Published in final edited form as:
Neuroimage. 2009 November 15; 48(3): 497–498. doi:10.1016/j.neuroimage.2009.05.004.

Improving the Reliability of Manual and Automated Methods for Hippocampal and Amygdala Volume Measurements

Khader M. Hasan[1,CA] and **Otto Pedraza**[2]

[1] Department of Diagnostic and Interventional Imaging, University of Texas Health Science Center at Houston, Houston, Texas, USA

[2] Department of Psychiatry and Psychology, Mayo Clinic, Jacksonville, Florida, USA

Dear Editor,

We read with interest an article published recently in NeuroImage by Morey et al (2009). The paper is informative and timely as it tackles the issue of volume estimation of the hippocampus and amygdala which have been implicated in studies on the structural substrates of human brain gray matter development (Gogtay et al., 2006) and aging (Allen et al., 2005; Raz et al., 1997; Walhovd et al., 2005; Xue et al., 2008), cognition (Zimmerman et al., 2006) and disease (Bigler et al., 1997;Makris et al., 2002; McDonald et al., 2008; Plessen et al., 2006; Tae et al., 2008). The authors manually delineated the borders of the amygdala and hippocampus bilaterally on 20 healthy controls and compared the volumes obtained by two freely distributed segmentation packages (e.g. FreeSurfer and FSL-FIRST).

We believe the conclusion made by the authors in regards to the accuracy or "superiority" of one automated method over another automated approach has to be taken with scrutiny as the paper did not present some needed details to help interested readers.

First, Morey et al. (2009) did not provide basic relevant demographics such as gender, handedness and age range, mean and standard deviation of the 20 controls. Moreover, the authors did not report the sensitivity of the manual, FreeSurfer and FSL-FIRST normalized volumetry (volume per unit total brain volume) to important variables such as age. It is noteworthy that the dependence of hippocampus absolute or normalized volume on age across the healthy lifespan seems to be contradictory in published literature (Allen et al., 2005; Gogtay et al., 2006; Jernigan et al., 2005; Liu et al., 2003; Walhovd et al., 2005; Liu Zimmerman et al., 2008) and hence the paper could have documented and compared such age correlations. An automated method that provides systematically biased absolute values compared to a well-established gold standard may not necessarily be less sensitive to age (Walhovd et al., 2005) or to pathology effects (Tae et al., 2008).

Second, the authors did not consider two recent and relevant Neuroimage publications (Allen et al., 2008; Shattuck et al., 2008) that tabulated the hippocampus volumetry manually versus the values obtained using other methods. The absolute volume results

[CA]Corresponding Author: Khader M. Hasan, Ph.D., Associate Professor of Diagnostic and Interventional Imaging, Department of Diagnostic and Interventional Imaging, University of Texas Health Science Center at Houston, 6431 Fannin Street, MSB 2.100, Houston, Texas 77030, Tel: Office (713) 500-7690, Fax: (713) 500-7684, Email: Khader.M.Hasan@uth.tmc.edu.

obtained by Morey et al. (2009) via manual delineation were not tabulated for each of the 20 controls, and hence they could not be compared with published “manually-delineated” hippocampal volumetry literature nor can they be considered more accurate than what FreeSurfer or FSL provided. The absence of demographic data and manually delineated volumes thus renders it difficulty for a reader to gauge the extent to which the Morey et al. (2009) findings can be generalized with confidence.

Third, despite a sample of 20 healthy controls, scatter plots and degrees of freedom in the statistical analyses suggest a sample of 40. Therefore, right and left volume measurements were treated as independent measurements, which can be problematic given the high correlation between left and right volumetry of similar neuroanatomic structures. In such a scenario, the correlations and Bland-Altaman analysis between methods would be misleading. The information in the figures should have been presented on the mean values or even better separately or clearly marked so that linear model noise reflects only random effects.

Fourth, since there is supporting evidence that hippocampal volumes are right-ward asymmetric (Pedraza et al., 2004; Tae et al., 2008), a clear tabulation of the right and left volumes of all methods would have been preferred. To assist readers we compiled Table 1 from recent publications on the hippocampus using healthy controls. The Table shows clearly the rightward-asymmetry of the hippocampus, consistent with an earlier meta-analysis by Pedraza et al. (2004). The asymmetry of the hippocampus could have been used as an additional index to compare the three approaches (Tae et al., 2008).

Fifth, a key paper by Shattuck et al. (2008) -- not cited in Morey at al. (2009)-- may have provided important clues to why automated methods do differ compared to manually-delineated procedures. Since the same MRI data are used in both automated approaches, then differences between automated methods such as FSL-FIRST and FreeSurfer may result from sensitivity to data acquisition parameters (Bigler et al., 1997; Jack et al., 1995; Jovicich et al., 2009; Laakso et al., 1997), preprocessing stages, registration, spatial normalization (Allen et al., 2008), tissue segmentation details, and the use of specialized anatomical labeled atlases (Rodinov et al., 2009) with different spatial resolutions (Shattuck et al., 2008). The hippocampus and amygdale volumes in FSL and FreeSurfer atlases may not be identical as these are generated from different populations. It would have been interesting if Morey et al. (2009) had correlated the systematic bias with respect to the manual delineation (~ 30%) in the FreeSurfer-estimated hippocampus volume (see also Tae et al., 2008) with the volume of the hippocampus in the anatomical atlas used by FreeSurfer itself.

In conclusion, we suggest that manually-delineated hippocampal volumes need to be compared with published works after providing age, gender, handedness, and total brain volume. Without exploring sensitivity to variables such as age and gender on large samples of children and adults to account for development and aging trends, it remains premature at this stage to describe automated methods for tissue estimation as “Superior” over other methods.

Acknowledgments

National Institutes of Health-National Institute for Neurological Diseases and Stroke grant R01-NS052505-04 (K.M.H.).

References

Allen JS, Bruss J, Mehta S, Grabowski T, Brown CK, Damasio H. Effects of spatial transformation on regional brain volume estimates. Neuroimage 2008;42:535–547. [PubMed: 18599317]

Allen JS, Bruss J, Brown CK, Damasio H. Normal neuroanatomical variation due to age: the major lobes and a parcellation of the temporal region. Neurobiol Aging 2005;26:1245–60. discussion 1279–1282. [PubMed: 16046030]

Bigler ED, Blatter DD, Anderson CV, Johnson SC, Gale SD, Hopkins RO, Burnett B. Hippocampal volume in normal aging and traumatic brain injury. AJNR Am J Neuroradiol 1997;18:11–23. [PubMed: 9010515]

Gogtay N, Nugent TF 3rd, Herman DH, Ordonez A, Greenstein D, Hayashi KM, Clasen L, Toga AW, Giedd JN, Rapoport JL, Thompson PM. Dynamic mapping of normal human hippocampal development. Hippocampus 2006;16:664–672. [PubMed: 16826559]

Jack CRJ, Theodore WH, Cook M, McCarthy G. MRI-based hippocampal volumetrics: data acquisition, normal ranges, and optimal protocol. Magn Reson Imaging 1995;13:1057– 1064. [PubMed: 8750317]

Jovicich J, Czanner S, Han X, Salat D, van der Kouwe A, Quinn B, Pacheco J, Albert M, Killiany R, Blacker D, Maguire P, Rosas D, Makris N, Gollub R, Dale A, Dickerson BC, Fischl B. MRI-derived measurements of human subcortical, ventricular and intracranial brain volumes: Reliability effects of scan sessions, acquisition sequences, data analyses, scanner upgrade, scanner vendors and field strengths. Neuroimage. 2009 10.1016/j.neuroimage.2006.02.051.

Laakso MP, Juottonen K, Partanen K, Vainio P, Soininen H. MRI volumetry of the hippocampus: the effect of slice thickness on volume formation. Magn Reson Imaging 1997;15:263–265. [PubMed: 9106155]

Liu RS, Lemieux L, Bell GS, Sisodiya SM, Shorvon SD, Sander JW, Duncan JS. A longitudinal study of brain morphometrics using quantitative magnetic resonance imaging and difference image analysis. NeuroImage 2003;20:22–33. [PubMed: 14527567]

Makris N, Gasic GP, Seidman LJ, Goldstein JM, Gastfriend DR, Elman I, Albaugh MD, Hodge SM, Ziegler DA, Sheahan FS, Caviness VS Jr, Tsuang MT, Kennedy DN, Hyman SE, Rosen BR, Breiter HC. Decreased absolute amygdala volume in cocaine addicts. Neuron 2004;44:729–740. [PubMed: 15541319]

McDonald CR, Hagler DJ, Ahmadi ME, Tecoma E, Iragui V, Dale AM, Halgren E. Subcortical and cerebellar atrophy in mesial temporal lobe epilepsy revealed by automatic segmentation. Epilepsy Res 2008;79:130–138. [PubMed: 18359198]

Morey RA, Petty CM, Xu Y, Pannu Hayes J, Wagner HR 2nd, Lewis DV, Labar KS, Styner M, McCarthy G. A comparison of automated segmentation and manual tracing for quantifying hippocampal and amygdala volumes. NeuroImage 2009;45:855–866. [PubMed: 19162198]

Pedraza O, Bowers D, Gilmore B. Asymmetry of the hippocampus and amygdala in MRI volumetric measurements of normal adults. J of Int Neuropsychol Soc 2004;10:664–678. [PubMed: 15327714]

Plessen KJ, Bansal R, Zhu H, Whiteman R, Amat J, Quackenbush GA, Martin L, Durkin K, Blair C, Royal J, Hugdahl K, Peterson BS. Hippocampus and amygdala morphology in attention-deficit/hyperactivity disorder. Arch Gen Psychiatry 2006;63:795–807. [PubMed: 16818869]

Raz N, Gunning FM, Head D, Dupuis JH, McQuain J, Briggs SD, Loken WJ, Thornton AE, Acker JD. Selective aging of the human cerebral cortex observed in vivo: differential vulnerability of the prefrontal gray matter. Cereb Cortex 1997;7:268–282. [PubMed: 9143446]

Rodionov R, Chupin M, Williams E, Hammers A, Kesavadas C, Lemieux L. Evaluation of atlas-based segmentation of hippocampi in healthy humans. Magn Reson Imaging. 200910.1016/j.mri.2009.01.008

Shattuck DW, Mirza M, Adisetiyo V, Hojatkashani C, Salamon G, Narr KL, Poldrack RA, Bilder RM, Toga AW. Construction of a 3D probabilistic atlas of human cortical structures. Neuroimage 2008;39:1064–1080. [PubMed: 18037310]

Tae WS, Kim SS, Lee KU, Nam EC, Kim KW. Validation of hippocampal volumes measured using a manual method and two automated methods (FreeSurfer and IBASPM) in chronic major depressive disorder. Neuroradiology 2008;50:569–581. [PubMed: 18414838]

Walhovd KB, Fjell AM, Reinvang I, Lundervold A, Dale AM, Eilertsen DE, Quinn BT, Salat D, Makris N, Fischl B. Effects of age on volumes of cortex, white matter and subcortical structures. Neurobiol Aging 2005;26:1261–1270. [PubMed: 16005549]

Xu Y, Valentino DJ, Scher AI, Dinov I, White LR, Thompson PM, Launer LJ, Toga AW. Age effects on hippocampal structural changes in old men: the HAAS. NeuroImage 40:1003–1015. [PubMed: 18280181]

Zimmerman ME, Brickman AM, Paul RH, Grieve SM, Tate DF, Gunstad J, Cohen RA, Aloia MS, Williams LM, Clark CR, Whitford TJ, Gordon E. The relationship between frontal gray matter volume and cognition varies across the healthy adult lifespan. Am J Geriatr Psychiatry 2006;14:823–833. [PubMed: 17001022]

Table 1

A list studies on healthy controls that manually delineated the hippocampus volume and provided population age and volume (mean and standard deviation) bilaterally.

Authors, year	N(H/S) Controls	Age (years)	Right HV (mL)	Left HV (mL)	AI (%)
Allen et al., 2008	21 (F/Rh)	33.0 ± 7.3 23–47	3.50 ± 0.41	3.20 ± 0.41	9.0
Makris et al., 2002	27 (21M)	35.6 23–46	4.19 ± 0.13	4.01 ± 0.13	4.4
Plessen et al., 2006	63 (Rh)	11.5 ± 3.04	3.18 ± 0.40	3.14 ± 0.43	1.3
Shattuck et al., 2008	40 (20F)	29.2 6.3 19.3 ± 39.5	4.12 ± 0.52	3.91 ± 0.52	5.2
Tae et al., 2008	20 (F/Rh)	41.9 ± 10.3 21–57	2.87 ± 0.27	2.81±0.22	2.2
Xue et al., 2008	104 (M)	83.2±4.4 75–95	2.91±0.49	2.78±0.47	4.8

* Manually delineated as reported by these papers which are alphabetized by leading author last name. Tabulated values are Mean ± standard deviation

Hippocampus volume asymmetry index (AI) = 2*(RHV-LHV)/(RHV+LHV) × 100 %, Where, LHV = Left Hippocampus Volume in mL; RHV = Right Hippocampus Volume in mL; Rh = Right handed; S = Sex; H = Handedness; M = Males; F = Females.