF (discrete).

Discussion

Reliability and reproducibility – comparison with results from Andreasen's group

The inter- and intraoperator reliability as well as the scan-rescan reproducibility were found to be excellent (inter- and intrareliability slightly superior to scan-rescan reproducibility) for both the discrete and the continuous segmentation. The discrete and continuous classification appeared equally reliable although the continuous classification had a slightly higher intraclass r^2 for the right and left hemisphere GM compared with the discrete classification. Andreasen's group, who did not report values for right and left hemisphere separately, also found equal excellence for inter- and intrareliability for both classification types (Harris et al. 1999). They measured total CSF (internal and sulcal) and total GM and WM. The reliability of the automated classifications (both inter- and intraoperator reliability) were in the 0.99 range which was also the case for the majority of our measures. However, Andreasen's group reported that the continuous segmentation was marginally superior in scan-rescan data with all correlations above 0.9. From the reliability of the current study, we could however not confirm the superiority of one classification method to the other. The reliability for smaller regions in the brain was not tested in the present study and may yield different results.

Proportions of relative and absolute tissue class volumes

The proportional distribution (relative quantities) of GM, WM and CSF tissue class volumes for the set of ten scans used for the determination of inter- and intraoperator reliability and for the set of eleven scans used for scan-rescan reproducibility was considered for both types of classifications. There was less than 1% difference in the volume measures from the first set of scans compared with the second set. However, there was a notable difference between the volumes obtained with the continuous and discrete classification, most prominent for the CSF. In our study, for both sets of scans, the total brain tissue volume was 7% larger and the CSF volume correspondingly 7% smaller by the discrete classification as compared with the continuous classification. As shown in Tables 1a, 2a and 3a, the 7% difference in the CSF/total brain tissue proportions between the two classifiers was apparent in the absolute values as a 50% difference for the CSF volume and 3–5% and 10–12% for GM and WM volumes, respectively. The reason for this discrepancy is likely to be the partial voluming effect. CSF in narrow sulci has a tendency to become classified as GM, when both classes are contained in the same voxel. Since also the white matter volume was similarly classified to be larger in the discrete classification relative to the continuous classification, this may be true for the grey/white matter border as well. The continuous classification should, hypothetically, give more accurate measurements of tissue volumes. The reason for this is that a segmentation method that creates homogeneous segments such as the discrete segmentation excludes the possibility to detect "partial voluming" effects. Such effects can make an important contribution to the total volume particularly of CSF. However, since both classification methods demonstrated excellent reliability and since the validity of the tissue class volume measurements in the current study was not tested, we were not able to claim one method's superiority over the other.

We also did not measure the ventricular CSF and sulcal CSF separately, and therefore we do not know if the difference predominately occurred for sulcal CSF or not. This will be clarified in future analyses.

Comparison of tissue proportions with results from Andreasen's group

Andreasen's group reported the proportion of different intracranial tissue class volumes from the discrete classification only. They reported 55.4% GM, 35.2% WM and 9.3% CSF (Harris et al. 1999). Absolute volumes were not reported. The small discrepancies compared with our results from the discrete classification may be accounted for by the fact that Harris et al. examined 37 healthy control subjects of whom 11 individuals were 50–90 years of age, which is above our selected age range of 18–50 years. This may explain the comparatively larger CSF volumes found by Harris et al. Mean age for their subjects was 40.7 ± 17.3 years which is above the mean ages for the two sets of subjects in our study, which were 30.3 ± 8.9 and 37.5 ± 11.7 years, respectively. The fact that we chose a different definition of the inferior limit of the ICV may also have partly influenced the results. The gender distribution of the 37 subjects examined by Harris et al. was not reported. It is unlikely, however, that differences in gender distribution could account for the discrepancy since gender differences in global grey and white matter tissue proportions have not been found using the same segmentation method (Nopoulos et al. 1999).

Differences in scanning protocol and pulse sequences may also account for part of the discrepancy between the results of our group and Harris et al. The main differences were that we used thinner slices for T2 scanning (two-mm slice thickness compared with 3–4 mm) and did not include proton density images which Andreasen's group did. There may also be differences in performance between MR scanners, which we tried to minimize by using similar GE instrumentation (Signa). Presently, we are studying the same brains at the different centers using the same MR protocol. It is also not exactly clear how the proportional relationships between different tissue types in the brain are affected by higher resolution scans, choice of pulse sequences, thinner slice thickness, and even smaller voxel size. The variability in this respect remains to be determined.

Validity

We did not test the validity of the segmented tissue volume measurements. There are indirect ways to consider this issue. In a previous article, Harris et al. (1999) presented validity assessments by using indirect indicators such as sensitivity of the method to detecting changes associated with aging and agreement between the automated segmentation values and those produced through manual segmentation of tissue types. They found that the continuous segmentation had a slightly lower percent agreement with the "gold standard" of manual segmentation as well as a slightly weaker performance in age regression measures. The absolute volume measures that we acquired were in agreement with the findings of Andreasen's group. The proportion between grey and white matter is also in agreement with previous MRI and post mortem findings (Miller et al. 1980). Few groups report the raw values of segmented tissue volumes. This however would be of considerable interest, since one way to test the validity is to compare results across studies that use similar classification techniques.

Limitations

In the current study, we only addressed the reliability for the grey/white/CSF segmentation of the ICV within each hemisphere. Reliability assessments of smaller brain regions cannot be expected to yield as good of results. We also did not differentiate between ventricular and surface CSF, which is planned for a subsequent analysis.

Advantages

An advantage of the BRAINS program is its utility in the collection of large amounts of MR data since it is a highly automated procedure. Previous methods that have used manual deskulling have been extremely time consuming with additional problems of operator's drift and inconsistency. The current method combines high-speed image processing with high precision and reliability of measurements. If errors are introduced they tend to be systematic. Therefore, in future studies, the issue of validity should be specifically addressed.

Conclusion

Inter- and intraoperator reliability and scan-rescan reproducibility of the segmented intracranial tissue volume measures were found to be excellent. The brainstem cut off at the level of the foramen magnum gives an anatomically correct measure of the intracranial volume and is consistent across scans. The volume results that we report are considered to be valid. The proportional intracranial distributions of GM, WM and CSF obtained are very close to those obtained by Andreasen's group although the investigations were made at geographically different research sites on different MR scanners and on different subject groups. The results of our study support the excellent reproducibility of this method to obtain morphological data on the brain in schizophrenia and healthy individuals. The results from the current study support the versatility of combining results from these two different sites in the same database.

Acknowledgements Prof. Nancy Andreasen, MD, Ph. D., and her research group are acknowledged for excellent support and advice on the implementation of BRAINS. Lena Brandt, BSc., is acknowledged for statistical help and useful advice.

References

1. American Psychiatric Association (1994) Diagnostic and Statistical Manual of Mental Disorders, 4th edn. American Psychiatric Association Press, Washington, DC
2. Andreasen NC, Cohen G, Harris G, Cizaldo T, Parkkinen J, Rezai K, Swayze VW (1992b) Image processing for the study of brain structure and function: problems and programs. J Neuropsychiat Clin Neurosci 4:125–133
3. Andreasen N, Cizaldo T, Harris G, Swayze V, O'Leary DS, Cohen G, Ehrhardt J, Yuh WTC (1993a) Voxel processing techniques for the antemortem study of neuroanatomy and neuropathology using MRI. J Neuropsychiat and Clin Neurosci 5:121–130
4. Andreasen NC, Rajarethinam R, Cizaldo T, Arndt S, Swayze VW II, Flashman LA, O'Leary DS, Ehrhardt JC, Yuh WTC (1996) Automatic atlas-based volume estimation of human brain regions from MR images. J Comp Ass Tomography 20:98–106
5. Cohen G, Andreasen NC, Alliger R, Arndt S, Kuan J, Yuh WT, Ehrhardt J (1992) Segmentation techniques for the classification of brain tissue using magnetic resonance imaging. Psychiat Res: Neuroimaging 45:33–51
6. Shrout P, Fleiss J (1979) Intraclass correlations: use in assessing rater reliability. Psychol Bull 86:420–429
7. Harris G, Andreasen NC, Cizaldo T, Bailey JM, Bockholt HJ, Magnotta VA, Arndt S (1999) Improving tissue classification in MRI: a three-dimensional multispectral discriminant analysis method with automated training class selection. J Comp Ass Tomography 231:144–154
8. Magnotta VA, Heckel D, Andreasen NC, Cizadlo T, Corson PW, Ehrhardt JC, Yuh WT (1999a) Measurement of brain structures with artificial neural networks: two-and three-dimensional applications. Radiology 211:781–790
9. Magnotta VA, Andreasen NC, Schultz SK, Harris G, Cizadlo T, Heckel D, Nopoulos P, Flaum M et al. (1999b) Quantitative in vivo measurement of gyrification in the human brain: changes associated with aging. Cerebral Cortex 9:151–160
10. Miller A, Alston R, Corsellis J (1980) Variation with age in the volumes of grey and white matter in the cerebral hemispheres of man: measurements with an image analyzer. Neuropathol Appl Neurobiol 6:119–132
11. Nopoulos P, Flaum M, O'Leary D, Andreasen NC (1999) Sexual dimorphism in the human brain: evaluation of tissue volume, tissue composition and surface anatomy using magnetic resonance imaging. Psychiatry Res: Neuroimaging Sec 98:1–13
12. Talairach J, Tournoux P (1988) Co-planar Stereotaxic Atlas of the Human Brain. Thieme Medical, New York
13. Woods RP, Grafton ST, Watson JD, Sicotte NL, Mazziotta JC. (1998) Automated image registration: II. Intersubject validation of linear and nonlinear models. J Comp Ass Tomography 22:155–165