Hippocampal subfields in a developmental population: Assessing the reliability of fully-automated segmentation

Kelsey L. Canada, Morgan Botdorf, & Tracy Riggins
University of Maryland, College Park

UNIVERSITY OF MARYLAND

Neurocognitive Development Lab

Introduction

- Approaches to segmenting the hippocampus vary in the field of developmental cognitive neuroscience.
- The current "gold-standard" method relies on manual tracing of hippocampal subfields.
- This standard may be a barrier for researchers who do not focus on the hippocampus or are using large datasets as it requires expertise in neuroanatomy, is time-consuming, and often relies on higher image quality than is typically collected.
- **The proposed preliminary study sought to explore the reliability of fully-automated hippocampal subfield segmentation in 4- to 8-year-old children.**

Approach:
- Manual ("gold-standard") segmentations from T2 images were compared to:
  1) Semi-automated segmentations derived from a study-specific atlas using T2 images in ASHS
  2) Fully-automated segmentations derived from T1 images in Freesurfer
  3) Fully-automated segmentations derived from T1 and additional T2 images in Freesurfer
- These comparisons provide an initial assessment of the use of manual, semi-automated, and fully-automated segmentations in a pediatric population.

Methods

Participants
- 20 4- to 8-year-old participants

MRI Data Collection
- High resolution T1-weighted magnetization-prepared rapid gradient-echo (MPRAGE) sequence of 176 contiguous sagittal slices (.9 mm isotropic voxel size; 1900 ms TR; 2.32 ms TE; 900 ms inversion time; 9-degree flip angle; 256 x 256 pixel matrix).
- Ultra-high resolution (.4mm x .4mm x 2mm) structural scans of medial temporal lobe (MTL) were acquired with a T2-weighted fast spin echo sequence (TR=4120ms, TE=41ms, 24 slices, 149 degree flip angle).
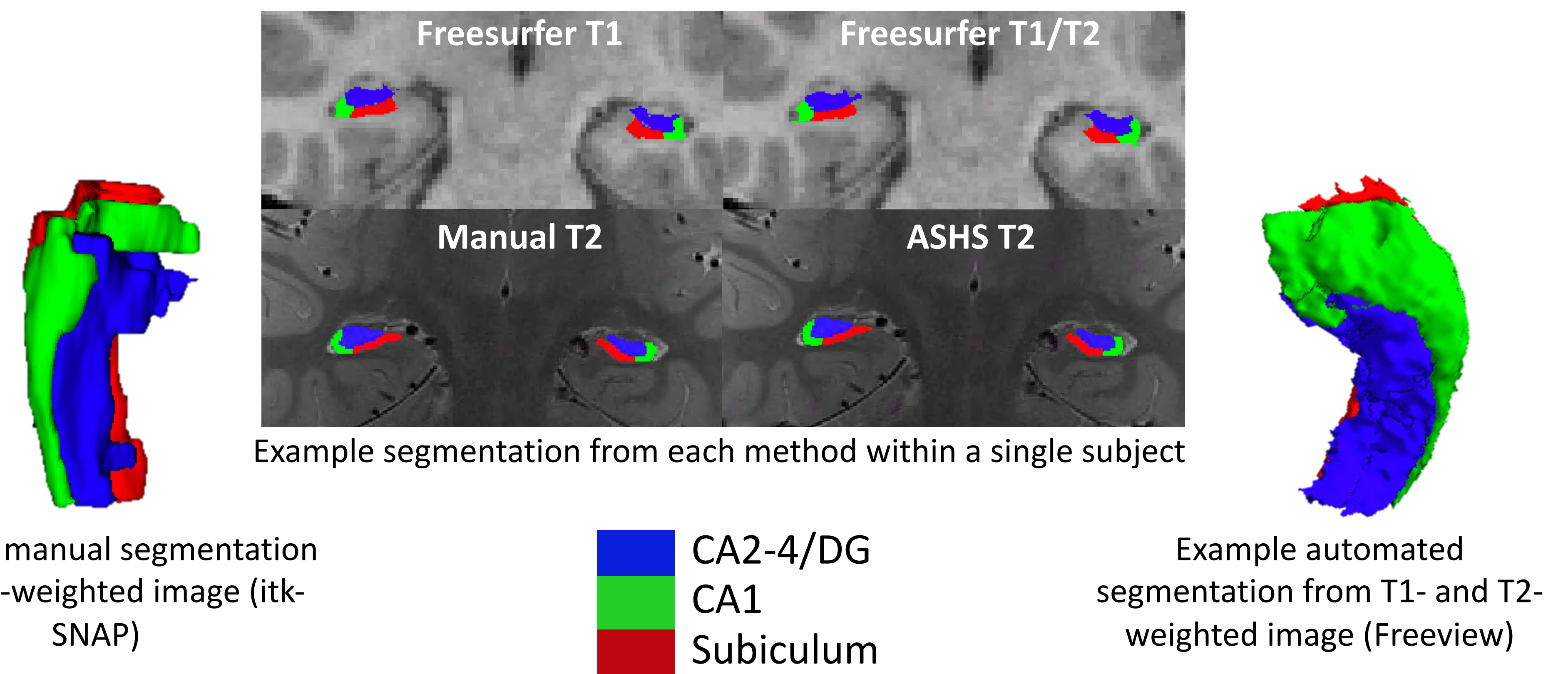
MRI Analysis
- Manual segmentations derived from T2 images for bilateral subiculum, CA1, and CA2-4/DG volumes using a protocol adapted from Joie et al. (2010).
- Semi-automated segmentations for bilateral subiculum, CA1, and CA2-4/DG volumes were derived from T2 images using a protocol adapted from Joie et al. (2010) used in conjunction with the Automatic Segmentation of Hippocampal Subfields software (ASHS, Yushkevich et al. 2014).
- Automated segmentations generated by Freesurfer (Version 7.1.0) using 1) a T1 image only and 2) T1 and T2 images and the "CA" segmentation, combining CA3 & CA4 labels.
- This resulted in bilateral CA2-4/DG, CA1, and subiculum volumes derived from Freesurfer.

Statistical Analyses
- Reliability was assessed using ICC (2,1) for absolute agreement
- Correlations between volumes were assessed using Pearson's correlation coefficient (r)

Results

1) ASHS segmentations are reliable when compared to manual segmentations and most closely align with manual segmentations.

2) Manual and Freesurfer volumes did not show reliable agreement.

3) Descriptively, Freesurfer segmentations using an additional T2-weighted image did not show marked improvement upon segmentations using only T1-weighted images.

Example segmentation from each method within a single subject

Example manual segmentation from T2-weighted image (itk-SNAP)

- CA2-4/DG (blue)
- CA1 (green)
- Subiculum (red)

Example automated segmentation from T1- and T2-weighted image (Freeview)

Comparison of agreement between methods

Subfield	Manual v ASHS ICC	Manual v ASHS r	Manual v Freesurfer T1 only ICC	Manual v Freesurfer T1 only r	Manual v Freesurfer T1 and T2 ICC	Manual v Freesurfer T1 and T2 r
Subiculum						
Right	0.736 (-.069-.930)	0.927**	.038 (-.068-.231)	.206	.033 (-.051-.198)	.282
Left	0.835 (.123-.954)	0.949**	.086 (-.082-.348)	.449*	.071 (-.069-.308)	.513*
CA1						
Right	0.763 (.029-.930)	0.927**	.361 (-.052-.680)	.528*	.354 (-.048-.672)	.505*
Left	0.627 (-.069-.879)	0.838**	.364 (-.049-.683)	.535*	.366 (-.049-.684)	.527*
CA2-4/DG						
Right	0.919 (.778-.969)	0.937**	.428 (.013-.723)	.515*	.377 (-.062-.700)	.534*
Left	0.803 (-.047-.951)	0.940**	.300 (-.087-.634)	.493*	.341 (-.092-.679)	.616**

Note: ** denotes significant effect at p < .001 level; * denotes significant effect at p < .05 level. Raw volumes were used for comparisons between segmentations. Numbers in parentheses represent the 95% confidence interval.

Discussion

- Suggests difficulty in using fully-automated subfield segmentations to examine subfield volume in a pediatric population, however the sample size is quite limited and additional work is needed due to the inter-individual variability in hippocampal morphometry and volume.

- The use of T1-weighted images with lower resolutions for fully-automated segmentations may have also contributed to the differences in estimated volumes from semi-automated segmentations.

- Given the specificity of the manual segmentations and semi-automated segmentations, it is possible that future work using a harmonized protocol in comparison to Freesurfer volumes would yield more reliable results with higher agreement between estimated volumes.

Acknowledgements