Automated measurement of hippocampal subfields in PTSD: Evidence for smaller dentate gyrus volume

Jasmeet P. Hayes a, b, c, *, Scott Hayes b, c, d, Danielle R. Miller a, b, Ginette Lafleche d, Mark W. Logue a, c, f, Mieke Verfaellie b, d

a National Center for PTSD, VA Boston Healthcare System, Boston, MA, USA
b Department of Psychiatry, Boston University School of Medicine, Boston, MA, USA
c Neuroimaging Research for Veterans Center, VA Boston Healthcare System, Boston, MA, USA
d Memory Disorders Research Center, VA Boston Healthcare System and Boston University School of Medicine, Boston, MA, USA
e Biomedical Genetics, Boston University School of Medicine, Boston, MA, USA
f Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA

ARTICLE INFO

Article history:
Received 13 March 2017
Received in revised form 25 August 2017
Accepted 7 September 2017

Keywords:
MRI
FreeSurfer
Veterans
Stress disorders
Memory
Hippocampus

ABSTRACT

Smaller hippocampal volume has been consistently observed as a biomarker of posttraumatic stress disorder (PTSD). However, less is known about individual volumes of the subfields composing the hippocampus such as the dentate gyrus and cornu ammonis (CA) fields 1–4 in PTSD. The aim of the present study was to examine the hypothesis that volume of the dentate gyrus, a region putatively involved in distinctive encoding of similar events, is smaller in individuals with PTSD versus trauma-exposed controls. Ninety-seven recent war veterans underwent structural imaging on a 3T scanner and were assessed for PTSD using the Clinician-Administered PTSD Scale. The hippocampal subfield automated segmentation program available through FreeSurfer was used to segment the CA4/dentate gyrus, CA1, CA2/3, presubiculum, and subiculum of the hippocampus. Results showed that CA4/dentate gyrus subfield volume was significantly smaller in veterans with PTSD and scaled inversely with PTSD symptom severity. These results support the view that dentate gyrus abnormalities are associated with symptoms of PTSD, although additional evidence is necessary to determine whether these abnormalities underlie fear generalization and other memory alterations in PTSD.

Published by Elsevier Ltd.

1. Introduction

Posttraumatic stress disorder (PTSD) is a debilitating psychiatric condition that develops following exposure to highly distressing life events. One of the most consistently observed biological markers of PTSD is smaller hippocampal volume, with several meta-analyses now supporting a small but significant association (Kitayama et al., 2005; Smith, 2005; Karl et al., 2006; O’Doherty et al., 2015). The hippocampal formation is composed of several subfields that are thought to mediate different memory functions including the dentate gyrus (DG), cornu ammonis (CA) 1–4, presubiculum, and subiculum. The DG subfield is putatively involved in “pattern separation,” which is the process by which incoming neural signals are made more distinct from each other at the time of encoding (O’Reilly and McClelland, 1994). Recently, it has been proposed that pattern separation deficits may underlie fear generalization (Kheirbek et al., 2012), a process that occurs in anxiety and stress based disorders including PTSD (Morey et al., 2015).

Consistent with the notion that the DG may be impaired in PTSD, Wang et al. (2010) observed smaller DG/CA3 volume in 17 subjects with PTSD relative to 17 trauma-exposed control subjects. Further, lower CA1 volume was found as a function of age but not PTSD. This study used manual tracing of hippocampal subfield regions, the gold standard in the field. However, manual tracing is an extremely time-consuming and labor intensive process. As trends in science have moved toward large datasets to study subtle relationships between brain anatomy, psychiatric symptoms and other behavioral outcome measures, hand-tracing of hippocampal
subfields has become prohibitive in terms of cost and time. Further, the wide range in manual tracing methods with lack of consistent protocols (Geuze et al., 2005) is a key limitation of manual volumetric assessment. By contrast, automated segmentation can foster standardization of methods across different laboratories and represents a time-efficient strategy to segment the hippocampus without sacrificing inter-rater reliability.

Recent developments in automated measures put forth by the creators of FreeSurfer provide a method to delineate hippocampal subfield volume on standard T1 high resolution images (Van Leemput et al., 2009). This method is based on a Bayesian modeling approach that predicts the location of neuroanatomical labels based on probabilistic atlases and learned locations of manual hippocampal segmentations from training subjects. The automated segmentations have been validated against manual morphometric measurements of ultra-high resolution scans. The automated hippocampal subfield extraction tool outputs left and right volumes of the following structures: CA4/DG, CA1, CA2/3, presubiculum, subiculum, fimbria, hippocampal fissure, and the tail of the hippocampus. Several studies have now used these methods to examine differences among psychiatric groups including schizophrenia and bipolar disorder (Mathew et al., 2014; Haukvik et al., 2015). In the current study, we examined hippocampal subfield volumes in a large group of trauma-exposed individuals. We hypothesized that we would observe lower DG volume in the PTSD group, similar to the findings reported by Wang et al. (2010). As age was shown to influence certain volumes in that study, we also examined the effects of age and the PTSD by age interaction on subfield volumes.

2. Materials and methods

2.1. Participants

Participants were 97 (mean age = 30) veterans who had been deployed in support of Operation Enduring Freedom, Operation Iraqi Freedom, or Operation New Dawn (OEF/OIF/OND). Participants were excluded from the study if they reported a history of pre-deployment traumatic brain injury (TBI) with loss of consciousness (LOC) of any duration or TBI without LOC with symptoms persisting more than three months after the injury. Additional exclusion criteria were moderate or severe TBI at any time, structural brain abnormalities as determined by a board-certified neuroradiologist, and high levels of current alcohol use (>25 drinks per week). Participants were recruited through the VA Boston Polytrauma Network and through flyers and outreach events in the community. Study procedures were approved by the VA Boston Institutional Review Board and all participants provided written informed consent consistent with the Declaration of Helsinki. A summary of the demographic characteristics is shown in Table 1.

Table 1
Demographic and clinical characteristics.

<table>
<thead>
<tr>
<th></th>
<th>PTSD (n = 58)</th>
<th>Controls (n = 39)</th>
<th>Group Comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age in years, M (SD)</td>
<td>29.7 (7.1)</td>
<td>29.3 (6.4)</td>
<td>t(95) = −0.271, p = 0.787</td>
</tr>
<tr>
<td>Males, n (%)</td>
<td>56 (96.6)</td>
<td>35 (89.7)</td>
<td>χ²(1) = 1.863, p = 0.172</td>
</tr>
<tr>
<td>Education in years, M (SD)</td>
<td>12.9 (1.8)</td>
<td>14.1 (2)</td>
<td>t(95) = 3.15, p = 0.002</td>
</tr>
<tr>
<td>WTAR, M (SD)</td>
<td>0.35 (0.7)</td>
<td>0.42 (0.8)</td>
<td>t(94) = 0.415, p = 0.679</td>
</tr>
<tr>
<td>Current alcoholic drinks per week, M (SD)</td>
<td>71 (18.8)</td>
<td>25.1 (14.8)</td>
<td>t(95) = −12.79, p &lt; 0.001</td>
</tr>
<tr>
<td>CAPS total, M (SD)</td>
<td>12.9 (1.8)</td>
<td>14.1 (2)</td>
<td>t(95) = 3.15, p = 0.002</td>
</tr>
<tr>
<td>PCL total, M (SD)</td>
<td>4.5 (6.8)</td>
<td>3.9 (5.5)</td>
<td>t(95) = −0.44, p = 0.658</td>
</tr>
</tbody>
</table>

Note: WTAR is listed as average z-score. WTAR = Wechsler Test of Adult Reading; CAPS = Clinician-Administered PTSD Scale.

2.2. Clinical assessment procedure

PTSD diagnosis was assessed using the Clinician-Administered PTSD Scale (CAPS) for DSM-IV (Blake et al., 1995) by a trained clinical psychologist. On this interview, each DSM-IV PTSD criterion was assessed with two CAPS sub-items, one which reflects the frequency of the symptom on a 0–4 scale and one which reflects the intensity of the symptom on a 0–4 scale; the two sub-items can be combined to reflect either presence/absence of that symptom for diagnosis or symptom severity. A score for each individual was derived using the total score for PTSD symptoms occurring within the past month. Subjects were placed in the PTSD group if they met criteria for PTSD diagnosis or in the trauma-exposed control group if they experienced a criterion A trauma event but did not meet PTSD diagnosis. Continuous CAPS scores were used as a measure of PTSD symptom severity. CAPS scores were unavailable for three participants and were estimated based on their score on the PTSD Checklist—Military version (PCL-M; Weathers et al., 1991). The PCL is a self-report questionnaire that has good convergent validity with the CAPS (Wilkins et al., 2011). A linear regression model was generated using PCL scores of the remaining dataset as predictors of their CAPS scores. Using this model, the three individuals’ PCL scores were entered into the equation to generate their predicted CAPS scores.

As a measure of trauma chronicity, we calculated the number of days that had elapsed between the date of first military deployment and PTSD interview. This metric was selected due to the observation that war-zone traumatic experiences were not always limited to a single event but were often the result of repeated life-threatening events throughout the course of deployment. Thus, an advantage of using military deployment date is that it is objective, verifiable, and tied to the start of stress exposure. Most participants (92; 95%) had a deployment-related trauma event. For analyses using the trauma chronicity variable, the five subjects without a deployment-related trauma were removed because the trauma chronicity variable in those instances could not be anchored to an objective date. Of the 92 subjects with deployment related trauma, six subjects reported additional early childhood trauma. Actual date of childhood trauma in those individuals was not available. Thus, we substituted an arbitrary trauma chronicity date of 20 years for these six subjects.

2.3. Image acquisition & processing

Structural imaging data were acquired on a 3-T Siemens Trio whole-body MRI scanner located at the VA Boston Healthcare System, Jamaica Plain campus. Two 3-dimensional (3D)
magnetization-prepared rapid gradient-echo (MP-RAGE) scans were acquired in the sagittal plane for each individual and averaged to create a single high contrast-to-noise image. The first 28 participants recruited into the study were scanned using the following parameters: FOV = 240, Matrix = 240 × 256, 160 slices, 1 × 1 × 1.2 mm voxels, TR = 2300 ms, TE = 2.98 ms, flip angle = 9° and thereafter, a slightly modified TI sequence was used: FOV = 256, Matrix = 256 × 256, 176 slices, 1 × 1 × 1 mm voxels, TR = 2530 ms, TE = 3.32 ms, flip angle = 7°. The sequences were modified in order to align them with the pulse sequence of a separate study protocol for data sharing. There were no differences between participants in sequence 1 and sequence 2 in any subfield volume, whole hippocampal volume, or intracranial volume (ICV). Sequence was added as a covariate to all analyses.

The FreeSurfer image analysis suite (version 5.1, freely available for download online http://surfer.nmr.mgh.harvard.edu) was used to process the data including motion correction and averaging of multiple volumetric T1-weighted images, removal of non-brain tissue using a hybrid watershed/surface deformation procedure, automated Talairach transformation, and segmentation of the subcortical white matter and deep grey matter volumetric structures. An additional flag was appended to calculate each subfield volume. The automated hippocampal subfield extraction tool outputs left and right volumes of the following structures: CA4/DG, CA1, CA2/3, presubiculum, subiculum, fimbria, hippocampal fissure, and the tail of the hippocampus. As measurement of smaller subfields may be unreliable (Van Leemput et al., 2009), we excluded the fimbria, hippocampal fissure, and the tail of the hippocampus from analyses. However, the fimbria and tail of the hippocampus were included to calculate total hippocampal volume, which was the sum of all volumes except the hippocampal fissure. Left and right volumes for all structures were averaged to get a single score for each participant. Segmentations were visually inspected for accuracy. No subfield segmentations required manual adjustment. Subfield segmentation for a sample subject is displayed in Fig. 1.

2.4. Statistical approach

Statistical analyses were performed using SPSS, version 22 (IBM Corp., Armonk, NY). Demographic analyses used t-tests and chi-square where appropriate to compare the PTSD and trauma-exposed control groups (Table 1). Hierarchical linear regression models were run to examine differences in global hippocampal volume between the PTSD and control group. In SPSS, hierarchical regression can be performed using the linear regression command and predictors can be modeled in a hierarchical fashion by entering covariates in separate blocks with the forced entry method. In our analyses, covariates age, imaging sequence, sex, and education were entered into the first block, PTSD diagnosis in the second block, and the age × PTSD interaction in the third block. Analyses were repeated to examine the effect of trauma chronicity by adding this regressor to the first block (along with age, imaging sequence, sex, and education). Post-hoc analyses examined the effect of PTSD and the PTSD by age interaction on left and right hippocampal volume separately. Global hippocampal volume was adjusted for ICV using the covariance formula: adjusted hippocampal volume = raw hippocampal volume - b*(ICV - mean ICV), where b is the slope of a regression of a region-of-interest volume on ICV, as recommended in (Buckner et al., 2004). This approach yields a more Gaussian distribution than a ratio approach (Jack et al., 1989).

To examine differences in volume among hippocampal subfields, hierarchical linear regression models were run separately for CA4/DG, CA1, CA2/3, presubiculum, and subiculum. As there were no laterality effects in global hippocampal volume and we did not have an apriori hypothesis regarding laterality effects in subfield volumes, left and right subfield volumes were averaged and examined bilaterally. The same covariates indicated above were entered into the first block, PTSD diagnosis was entered in the second step, and the age x PTSD interaction was entered in the final step. Each subfield volume was adjusted by mean global hippocampal volume using the covariance formula: adjusted hippocampal subfield volume = raw subfield volume - b* (global hippocampal volume - mean global hippocampal volume), where b is the slope of a regression of the subfield volume on that subject’s global hippocampal volume. Global hippocampal volume was selected as a volume correction factor over ICV because all of the subfields scaled with hippocampal volume (thus, those with larger hippocampal volumes had larger subfield volumes) but did not scale with ICV. Nominal (uncorrected) as well as multiple-testing corrected significance (adjusting for analysis of five subfields) for main effects and interactions were determined using Monte-Carlo null simulation (Churchill and Doerge, 1994) with 10,000 replicates using a script written in the R statistical programming framework. This Monte-Carlo procedure permutes case/control status across subjects. Then, the p-values for the test of association with each of the five regions are computed to estimate the distribution of the minimum p-value across the five tested subfields. The percentile of the observed p-value in this minimum p-value distribution was taken as our (estimated) corrected significance level. This analysis imposes strict multiple-testing control while taking into account the correlation between subfields. To examine hippocampal subfield volume as a function of PTSD symptom severity, analyses were repeated replacing the dichotomous PTSD diagnosis variable with CAPS symptom severity scores, but only for subfields that showed a significant relationship between PTSD diagnosis and
Hierarchical linear regression showed no significant overall model for the relationship between PTSD, age and global hippocampal volume ($p > 0.4$). After including trauma chronicity, the overall model was again not significant ($p > 0.3$), but there was a significant change in the third step of the model corresponding to a PTSD by age interaction on ICV-adjusted global hippocampal volume ($p = 0.047$). We next examined global hippocampal volume laterality effects by examining right and left hemispheres separately.

In CA2/3, there was no effect of PTSD or the age x PTSD interaction ($p = 0.037$) that did not survive multiple comparisons correction. The regression model was re-run with trauma chronicity as a predictor. Trauma chronicity was not significantly associated with CA4/DG volume ($p = 0.25$). In this model, there was a significant main effect of PTSD diagnosis ($p = 0.004$, corrected $p = 0.018$) but no main effect of age or an age x PTSD interaction. We next examined the influence of handedness as a covariate in the analysis. Handedness was not a significant predictor in the model ($p = 0.56$) and the main effect of PTSD on CA4/DG remained significant ($p = 0.008$).

In CA2/3, there was no effect of PTSD or the age x PTSD interaction. After including trauma chronicity, a nominal effect of PTSD emerged ($p = 0.034$), that did not survive multiple comparisons correction ($p = 0.12$). There were no main effects of PTSD or age on CA1 or the presubiculum with or without trauma chronicity included as a regressor. The subiculum showed a nominal age x PTSD interaction effect ($p = 0.018$) but this effect did not survive multiple comparisons correction ($p = 0.076$). No other subfields showed a nominal or significant age x PTSD interaction.

Given the high comorbidity of PTSD and mTBI in our sample, and the fact that those with mTBI had significantly higher PTSD symptom severity scores than those without mTBI, we examined whether mTBI explained significant variance in subfield volume. Regression analyses revealed that mTBI was not significantly associated with any of the five hippocampal subfield volumes. There was no significant $F$ change in the second block for CA4/DG: $[\Delta F(1,91) = 2.82, p = 0.096, \Delta R^2 = 0.03]$; CA1: $[\Delta F(1,91) = 0.75, p = 0.39, \Delta R^2 = 0.008]$; presubiculum: $[\Delta F(1,91) = 2.35, p = 0.13, \Delta R^2 = 0.02]$; or subiculum: $[\Delta F(1,91) = 1.53, p = 0.22, \Delta R^2 = 0.02]$. There was a nominal association between mTBI and CA2/3: $[\Delta F(1,91) = 4.71, p = 0.03, \Delta R^2 = 0.05]$ but this association did not survive multiple comparisons correction.

### Table 2
Summary of hierarchical regression analysis for association with CA4/DG volume.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$B$</td>
<td>$SE(B)$</td>
<td>$\beta$</td>
</tr>
<tr>
<td>Age</td>
<td>0.107</td>
<td>0.377</td>
<td>0.031</td>
</tr>
<tr>
<td>Sex</td>
<td>12.33</td>
<td>10.75</td>
<td>0.125</td>
</tr>
<tr>
<td>Sequence</td>
<td>8.831</td>
<td>5.432</td>
<td>0.168</td>
</tr>
<tr>
<td>Education</td>
<td>−1.13</td>
<td>1.354</td>
<td>−0.092</td>
</tr>
<tr>
<td>PTSD</td>
<td>−13.5</td>
<td>5.04</td>
<td>−0.280</td>
</tr>
</tbody>
</table>

Note: HC = hippocampal; PTSD = posttraumatic stress disorder. \* corrected value = 0.036, \* = statistically significant.

### 4. Discussion

In a large cohort of trauma-exposed Iraq and Afghanistan war veterans, we examined hippocampal subfield volume using automated segmentation methods. Results revealed that PTSD diagnosis was associated with smaller volume in CA4/DG, and CA4/DG volume scaled inversely with symptom severity. In addition, CA2/3 volume was nominally smaller as a function of PTSD after taking into account trauma chronicity, but this effect did not survive multiple comparisons correction. These results are consistent with a previous study in PTSD using a time-intensive hand tracing method (Wang et al., 2010), which found that CA3/DG volumes were lower in PTSD. It should be noted, however, that the automated procedures used here grouped CA volumes differently than Wang and colleagues (CA4/DG versus CA3/DG, respectively) and for both, the rationale for their respective grouping method was difficult in distinguishing among these regions. Some investigators have suggested that CA4 should be considered part of CA3 as its boundaries with CA3 are not well defined (Amaral et al., 2007). In spite of this difference between studies, the common region that was smaller in both studies was the DG, consistent with theories of this region’s role in PTSD. A better understanding of the role of CA3 in PTSD may require higher resolution imaging that will allow more reliable spatial separation of subfields. Nonetheless, these results, although preliminary and requiring replication, show promise for automated methods in parsing hippocampal subfields in PTSD.

The DG is distinct from other subfields in that it has a specific cell type called mossy cells, which can function as excitatory or inhibitory inputs to granule cells. Recent optogenetics work has
revealed that the primary effect of mossy cells on granule cells is inhibitory (Hsu et al., 2016). Inhibition of granule inputs may underlie the phenomenon known as pattern separation, in which two overlapping inputs are made more distinct during memory encoding (see Scharfman, 2016). Recently, it has been proposed that pattern separation deficits may underlie generalization of fear (Kheirbek et al., 2012), which is thought to be one of the central mechanisms responsible for heightened fear responses and symptoms in anxiety and stress-based disorders. Although the link between the underlying neurobiology of DG volume and PTSD symptoms is currently speculative, the present results motivate further study of this possibility. Individuals with PTSD often report that a broad range of stimuli serve as reminders of trauma, suggesting overgeneralization of stimuli. Experimentally, individuals with PTSD have difficulty distinguishing new and old trauma items, with higher rates of false alarms for novel trauma-related items (Hayes et al., 2011). Although additional research is necessary, we believe that these findings have potential clinical relevance. Given the well-established link between DG and pattern separation, our finding of reduced DG volume in PTSD points to a potential mechanism by which PTSD symptoms are formed and maintained. Namely, PTSD may be associated with increased interference leading to symptoms such as reminders re-instantiating the trauma memory. As such, DG volume may serve as a useful biomarker for individuals at risk for fear generalization. If confirmed in future work, this hypothesis might also spur the development of treatment strategies that specifically target DG function and structure. For example, there is now evidence that exercise is associated with neurogenesis in the DG (Sahay et al., 2011). Whether this could improve PTSD symptoms remains to be determined.

Use of an automated approach to segment hippocampal subfields is a primary strength of this study, but also a primary limitation. There remains much controversy regarding the accuracy of these methods, particularly when applied to structural data at a lower resolution than required to define the subfields (see Wisse et al., 2014). However, studies have shown that this method has good correlation with hand tracing methods (Van Leemput et al., 2009) and the results reported here are consistent with a previous study using a hand-tracing approach (Wang et al., 2010). Another limitation was that date of the trauma was estimated and thus was not an ideal measure of trauma chronicity. Although we
did not see evidence of trauma chronicity effects on subfield volume, additional studies should be conducted with a more precise chronicity variable.

In conclusion, we report smaller DG volume as a function of PTSD diagnosis and severity in a large cohort of war veterans. Our automated subfield segmentation approach results are consistent with a previous study that used time intensive hand drawing methods and also found smaller DG volume. The results from this study support the hypothesis suggesting that the DG plays a role in maintaining PTSD symptoms, although additional studies are required to examine whether a pattern separation deficit is the mechanism that links volume to symptom profiles.

**Role of funding source**

This work was supported by VA Rehabilitation Research and Development grant # I01RX000216, VA Biomedical Laboratory Research and Development grant # I01BX003477, the National Center for PTSD, and the VA Clinical Science Research and Development Service. The funding sources had no other involvement in the study design and interpretation, writing or the decision to submit the article for publication.

**Conflicts of interest**

None.

**Appendix A. Supplementary data**

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.jpsychires.2017.09.007.

**References**


