

NIH -- W1 PS358

ELIZABETH MOLLOY
Building 10 Room 4C110
10 Center Drive
Bethesda, MD 20892

ATTN:	SUBMITTED:	2001-10-01 10:51:27
PHONE: 301-435-4515	PRINTED:	2001-10-04 08:33:43
FAX: 301-480-8898	REQUEST NO.:	NIH-10031906
E-MAIL:	SENT VIA:	LOAN DOC 4513850

```

-----
NIH                               Fiche to Paper       Journal
-----
TITLE:                            PSYCHIATRY RESEARCH
PUBLISHER/PLACE:                   Elsevier/North-Holland Biomedical Press Limerick :
VOLUME/ISSUE/PAGES:               1995 Aug 8;61(2):113-9      113-9
DATE:                              1995
AUTHOR OF ARTICLE:                 Giedd JN; Kozuch P; Kaysen D; Vaituzis AC; Hamburger SD; Bar
TITLE OF ARTICLE:                  Reliability of cerebral measures in repeated exami
ISSN:                              0165-1781
OTHER NOS/LETTERS:                Library reports holding volume or year
                                   7911385
                                   7480388
SOURCE:                             PubMed
CALL NUMBER:                       W1 PS358
REQUESTER INFO:                    ELIZAMOLLOY
DELIVERY:                           E-mail: emolloy@codon.nih.gov
REPLY:                              Mail:

```

NOTICE: THIS MATERIAL MAY BE PROTECTED BY COPYRIGHT LAW (TITLE 17, U.S. CODE)

-----National-Institutes-of-Health,-Bethesda,-MD-----



Reliability of cerebral measures in repeated examinations with magnetic resonance imaging

Jay N. Giedd^{*a}, Patricia Kozuch^a, Debra Kaysen^a, A. Catherine Vaituzis^a,
Susan D. Hamburger^b, John J. Bartko^b, Judith L. Rapoport^a

^aChild Psychiatry Branch, National Institute of Mental Health, Building 10, Room 6N240, 9000 Rockville Pike, Bethesda, MD 20892, USA

^bDivision of Epidemiology and Services Research, National Institute of Mental Health, 9000 Rockville Pike, Bethesda, MD 20892, USA

Received 2 February 1994; revision received 19 July 1994; accepted 2 November 1994

Abstract

This study examined the reliability of quantitative measures of cerebral magnetic resonance images (MRI) in repeated scans. Ten subjects were scanned twice, at 2- to 4-week intervals. Volumetric data from 14 regions of the cerebrum, the caudate nucleus, and the lateral ventricles and area measures of the corpus callosum were acquired. Intrarater and scan-rescan reliabilities, including the relative percent error from each of these two sources, were determined for each structure. Intraclass correlations ranged from 0.88 for the head of the caudate nucleus to 0.99 for the ventricular volume. Quantitative cerebral MRI measures of these structures are stable over time intervals of 2-4 weeks.

Keywords: Brain, morphology; MRI rescan reliability; Caudate nucleus; Cerebrum; Corpus callosum

1. Introduction

Quantitative cerebral magnetic resonance imaging (MRI) often relies on the assumption that differences in measures of brain structure size are not due to methodological factors or normally occurring physiological fluctuations such as hydration status which may produce transient changes in morphology. The validity of this assumption is important for longitudinal developmental studies of

neuroanatomy and studies of disease progression or resolution.

Prior studies of reliability of cerebral MRI measures have yielded mixed results. Good intrascanner and interscanner reliabilities have been reported for the hippocampus and other brain structures on subjects scanned two or three times on the same or different machines (Bartzokis et al., 1993). A multicenter MRI study quantifying brain, liver, and skeletal muscle also concluded that quantitative measures could be reliably reproduced given adequate control of machine calibration and measurement technique (de Certaines et

* Corresponding author, Tel: +1 301 496-6081; Fax: +1 301 402-0296.

1993). The Consortium to Establish a Registry for Alzheimer's Disease did not find satisfactory inter-rater agreement by experienced neuroradiologists in interpreting MRI findings acquired from different centers with different imaging parameters, although the extent to which this reflected quantitative differences in the images themselves is unclear (Davis et al., 1992). Plante and Turkstra (1991) have addressed several sources of bias in the quantitative analysis of MRI scans, including partial volume effects, head tilt, plane of view, use of noncontiguous slices, contrast/intensity manipulations, and magnetic field inhomogeneities. The present study sought to minimize these sources of error and assess the stability of measures from scans of 10 healthy adults acquired 2–4 weeks

apart. Partial volume effects were minimized by using thin slices (1.5 mm), head tilt was minimized by using standardized external landmarks, different planes of view were used to assess structures along their longest axis, contiguous slices were used, contrast and intensity were held constant, and the same General Electric 1.5 Tesla Signa scanner (General Electric Medical Systems, Milwaukee, WI) was used to minimize magnetic field inhomogeneities.

Structures of clinical interest were chosen to encompass a variety of sizes, shapes, tissue characteristics, and ease of measurement, including measurements from the sagittal, axial, and coronal planes. Total intraclass correlation coefficients (ICCs) and the relative contributions of

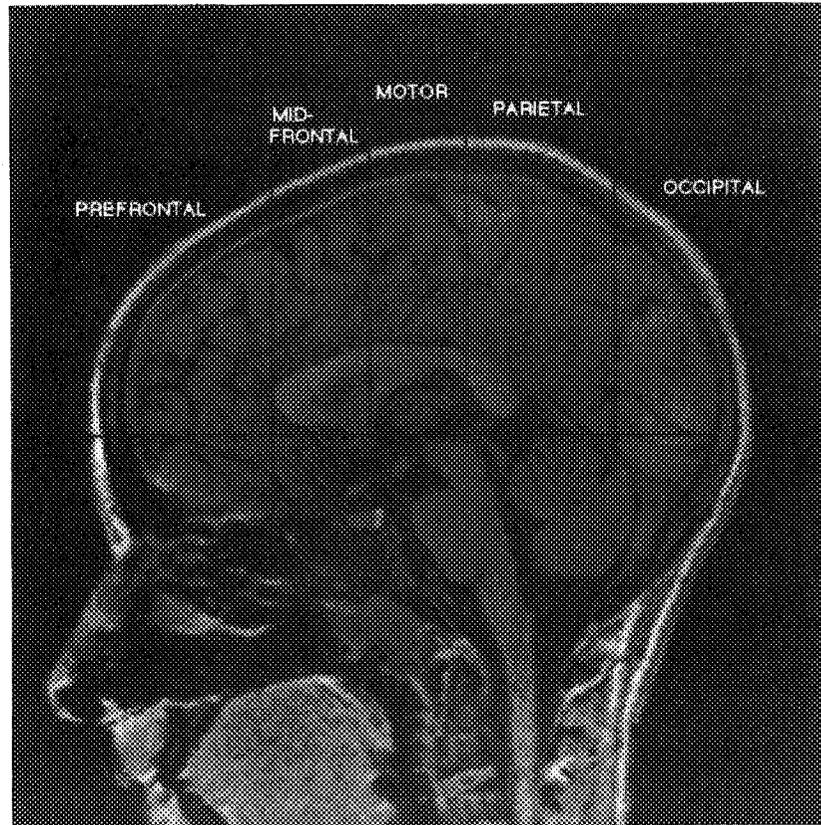


Fig. 1a. Midsagittal view showing boundaries used to demarcate cerebral regions

scan-rescan and intrarater measurement error were determined for the size and symmetry index of each structure.

2. Methods

2.1. Subjects

Ten healthy adult volunteers were recruited from the community (6 males, 4 females, mean age = 33.2 years, SD = 10.5). All were screened for contraindications to MRI scanning including a history of a metallic foreign body in the eye, cardiac pacemaker, aneurysmal or other surgical clips, shrapnel, cochlear implants, pregnancy, or severe claustrophobia. Written informed consent for participation in the study was obtained. Sub-

jects were scanned on a GE 1.5 Tesla Signa scanner located at the National Institutes of Health Clinical Center.

2.2. MRI protocol

Vitamin E capsules, wrapped in gauze and placed in the meatus of each ear, were used to help standardize head placement. A third capsule was taped to the lateral aspect of the left inferior orbital ridge. Vitamin E capsules emit a characteristic signal with our scanning parameters, and the three capsules were used to define a reference plane for our images. The patient's head was aligned in a head holder so that a narrow guide light passed through each of the vitamin E capsules. Foam padding was placed on both sides of the pa-

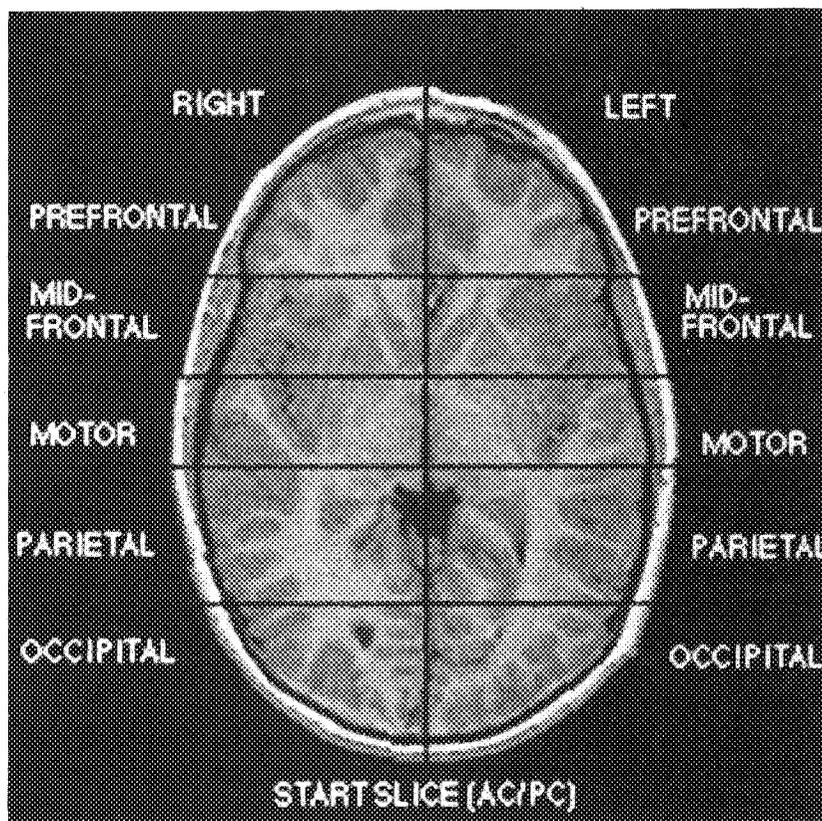


Fig. 1b. Axial view showing boundaries demarcating cerebral regions. A line drawn at the level of the anterior and posterior commissures (AC, PC) separated the left and right hemispheres. Perpendicular lines drawn at the anterior-most point of the genu of the corpus callosum, the AC, and the PC demarcated prefrontal, midfrontal, and motor regions. A perpendicular line drawn 1.5 times the AC-PC distance behind the PC line divided the posterior areas into parietal and occipital regions.

tient's head to minimize head movement. A multi-echo sagittal localizing plane was acquired (slice thickness = 5 mm, interslice gap = 1.5 mm, echo time [TE] = 10 ms, repetition time [TR] = 400 ms, acquisition matrix = 256×128 , NEX = 1, field of view [FOV] = 30 cm) and from this a multi-echo axial series was acquired (slice thickness = 3 mm, TE = 10 ms, TR = 400 ms, acquisition matrix = 256×128 , NEX = 1, FOV = 24 cm) to assure that one of the axial slices contained all three of the capsules. If no slice contained all three capsules, the patient was realigned until this criterion was met. To control for orientation in all three axes, an additional alignment criterion was that the nose be centered at the "12 o'clock" position. Once patient alignment was established, a volumetric sagittal scan using three-dimensional spoiled gradient recalled echo in the steady state was obtained (slice thickness = 1.5 mm, TE = 5, TR = 24, flip angle = 45° , acquisition matrix = 256×192 , NEX = 1, FOV = 24 cm).

2.3. Image analysis

All scans were evaluated as free from gross abnormalities by a clinical neuroradiologist. Images were transferred to a Macintosh II FX computer workstation and analyzed with IMAGE (Rasband, 1993), a software package developed at the National Institutes of Health that is in the public domain. Raters were unaware of all subject characteristics and whether they were evaluating the first or the second scan.

A three-dimensional data set acquired from the volumetric sagittal series was used to generate images in the axial plane for assessment of the regional brain volumes and in the coronal plane for assessment of the caudate nucleus and ventricular volumes. The area of the corpus callosum was measured directly from the midsagittal series. Fig. 1 shows the boundaries for the regional brain volumes.

Consistent with the Talairach atlas technique, the anterior commissure (AC) and the posterior commissure (PC) were used as internal landmarks to position and provide reference points for subdivision of the brain. All of the planes used to demarcate boundaries were perpendicular to the AC-PC line. The grid described in Fig. 1 was

overlaid on all axial images in which brain matter was visible. A supervised thresholding technique, which separated brain matter from cerebrospinal fluid (CSF), was used to calculate area measures on each of the axial sections. Gray matter and white matter were not individually segmented. Area measures were summed and multiplied by the slice thickness (1.5 mm) to derive volumes of the various regions. All brain matter anterior to the genu of the corpus callosum was deemed *prefrontal*. The volume between the tip of the genu of the corpus callosum and the AC was designated as *midfrontal*. The volume between the AC and the PC was designated as *motor*. The volume between the PC and 1.5 times the AC-PC length posterior to the PC was designated as *parietal*, and all brain matter posterior to that plane was designated as *occipital*. It should be noted that these boundaries do not take into consideration any sulcal patterns or cytoarchitectonic information and thus should be interpreted as volumes that are *mostly* prefrontal, *mostly* midfrontal, and so on. Measures of total hemisphere volumes do not include brainstem, cerebellum, subdural cerebrospinal fluid, or ventricular volume.

The boundaries of the caudate nucleus in the coronal plane were defined posteriorly as the first slice in which the body of the caudate was clearly distinguishable from the surrounding white matter and anteriorly on every slice in which they were visible. The head and body of the caudate nucleus were separated by the interventricular foramina in the coronal plane as defined in *Gray's Anatomy* (Williams, 1989). The caudate nuclei were manually traced without thresholding techniques.

Ventricular areas were measured in the coronal plane with a thresholding technique that separated brain matter from CSF. Ventricular volume was derived by multiplying by the slice thickness (1.5 mm).

The corpus callosum was measured from a slice designated as midsagittal based on the presence of the septum pellucidum and patency of the cerebral aqueduct.

Each of the 20 scans (10 subjects scanned twice) was measured twice to determine intrarater reliability. The ICCs for the four ratings \times subjects data matrix were determined using a one-way

Table
Meas
Regio

Caud
Caud
Vent
Prefr
Midf
Mot
Pari
Occ
Tot

ana
Cau
use
squ
5) r
for
T
des
rati

Tabl
Intr
resc

Regio

Cau
Cau
Cor
Lat
Pref
Mid
Mot
Par
Occ
Tot

M

Table 1
Mean volume (in ml) of cerebral regions, caudate nucleus, and lateral ventricles

Region	Right volume				Left volume			
	Scan 1		Scan 2		Scan 1		Scan 2	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Caudate, head	6.2	0.23	6.2	0.23	5.9	0.19	6.0	0.26
Caudate, body	0.82	0.33	0.81	0.32	0.75	0.26	0.75	0.27
Ventricles	12.2	5.9	12.2	5.9	12.3	5.0	12.6	4.8
Prefrontal	81.3	9.5	80.3	8.7	79.1	8.5	78.4	6.9
Midfrontal	118.7	10.7	118.6	10.3	117.9	12.2	118.3	11.9
Motor	112.9	11.0	111.9	9.3	111.7	10.3	110.8	9.3
Parietal	166.2	17.6	163.2	16.6	165.0	17.2	163.3	14.7
Occipital	63.0	9.2	61.6	9.5	66.0	6.8	64.2	7.6
Total hemisphere	542.0	45.6	535.6	41.6	539.5	41.7	535.0	38.2

analysis of variance (ANOVA) model (Bartko and Carpenter, 1976). ICC computation requires the use of the between- and within-subjects mean squares from the ANOVA. Table 2 (columns 2 and 5) presents the regional ICCs of the four ratings for volume and symmetry measures.

The four ratings occurred in a scan \times ratings design, where the two scan sessions each had two ratings. Both rescanning and raters are a source of

variance. Arithmetically, the within-subjects variance from the ANOVA is composed of the sum of these two variance components. Table 2 (columns 3 and 4) presents the percent contribution from these two sources for the volume measures, while Table 2 (columns 6 and 7) presents that for the symmetry measures. The greater the ICC reliability, the smaller the within-subjects mean square. Because symmetry measures are particularly sensi-

Table 2
Intraclass correlation coefficients (ICCs) of volume and symmetry index measures and percent of within-subjects variance due to rescan and rater reliability components in 10 adults scanned at 2- to 4-week intervals

Region	Volume			Symmetry index		
	ICC	Rescan	Rater	ICC	Rescan	Rater
Caudate, head	0.884	14%	86%	0.645	25%	75%
Caudate, body	0.981	5%	95%	0.889	36%	64%
Corpus callosum	0.964	70%	30%	—	—	—
Lateral ventricle	0.998	90%	10%	0.945	35%	65%
Prefrontal	0.946	62%	38%	0.791	40%	60%
Midfrontal	0.961	33%	67%	0.781	38%	62%
Motor	0.961	54%	46%	0.824	13%	87%
Parietal	0.950	69%	31%	0.864	36%	64%
Occipital	0.905	35%	65%	0.927	12%	88%
Total hemisphere	0.969	74%	26%	0.938	33%	67%

Note. Symmetry index = $[(R - L) \times 2]/(R + L)$. The corpus callosum measures are midsagittal area (cm^2).

tive to variations in head tilt, a symmetry index, $[(R - L) \times 2]/(R + L)$, was calculated for all bilateral structures.

3. Results

Table 1 shows the means and SDs of the left and right volumes of structures for each of the scans. Table 2 shows the ICCs for the size and symmetry index measures of each of the regions and the percent of variance accounted for by for rescan and rater measurement effects.

For volume measures, ICCs ranged from 0.88 for the head of the caudate nucleus to 0.99 for the lateral ventricles. For symmetry measures, the range was from 0.65 for the caudate head to 0.95 for the lateral ventricles. No consistent pattern emerged regarding the relative contribution of error from the methodology of rescanning versus the error from rater measurement for structure size. For symmetry measures, rater effects were consistently higher than rescan effects. Given the high ICCs, however, neither source of error was substantial. Analysis of gender interactions revealed no particular effects of gender on reliability measures.

4. Discussion

Variations in measures of cerebral structures, whether due to actual changes in structure size, rater measurement error, or rescan methodology, were small. Effects of head positioning can be adequately controlled for in quantitative MRI studies, even without post-acquisition computerized realignment. Potential transient fluctuations in brain size related to hydration status or hormonal changes do not seem to be major confounding factors in quantitative MRI analysis. The rescan measurement error is well within the limits seen for many laboratory assays. When reliabilities are high, the number of subjects needed to observe a given change in volume over time is driven by the mean volume and variance of the structure under question. Examination of the mean volumes and variances of a variety of brain structures in our data set suggests that a general rule of thumb, with the use of paired *t* tests and a 95% confidence level,

is the need for about 15 subjects in each group to detect a 10% difference and about 60 subjects in each group to detect a 5% difference.

The power to detect a given percentage of change over time varies with the individual mean and variance of each structure. Analysis for left hemisphere volume is presented as an example. From Table 1, the effect size for *Left volume, Scan 1 vs. Scan 2, Total hemisphere* (last line of Table 1) is 0.42. With a statistical Type 1 error of 5%, an $n = 10$, and the paired *t* test, the power for detecting this effect size of 0.42 or a difference in means of 4.5 units is 0.25. To detect the effect size of 0.42 (power 80%, Type 1 error of 5%, paired *t* test), 40 subjects would be required. By themselves, the 10 subjects in the present study have 80% power to detect an effect size of 0.90 or a difference in means of 9.6 units.

The symmetry indexes, being derived from two measures, were less reliable, although the ICCs were still quite high (see Table 2). Previous investigators have noted that area measures, such as a ventricle to brain ratio calculated from a single "best" slice are less reliable than volumetric measures (Woods et al., 1991). For studies assessing right-left asymmetries in the brain, it seems especially critical to control for head positioning and to use volumetric measures whenever possible.

Overall, rescan reliabilities were high for a variety of brain structures acquired at 2- to 4-week intervals. This is an encouraging finding with regard to longitudinal studies of pathological groups and normal developmental studies.

Acknowledgment

The work reported was supported, in part, by a grant from the Gulton Foundation.

References

- Bartko, J.J. and Carpenter, W.T. (1976) On the methods and theory of reliability. *J Nerv Ment Dis* 163, 307–317.
- Bartzokis, G., Mintz, J., Marx, P., Osborn, D., Gutkind, D., Chiang, F., Phelan, C.K. and Marder, S.R. (1993) Reliability of in vivo volume measures of hippocampus and other brain structures using MRI. *Magn Reson Imaging* 11, 993–1006.

in each group to
out 60 subjects in
ence.

n percentage of
individual mean
Analysis for left
l as an example.
Left volume, Scan
st line of Table 1)
error of 5%, an
power for detect-
fference in means
effect size of 0.42
paired *t* test), 40
nemselves, the 10
ve 80% power to
fference in means

derived from two
hough the ICCs
2). Previous in-
neasures, such as
ted from a single
volumetric mea-
studies assessing
brain, it seems
head positioning
henever possible.
e high for a vari-
at 2- to 4-week
ng finding with
of pathological
al studies.

ted, in part, by a
on.

On the methods and
163, 307-317.

n, D., Gutkind, D.,
S.R. (1993) Reliabili-
ocampus and other
Reson Imaging 11,

Davis, P.C., Gray, L., Albert, M., Wilkinson, W., Hughes, J., Heyman, A., Gado, M., Kumar, A.J., Destian, S., Lee, C., Duvall, E., Kido, D., Nelson, M.J., Bello, J., Weathers, S., Jolesz, F., Kikinis, R. and Brooks, M. (1992) The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part III. Reliability of a standardized MRI evaluation of Alzheimer's disease. *Neurology* 42, 1676-1680.

de Certaines, J.D., Henriksen, O., Spisni, A., Cortsen, M. and Ring, P.B. (1993) In vivo measurements of proton relaxation times in human brain, liver, and skeletal muscle: a multicenter MRI study. *Magn Reson Imaging* 11, 841-850.

Plante, E. and Turkstra, L. (1991) Sources of error in the quantitative analysis of MRI scans. *Magn Reson Imaging* 9, 589-595.

Rasband, W. (1993) *IMAGE (1.6)*. National Institutes of Health, Bethesda, MD.

Williams, P.L. (1989) *Gray's Anatomy*. Churchill Livingstone, New York.

Woods, B.T., Douglass, A. and Gescuk, B. (1991) Is the VBR still a useful measure of changes in the cerebral ventricles? *Psychiatry Res: Neuroimaging* 40, 1-10.