Title: Hippocampal volume and integrity as predictors of cognitive decline in intact elderly

Running Head: Hippocampal Integrity and volume

Authors: Davide **Bruno**[1], Adam **Ciarleglio**[2], Michel J. **Grothe**[3], Jay **Nierenberg**[2,4], Alvin H. **Bachman**[2], Stefan J. **Teipel**[3,5], Eva **Petkova**[2,4], Babak A. **Ardekani**[2,4] & Nunzio **Pomara**[2,4]

[1] Department of Psychology, Liverpool Hope University, Liverpool, UK

[2] School of Medicine, New York University, New York City, NY, USA

[3] German Center for Neurodegenerative Diseases (DZNE) – Rostock/Greifswald, Rostock, Germany

[4] Nathan Kline Institute for Psychiatric Research, Orangeburg, NY, USA

[5] Department of Psychosomatic Medicine, University of Rostock, Rostock, Germany

Corresponding Author: Nunzio Pomara, MD

Geriatric Psychiatry Division, Nathan Kline Institute for Psychiatric Research,

Orangeburg, NY, USA

Phone: 1-845-398-5581

Email: pomara@nki.rfmh.org

Word count: 2480

Funding and Conflict of Interest: This research was funded in part by an NIMH grant (R01 MH-080405) to NP. No conflicts of interest to declare.

Abstract

Risk of Alzheimer's disease (AD) can be predicted by volumetric analyses of MRI data in the medial temporal lobe. The present study compared a volumetric measurement of the hippocampus to a novel measure of hippocampal integrity derived from the ratio of parenchyma volume over total volume.

Participants were cognitively intact and aged 60 or older at baseline, and were tested twice, roughly three years apart. Participants had been recruited for a study on late-life major depression (LLMD) and were evenly split between depressed and controls.

Linear regression models were applied to the data with a cognitive composite score as outcome, and hippocampal integrity (HI) and volume (HV), together or separately, as predictors. Subsequent cognitive performance was predicted well by models that include an interaction between HI and LLMD-status, such that lower HI scores predicted more cognitive decline in depressed subjects.

More research is needed, but tentative results from this study appear to suggest that the newly introduced measure HI is an effective tool for the purpose of predicting future changes in general cognitive ability, and especially so in individuals with LLMD.

Keywords: Alzheimer's disease; Hippocampus; MRI; Integrity; Depression.

Introduction

Volumetric analyses of MRI data have been shown to predict conversion to Alzheimer's disease (AD) from a cognitively intact baseline. Tondelli et al. [1], for example, used MRI data to predict conversion to AD at least four years prior to any symptoms being detected. Their findings showed that AD converters presented greater atrophy in the right medial temporal lobe, an area prominently including the hippocampus (HC), compared to non-converters. More recently, others [2] have demonstrated that lower volumes of the HC in the right hemisphere predicted time of onset of clinical symptoms of neurodegeneration from a cognitively intact baseline over an average of 10 years.

A few issues should be noted, however, with regards to hippocampal volumetric measurements (HV). Manual HV measurements are tedious and time-consuming activities, suffer from intra- and inter-observer variability, require extensive operator training, and have low comparability across laboratories due to differing tracing protocols [3]. Although recently, a harmonized HR protocol has been developed and evaluated that may overcome some of the inconsistencies in previous manual volumetric approaches, this protocol still requires intensive training of raters and is time consuming to apply [4]. Automated algorithms for HV measurements, however, tend to be generally less robust as compared to manual measurements [5], can be computationally expensive, sometimes requiring several hours of computer time, are not widely available, and often require extensive preprocessing of the MRI scans (e.g., inhomogeneity correction, tissue segmentation, distortion correction) and associated technical operator expertise. In addition, both manual and automated HV measurements ought to be corrected for intra-cranial volume (ICV) with which they are significantly correlated. However, accurate measurement of the ICV is itself a non-trivial problem.

As an alternative, Ardekani et al. [6] have proposed a measure of hippocampal volumetric integrity (HI) based on the notion that cerebrospinal fluid (CSF) replaces brain parenchyma in the process of neurodegeneration. Therefore, in given standardized regions-of-interest (ROIs), the ratio of parenchyma volume over total volume (parenchyma plus CSF) would decrease as a result of neuronal loss. Compared to HV measures, as described above, the proposed HI measurement would provide an indirect estimation of HC atrophy rapidly (i.e., in less than a minute), while not requiring any preprocessing. HI can also be applied to scans immediately after acquisition, does not require adjustment for ICV, and relies on very little, if any, user image processing expertise.

In this study, we set out to compare the newly developed HI to a standard method for automatic HV measurement in their respective potential for prediction of cognitive performance. If it can be shown that HI is comparable to at least one type of HV in predicting cognitive change over time, then this finding could have significant impact in the field, providing a useful and practical tool potentially for both research and clinical practice.

Our study population was cognitively intact at baseline and was followed over a period of three to four years on average. Importantly, this sample was recruited for a study on late-life major depression (LLMD) and half of all participants had a depression diagnosis at baseline. Major depression has been shown repeatedly to associate with AD risk [7], although the exact nature of this relationship and resulting risk have not been completely elucidated. Some evidence suggests that amyloid beta disturbances may be present in both conditions [e.g., 8].

We measured cognitive ability by constructing a composite score that included: the Mini-mental State Exam (MMSE) score, which provides a general cognition index; the Digit Symbol Substitution Test (DSST), a measure of executive function/attention; and the delayed

recall test of the AVLT, a memory task. The composite score was obtained by standardizing the individual test scores over the population and then adding these values together.

Methods

Subjects. Participants were recruited at the Nathan Kline Institute (NKI) and New York University (NYU) Langone Medical Center for a study on late-life depression; the total number of subjects was originally 131, from which 12 individuals were excluded who either had an MMSE score ≤ 27, or showed stroke, extensive white matter (WM) disease or severe ventriculomegaly on the MRI. Of the remaining 119 subjects, 94 returned for at least two follow up sessions, and for 90 of these participants both HV and HI could be determined successfully, thus forming our study sample. The group was evenly split between participants with a diagnosis of LLMD and those without. All participants provided informed consent prior to taking part in the study. The NKI and NYU institutional review boards authorized this study on ethical grounds. Table 1 reports the demographic characteristics of the sample, and differences across groups were assessed with t-tests.

Table 1 here

HV. MRI data processing followed a standard SPM-based procedure for atlas-based volumetry of the hippocampus based on high-dimensional image registration to MNI standard space and a manually traced hippocampal ROI following standardized delineation criteria [9]. ICV was calculated within this framework by summing up the total volumes of gray matter, white matter, and cerebrospinal fluid partitions from the automated tissue segmentation output. Details of this procedure have been described previously [10].

HI. The hippocampal volumetric integrity measure was computed using the following procedure: 1) The mid-sagittal plane (MSP) was detected automatically on the MRI volume using the method described in [11]; 2) The anterior and posterior commissures (AC-PC) were

automatically located on the MSP using the method described in [12]; 3) Using the information from steps (1) and (2), the MRI volume was reoriented into a standard orientation where the x-axis points to the posterior direction and is parallel to the AC-PC line, the y-axis points to the inferior direction and the z-axis points to the left; the xy-plane is the MSP and the origin of the coordinates system is the halfway point between the AC and PC on the MSP. We call this the Posterior-Inferior-Left (PIL) orientation; 4) Approximately 100 landmarks were detected automatically around the hippocampus using a previously trained supervised landmark detection method; after that, an affine transformation was estimated to map these landmarks as closely and possible to a set of standard locations that had been previously determined based on a training set of scans, as were the patterns used for landmark detection; 5) The rigid-body transformation of step (3) and the affine transformation of step (4) were multiplied and the result inverted to obtain a single linear transformation; this transformation maps information from a standard space to the space of the original MRI scan; 6) The linear transformation in step (5) was applied to a set of 65 manually delineated hippocampal atlases that had previously undergone the combinations of the transformations in steps (3) and (4) to obtain a probabilistic HC ROI on the original MRI scan; Finally, 7) An automatic threshold selection procedure was applied to segment the voxels in this ROI as brain parenchyma and CSF. HI was defined as the ratio of the parenchymal voxels to the total number of voxels in the ROI. Steps (4)-(7) were repeated for the left and right hippocampi independently to obtain HI for both sides. More details about this methodology can be found in [6]. The software for HI estimation (kaiba) is freely available online at: www.nitrc.org/projects/art.

Procedure. Participants were tested at the Nathan Kline Institute and at the New York University Medical School, over three visits on successive weeks. On the first visit, participants provided informed consent, were administered a general medical intake

questionnaire, and had their vital signs examined; during this session, the MMSE test was also administered. Participants received an MRI scan of the head on the second visit. Finally, on the third visit, participants underwent a comprehensive neuropsychological assessment, including administration of the DSST.

*Design and Analysis. R*egression modeling was used to investigate the association between the composite score at 3-year follow-up and measures of hippocampal volume (HV), and hippocampal integrity (HI) in both the right and left hemispheres. The outcome variable for each model considered was the composite score at 3-year follow-up. We were primarily interested in assessing the predictive ability of HV and HI separately or combined. We also investigated whether the LLMD status at baseline moderated the effect of either of the HV or HI measures. We used robust linear regression, implemented in R [13], since some of the follow-up composite scores were much smaller (negative values of large magnitude) than the majority of the scores. Observations corresponding to these large-magnitude scores showed evidence of strong influence on the estimates derived from ordinary least-squares regression. Specifically, we used M-estimation with Huber weighting [14] (tuning constant $k = 4.685\hat{\sigma}$ where $\hat{\sigma}$ is the estimate of the median absolute residual divided by 0.6745; this provides coefficient estimates that are about 95% as efficient as those produced by ordinary least squares, when the errors are normally distributed) to obtain regression estimates and test statistics for the models that we fit. Huber weighting gives more weight to observations with smaller residuals while giving smaller weight to observations with larger residuals, thus reducing the influence of those observations with larger residuals on the regression estimates. Besides having HV, HI, or both measures as predictors in a given model, we also adjusted for the following covariates: baseline composite score, sex, LLMD status, e4 status, and TIV. The HV and HI measures were centered and scaled before entering the model so that coefficient estimates are comparable. For each robust linear regression model, we computed

the pseudo weighted least-squares coefficient of determination, pseudo R^2_{WLS} [15]. *P*-values for the regression coefficients are based on standard normal approximations for the distributions of the corresponding test statistics

Results

First, it was investigated whether LLMD status modified the association between the composite score at 3-year follow-up and either of the HV or HI measures on either the right or left side. We found that LLMD status did modify the association between HI and the composite score on both the right and left side, but not the association between HV and the composite score on either side. Table 2 shows the standardized adjusted effect estimates and pseudo R^2_{WLS} values for each of the three relevant models fit using the right or left HV and HI measures as predictors. For comparison, we have also included the R^2 values from the corresponding ordinary least-squares fits.

Models with right side measures: The model with both HV and HI measures (and interaction HI*LLMD) included as predictors (Model 1R) suggests that there is a positive association between each of these measures and Composite score at 3-year follow-up, however only HI shows a significant association among depressed subjects with an adjusted effect estimate of 0.88 (p = 0.003) (i.e., among depressed subjects and adjusting for the other covariates, including HV, a one standard deviation increase in HI on the right side corresponds to a 0.88 point increase in Composite score at 3-year follow-up on average). The main effects model with only HV (Model 2R) suggests that, without adjusting for HI, there is an estimated positive association between HV and Composite score at 3-year follow-up with an adjusted effect estimate but the effect is not significantly significant [0.35 (0.07)]. The model with only HI and HI*LLMD (Model 3R) suggests that, without adjusting for HV, there is a significant positive association between HI and Composite score at 3-year follow-up among

depressed subjects with an adjusted effect estimate of 0.99 (p = 0.0003). The fact that the adjusted coefficient for HI among depressed subjects is larger than that for HV in both Models 1R and 3R and the pseudo R^2_{WLS} for model 3R is larger than that for Model 2R suggests that HI may be a better predictor of Composite score at 3-year follow-up than HV among depressed subjects.

Models with left side measures: Models with left side measures show similar relationships between Composite score at 3-year follow-up and the HV and HI measures although the predictive ability of these measures is not as strong as those measures from the right side as evidenced by the lower pseudo R^2_{WLS} values shown in the Table 2.

Non-composite analyses: When examining the individual components of the composite score, i.e., MMSE, DSST and AVLT delayed recall, we did not detect any modifying effects of LLMD status on the association between the outcome scores and either HI or HV. Therefore, for these analyses, we only employed three models (HV + HI, HV or HI) controlling for baseline scores, sex, LLMD status, e4 status and TIV. For the right hippocampus, we found both HI (p<.001) and HV (p=.034) to predict follow up MMSE performance, although neither measure predicted follow up DSST or delayed recall scores. For the left hippocampus, only HI was significantly associated (p=.025) with MMSE, and again neither measure predicted subsequent performance in DSST or delayed recall.

Tables 2 here

Discussion

The current study set out to compare the relative predictive values of two automated measures of hippocampal atrophy on MRI. HV provides a direct measure of the total hippocampal gray matter volume, whereas HI provides a novel measure of volumetric integrity, defined as the ratio of parenchymal voxels to the total number of voxels in a

linearly registered hippocampal probabilistic ROI. Our findings showed that, when predicting cognitive ability using a composite cognitive score, both the right and left HI measures stood out as significant predictors of decline in individuals with LLMD, although the right hippocampus provided relative better prediction than the left hippocampus. An advantage for the right over the left side is consistent with existing literature [1-2, 16], and may reflect the fact that, for verbal memory tasks, where the left hippocampus is likely to be affected before the right hippocampus, the right hippocampus may provide support in the form of a secondary network. Therefore, once this secondary network is significantly atrophied, it may be an indication that most available reserve has been depleted and that global decline is forthcoming.

A key finding in this study is that LLMD status interacted with HI to yield significant predictive value. LLMD is a well-established risk factor for AD [7] and may be a prodromal stage of the disease [17]. In relation to this, depression has been associated with both reductions in hippocampal neurogenesis [18] and HPA axis dysfunction, including increased cortisol levels [19].

Although more research is needed, initial results from this study appear to suggest that HI is comparable, if not superior, to HV for the purpose of predicting cognitive decline over a short period of time. An obvious limitation to note is that we only followed our participants for a relatively short period of time, and all were cognitively intact at baseline; therefore, no substantial change in generalized cognitive ability, or conversion to dementia, was detected in this cohort. To mitigate this, however, it should be noted that detecting subtle drops in performance in high-functioning individuals is more difficult than in relatively more impaired participants, thus testifying to the sensitivity of HI [20].

Acknowledgments

Part of this research was presented as a poster at the 2015 Alzheimer's Association International Conference (Washington, D.C., USA) and published as an abstract in *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*.

References

1. Tondelli M, Wilcock GK, Nichelli P, De Jager CA, Jenkinson M, Zamboni, G. Structural MRI changes detectable up to ten years before clinical Alzheimer's disease. Neurobiol Aging. 2012; 33:825-e25.

2. Soldan A, Pettigrew C, Li S, Wang MC, Moghekar A, Selnes OA, Albert M, O'Brien R, BIOCARD Research Team. Relationship of cognitive reserve and cerebrospinal fluid biomarkers to the emergence of clinical symptoms in preclinical Alzheimer's disease. Neurobiol Aging. 2013; 34:2827–34.

3. Boccardi M, Ganzola R, Bocchetta M, Pievani M, Redolfi A, Bartzokis G, Camicioli R, Csernansky JG, de Leon MJ, deToledo-Morrell L, Killiany RJ, Lehericy S, Pantel J, Pruessner JC, Soininen H, Watson C, Duchesne S, Jack CR Jr, Frisoni GB. Survey of protocols for the manual segmentation of the hippocampus: preparatory steps towards a joint EADC-ADNI harmonized protocol. Journal of Alzheimer's disease. 2011; 26:61-75.

4. Boccardi M, Bocchetta M, Morency FC, Collins DL, Nishikawa M, Ganzola R, et al. Training labels for hippocampal segmentation based on the EADC-ADNI harmonized hippocampal protocol. Alzheimers Dement. 2015.

5. Mulder ER, de Jong RA, Knol DL, van Schijndel RA, Cover KS, Visser PJ, Barkhof F, Vrenken, H. Hippocampal volume change measurement: quantitative assessment of the reproducibility of expert manual outlining and the automated methods FreeSurfer and FIRST. Neuroimage. 2014; 92:169–81.

6. Ardekani, BA, Convit A, Bachman AH. Analysis of the MIRIAD Data Shows Sex Differences in Hippocampal Atrophy Progression. J Alzheimer's Disease. 2016 (in press).

7. Byers AL, Yaffe K. Depression and risk of developing dementia. Nat Rev Neurol. 2011; 7:323-31.

8. Pomara N, Bruno D, Sarreal AS, Hernando RT, Nierenberg J, Petkova E, Sidtis JJ,

282 Wisniewski TM, Mehta PD, Pratico D, Zetterberg H, Blennow K. Lower CSF amyloid beta
283 peptides and higher F2-isoprostanes in cognitively intact elderly individuals with major
284 depressive disorder. Am J Psychiatry. 2012;169:523-30.
285 9. Frisoni, G.B., Jack, C.R Jr., Bocchetta, M., Bauer, C., Frederiksen, K.S., Liu, Y. et
286 al, The EADC-ADNI Harmonized Protocol for manual hippocampal segmentation on
287 magnetic resonance: Evidence of validity. *Alzheimers Dement*. 2015;11:111–125.
288 10. Bruno D, Grothe MJ, Nierenberg J, Zetterberg H, Blennow K, Teipel SJ, Pomara N.
289 A study on the specificity of the association between hippocampal volume and delayed
290 primacy performance in cognitively intact elderly individuals. Neuropsychologia. 2015;69:1-
291 11. Ardekani BA, Kershaw J, Braun M, Kanno I. Automatic detection of the mid-sagittal
292 plane in 3-D brain images. IEEE Trans Med Imaging 1997: 16(6), 947-52.
293 12. Ardekani BA, Bachman AH (2009) Model-based automatic detection of the anterior
294 and posterior commissures on MRI scans. Neuroimage 2009: 46(3), 677-82.
295 13. R Development Core Team. R: A language and environment for statistical computing.
296 Vienna, Austria: R Foundation for Statistical Computing. 2015.
297 14. Huber, P. J. Robust estimation of a location parameter. The Annals of Mathematical
298 Statistics. 1964; 35: 73–101.
299 15. Willett, J. B., Singer, J. D. Another cautionary note about R^2: Its use in weighted
300 least-squares regression analysis. The American Statistician.1988; 42(3):236–238.
301 16. Brüggen K, Dyrba M, Barkhof F, Hausner L, Filippi M, Nestor PJ, Haunestein K,
302 Kloppel S, Grothe MJ, Kasper E, Teipel SJ. Basal Forebrain and Hippocampus as Predictors
303 of Conversion to Alzheimer's Disease in Patients with Mild Cognitive Impairment: A
304 Multicenter DTI and Volumetry Study. Journal of Alzheimer's Disease. 2015;1-8.
305 17. Dal Forno G, Palermo MT, Donohue JE, Karagiozis H, Zonderman AB, Kawas CH.
306 Depressive symptoms, sex, and risk for Alzheimer's disease. Annals Neurol. 2005; 57:381-7.

18. Boldrini M, Hen R, Underwood MD, Rosoklija GB, Dwork AJ, Mann JJ, Arango V. Hippocampal angiogenesis and progenitor cell proliferation are increased with antidepressant use in major depression. Biol Psychiatry. 2012;72:562-71.

19. Tafet GE, Nemeroff CB. The Links Between Stress and Depression: Psychoneuroendocrinological, Genetic, and Environmental Interactions. J Neuropsychiatry Clin Neurosci. 2015.

20. Bruno D, Reiss PT, Petkova E, Sidtis JJ, Pomara N. Decreased recall of primacy words predicts cognitive decline. Archives of clinical neuropsychology. 2013;28:95-103.