

## Reliability analysis of hippocampal MRI volumetry at 3 Tesla

J. F. Jansen<sup>1</sup>, C. R. Jeukens<sup>2</sup>, M. C. Vlooswijk<sup>2</sup>, H. J. Majoie<sup>2</sup>, M. C. de Krom<sup>2</sup>, A. P. Aldenkamp<sup>2</sup>, P. A. Hofman<sup>2</sup>, and W. H. Backes<sup>2</sup>

<sup>1</sup>Department of Medical Physics & Radiology, Memorial Sloan-Kettering Cancer Center, New York, NY, United States, <sup>2</sup>Maastricht University Hospital, Maastricht, Netherlands

### Introduction

The hippocampus is an essential neuroanatomical part of the declarative long term memory system [1]. In a number of neurological disorders (including epilepsy [2]), a relation has been proposed between hippocampal volume changes and memory related function decline. Hippocampal volumes are usually determined on the basis of manual delineation on high-resolution three-dimensional T1-weighted MR image sets. Quantification of the hippocampal volume yields a more objective measure than visual inspection. Unfortunately, the hippocampal volumetry protocols to delineate the hippocampus often substantially differ between studies and are not reported in great detail [3]. This complicates the comparison and reproducibility of measurements between different research groups. The aim of this work was to provide a detailed reliability analysis for hippocampal volumetry of T1-weighted MR images acquired at 3T. Also, the validity of a standard method to correct the hippocampal volume for total brain size [4] will be critically assessed. The applicability of the protocol was tested in a study population of 60 subjects, comprising patients with epilepsy and healthy volunteers, using high resolution T1-weighted MR images.

### Materials and Methods

**Subjects** The study population included 40 patients with non-symptomatic epilepsy (20F, 20M, age 39±12y), and 20 healthy volunteers (11F 9M, age 40±13y). **MRI** MRI was performed with a 3.0-Tesla whole-body unit (Philips Achieva [software release 1.5.4.0], Philips Medical Systems, Best, The Netherlands). T1-weighted three-dimensional (3D) turbo field echo (TFE) images were acquired with the following parameters: repetition time (TR) 9.91 ms, echo time (TE) 4.6 ms, inversion time (TI) 3 s, flip angle 8°, matrix 256x256x160, field of view (FOV) 256x256x160 mm<sup>3</sup>, 1 mm adjacent coronal slices. **Volumetric Analysis** The hippocampal volumetry guidelines were adapted from Jack et al. [4] and Watson et al. [5]. T1w-Images were reformatted using customized software programmed in Matlab. All hippocampi were delineated by two image readers, who were blinded to all subject information, using MRIcron. The delineation was performed in the posterior to anterior direction on the reformatted slices. After agreement between the two readers on the position of the most posterior slice, each second slice was delineated. The hippocampal cross-sectional areas in the non-delineated slices were obtained by linear interpolation of the delineated slices. The hippocampal volume were calculated by multiplying the number of voxels by the voxel volume (1 mm<sup>3</sup>). The final volume of the hippocampus was determined as the mean of the volumes delineated by the two readers. **Reliability** Reliability measures served to determine the inter-reader agreement for the hippocampal volumes of the 60 subjects and the intra-reader agreement for the hippocampal volumes of 3 randomly chosen healthy subjects that were delineated twice by every reader. The following reliability measures were calculated: i) Intra-class Correlation Coefficient (ICC) which quantifies the measurement error relative to the biological variability among the subjects. ii) Interobserver Volume Difference (VD) =  $2 \cdot (V_A - V_B) / (V_A + V_B) \cdot 100\%$ , where  $V_A$  and  $V_B$  are the volumes delineated by reader A and B, respectively. The VD expresses how strong the volumes delineated by observer A and B relatively differ. iii) Overlap ratio (OR), which is defined as the ratio of the overlapping volume that both readers delineated and the entire volume delineated by any reader, and iv) Bland-Altman plot. In this plot the volume difference is displayed as a function of the average volume. The reliability criteria were  $ICC \geq 0.85$ ,  $VD \leq 15\%$ ,  $OR \geq 0.75$  and the criterion for the Bland-Altman plot is that there is no significant (Pearson) correlation between the volume difference and the average volume. Additionally, hippocampal volumes were corrected for total brain volume using the covariance method described by Jack et al [4]. The total intracranial volume was determined from the MR images by automatic segmentation using routines from SPM5.

### Results

The uncorrected, absolute hippocampal volumes (mean ± SD, in mm<sup>3</sup>) for the patients and healthy controls were (L 3923±500, R 3120±416) and (L 3185±411, R 3302±411), respectively. Corrected volumes were (L 2977±418, R 3187±351) and (L 3185±360, R 3302±323), for patients and controls, respectively. No significant differences between left or right, or patients and controls were observed. Table 1 shows the values of the inter-reader and intra-reader reliability measures calculated for the left and right hippocampus. OR and VD were calculated for each pair of delineations and the mean and SD are also listed in Table 1. All reliability measures met the criteria defined except for the ICC value of the intra-reader A. Figure 1 shows the Bland-Altman plot of all delineations made by the two readers. No significant correlation ( $r=0.07$ ,  $p=0.59$  (left) and  $r=-0.017$ ,  $p=0.9$  (right)) between the volume difference and mean volume was found. Figure 2 shows the volume of the left and right hippocampus as a function of total intracranial volume for the healthy control group with a linear regression of the data for the left hippocampus the slope was  $B = 0.0016$  ( $r = 0.48$ ,  $p=0.03$ ) and for the left hippocampus  $B = 0.0020$  ( $r = 0.62$ ,  $p=0.0037$ ). The data points show a relatively large spread which is reflected by the low  $r$  value. The variances of the corrected volumes are not significantly decreased as was determined using an F-test ( $\alpha < 0.05$ ).

### Discussion and conclusion

The hippocampal volumes that were determined in this study were in agreement with values found in literature [3]. The beneficial ICC, OR and VD values indicate that the reproducibility was high (Table 1). Several studies have advised to correct the hippocampal volumes for brain size, as larger brains on average have larger brain structures [4]. Although the brain size correction introduces an additional source of error because of the uncertainty in the regression, a study by Mathalon et al. showed that brain corrected volumes might give better correlations with age and diagnostic status [6]. Based on the results of this study the usefulness of the correction can be questioned for the following reasons. The (linear) relation between the total intracranial volume size and the hippocampal volume is weak (Figure 2). Also, an independent dataset consisting of 61 healthy subjects revealed a weak correlation ( $r=0.42$ ,  $p=0.0008$ ) [7]. Furthermore, application of the brain size correction did not significantly reduce the variances of the hippocampal volumes. Correction of the hippocampal volumes for total intracranial volume volume is not straightforward and that the correlation between the hippocampal volume and total intracranial volume should be critically assessed before using it for correction.

### References

[1] Marieb, Human Anatom & Phys 2004 Pearson Education; [2] Araujo, Epilepsia 2006 47:1354; [3] Geuze, Mol Psych 2005 10:147; [4] Jack, Radiology 1989 172:549; [5] Watson, Neurology 1992 42:1743; [6] Mathalon, Psych Res NeuroIm 1993 50:121; [7] Tisserand, Neurobiol of Aging 2000 21:569.

	n	Side	SD <sub>bs</sub>	SD <sub>ws</sub>	ICC	OR	VD (%)
Inter-reader	112	L&R	445	175	0.87	0.82 ± 0.03	6.5 ± 4.3
	56	L	466	178	0.87	0.82 ± 0.04	7.0 ± 4.7
	56	R	404	162	0.86	0.82 ± 0.03	6.0 ± 3.8
Intra-reader A	6	L&R	214	124	0.59	0.87 ± 0.01	4.3 ± 3.1
Intra-reader B	6	L&R	171	63	0.88	0.91 ± 0.01	1.8 ± 0.8

Table 1: Inter- and intra-reader variabilities and reliability measures for the left (L) and right (R) hippocampus. Notation: mean ± SD; n is the number of subjects. SD standard deviation, SD<sub>bs</sub> between-subjects SD, SD<sub>ws</sub> within-in subjects SD.

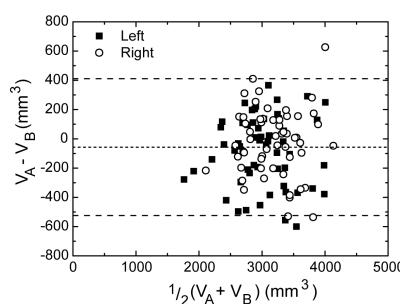


Figure 1: Bland-Altman plot showing the volume difference ( $V_B - V_A$ ) in mm<sup>3</sup> as a function of the mean volume ( $1/2(V_A + V_B)$ ) in mm<sup>3</sup> for the entire study population ( $n = 60$ ). Dotted lines: mean volume difference; dashed lines: 95 % confidence intervals.

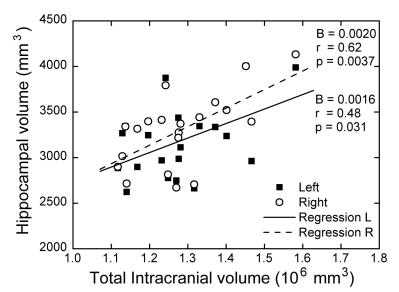


Figure 2: Left and right hippocampal volume as a function of the total intracranial volume for the healthy controls. Lines indicate linear regression of the data.