

Harmonized Segmentation Protocol of the Hippocampal Tail on High-Resolution *in vivo* MRI from the Hippocampal Subfields Group

R. de Flores¹, K.L. Canada², T. Brown³, N.J. Gervais⁴, A. Maass⁵, G. Rădman⁶, J. Shine⁵, H.L. Tucker⁶, E.N. Malloy⁵, J.N. Adams⁷, M.B. Reinke⁸, A. Bakker⁹, D. Berron⁵, M. Dalton¹⁰, K.M. Kennedy¹¹, R. La Joie¹², S.G. Mueller¹², N. Ofen¹¹, R. Olsen¹³, N. Raz¹⁴, T. Riggins¹⁵, K.M. Rodrigue¹¹, C.E. Stark⁷, L. Wang¹³, L.E.M. Wisse⁶, P.A. Yushkevich¹⁶, V. Carr³, A.M. Daugherty¹⁷ on behalf of the Hippocampal Subfields Groups (HSG)

¹Caen-Normandy University; ²University of Massachusetts Amherst; ³Georgia Institute of Technology; ⁴University of Groningen; ⁵Germany Centre for Neurodegenerative Disease (DZNE); ⁶Lund University; ⁷University of California Irvine; ⁸San Jose State University; ⁹Johns Hopkins University; ¹⁰University of Sydney; ¹¹University of Texas at Dallas; ¹²University of California San Francisco; ¹³University of Toronto; ¹⁴Stony Brook University; ¹⁵University of Maryland; ¹⁶University of Pennsylvania; ¹⁷Wayne State University

Aim Develop a harmonized segmentation protocol for hippocampal tail segmentation on high-resolution MRI (distinct from the body protocol)

Background

- Segmentation of human hippocampal subfields using MRI is of great interest for research on development, aging, and neurodegenerative disease
- Convergence of evidence across studies is hampered by differences in segmentation protocols
- The Hippocampal Subfields Group was formed to harmonize a segmentation protocol for medial temporal lobe structures
- A distinct protocol was created for segmenting the hippocampal tail using *in vivo* 3T structural MRI (T2-weighted with a 0.4x0.4x2mm³ resolution)

Methods

- Does not distinguish between subfields & is based on a novel histological reference data set labeled by multiple expert neuroanatomists
- Boundary rules were developed by 4 individuals within the working group & include: Posterior / Inferior / Superior / Lateral / Medial boundaries
- Initial check for feasible reliability
 - Two expert raters with at least 4 years of experience, 4 scans
- Feedback collection via Delphi procedure
 - Panelists were recruited via an open call to the HSG member email-list, website, and over various social media platforms
 - 9-point Likert scale to assess both agreement with the rule and clarity of the rule
- Formal reliability analysis of the consensus protocol
 - Four expert raters with at least 4 years of experience, 22 scans

Delphi procedure results

- 21 participating laboratories, 95% had expertise with hippocampal measures on 3T data, and 86% had experience with relevant T2-weighted images

Rule Definition	Question	N	Range	M	SD	% Agree	binomial test
Posterior boundary	How clear is the boundary description?	21	6-9	7.76	1.04	100	< 0.001
	Do you agree with the boundary rule?	21	6-9	8.67	0.80	100	< 0.001
Inferior boundary	How clear is the boundary description?	21	7-9	8.43	0.87	100	< 0.001
	Do you agree with the boundary rule?	21	7-9	8.90	0.44	100	< 0.001
Superior boundary	How clear is the boundary description?	21	5-9	8.05	1.12	95.24	< 0.001
	Do you agree with the boundary rule?	21	8-9	8.76	0.44	100	< 0.001
Lateral boundary	How clear is the boundary description?	21	5-9	8.14	1.24	90.48	< 0.001
	Do you agree with the boundary rule?	21	8-9	8.86	0.36	100	< 0.001
Medial boundary	How clear is the boundary description?	21	4-9	7.86	1.35	95.24	< 0.001
	Do you agree with the boundary rule?	21	6-9	8.62	0.80	100	< 0.001
Plan to not define subfields	Do you agree with the boundary rule?	20	5-9	8.45	1.15	95	< 0.001

Ratings were made on a 9-level Likert scale; responses > 5 were considered endorsement of clarity or agreement. Consensus was determined by statistical majority agreement with a significant binomial test (alpha = 0.05). N = 21 laboratories participated

- After a single iteration, all boundary definitions were deemed clear and reached consensus**

Reliability assessment

- Intra- and inter-rater reliability of the harmonized protocol**

Region	Volume (n=24)			Intra-Rater (3 raters, n = 11)		Inter-Rater (3 raters, n = 24)	
	Rater 1	Rater 2	Rater 3	Avg. ICC(2)	Avg. DSC	ICC(2,k)	Avg. DSC
HT-L	423.79 ± 188.46	472.65 ± 191.86	451.95 ± 189.54	0.99	0.88 ± 0.04	0.99	0.88 ± 0.04
HT-R	353.29 ± 163.65	395.89 ± 181.46	395.67 ± 197.72	0.98	0.86 ± 0.04	0.98	0.86 ± 0.04

Note: Reliability was tested on a blinded dataset by two expert raters and 1 novice rater (k = 3 raters), all naïve to the protocol. Average ICC, and average DSC (+ SD) are reported. ICC— intra-class correlation coefficient; DSC—Dice Similarity Coefficient; HT—Hippocampal Tail

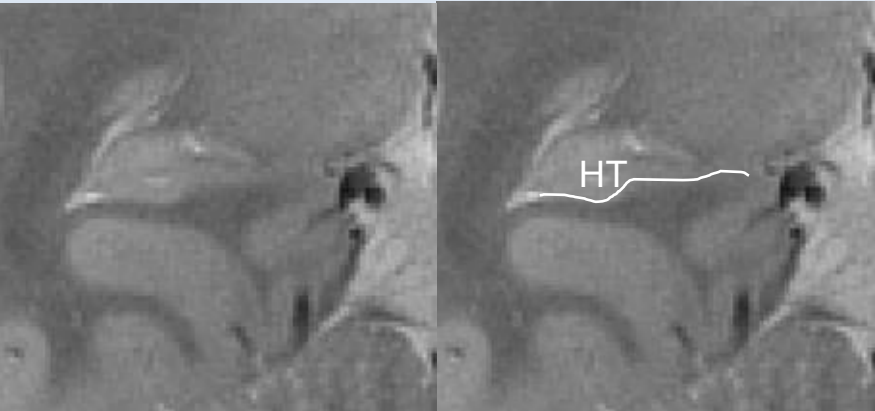
- Average inter-rater dice similarity coefficient by scan demographic**

Region	Age-Group Comparisons			Comparisons by Cognitive Diagnosis			
	Children (n=7)	Adults (n=17)	p-value	Cognitive Typical (n = 5)	MCI (n = 5)	Dementia (n=3)	p-value
HT-L	0.89 ± 0.02	0.88 ± 0.04	0.58	0.88 ± 0.04	0.89 ± 0.04	0.87 ± 0.03	0.69
HT-R	0.88 ± 0.03	0.85 ± 0.04	0.13	0.83 ± 0.03	0.87 ± 0.05	0.86 ± 0.03	0.14

Note: Average dice similarity coefficient (DSC) + SD is reported stratified by age group and cognitive diagnosis that was associated with the scan. Distributions were compared for statistically significant difference by Mann-Whitney U and Kruskal-Wallis tests (α = 0.05). Comparisons by cognitive diagnosis were made only among scans collected by ADNI and included standardized protocol for diagnosis. MCI—Mild cognitive impairment; HT—Hippocampal Tail.

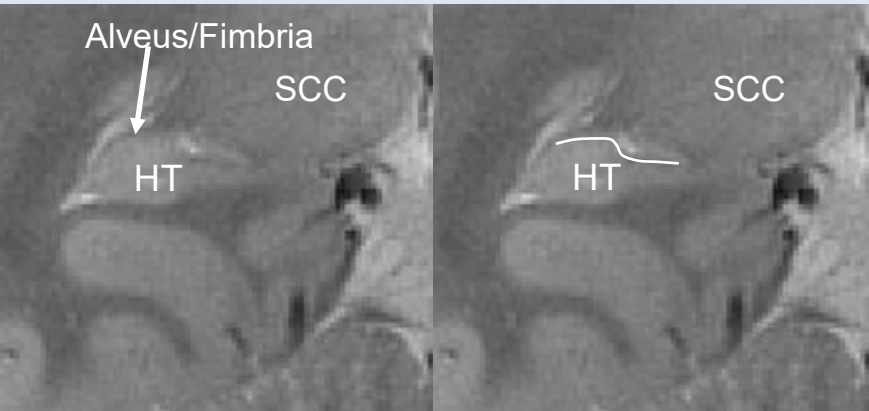
HSG Harmonized Protocol for segmenting the hippocampal tail

Inferior boundary



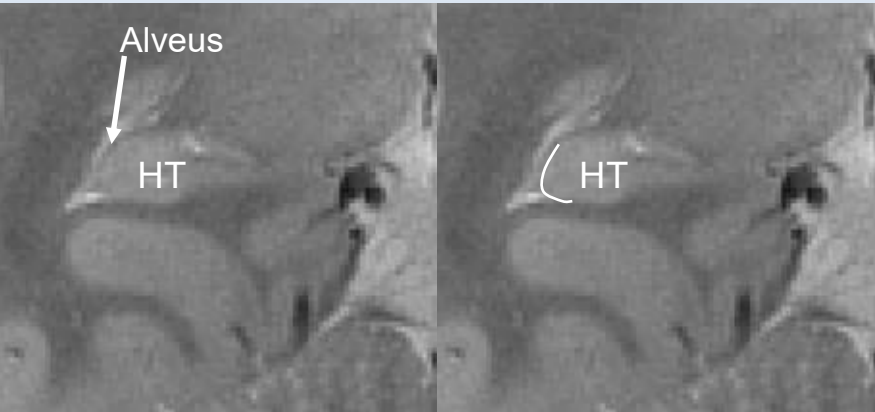
The border between the gray matter of the hippocampus and the white matter located inferior to it

Superior boundary



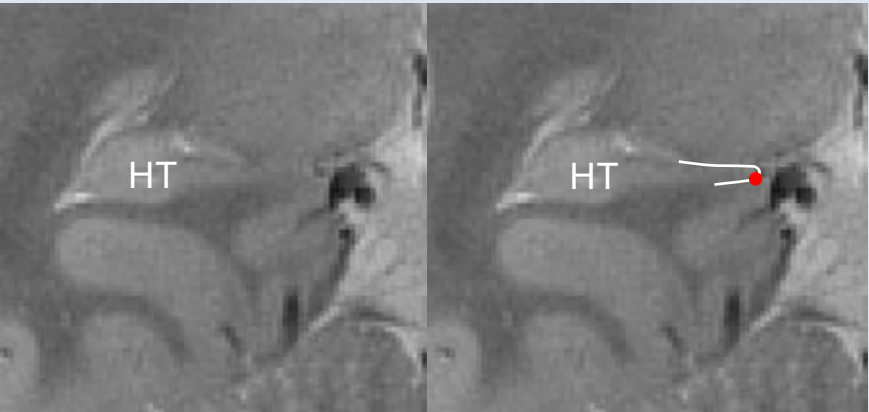
Alveus and/or CSF if the alveus is not visible

Lateral boundary



Fimbria/fornix, the white matter of the splenium of the corpus callosum and/or CSF

Medial boundary



Most supero-medial corner of the parahippocampal gyrus



Conclusion

- The HSG developed a reliable and accessible harmonized protocol for consistent hippocampal tail segmentation across research sites and MRI scanners.
- Complements the existing HSG hippocampal body protocol, with harmonized hippocampal head & MTL cortex protocols currently in development.
- Establishes standardized guidelines to improve comparability of hippocampal measurements across developmental, aging, and clinical studies.
- Optimized for high-resolution T2-weighted MRI while remaining flexible enough to accommodate future technological advances and methodological refinements.



We are looking for volunteers for working groups! Email if interested: ana.daugherty@wayne.edu