A comparison of manual and automated methods of generating hippocampal volume measurements

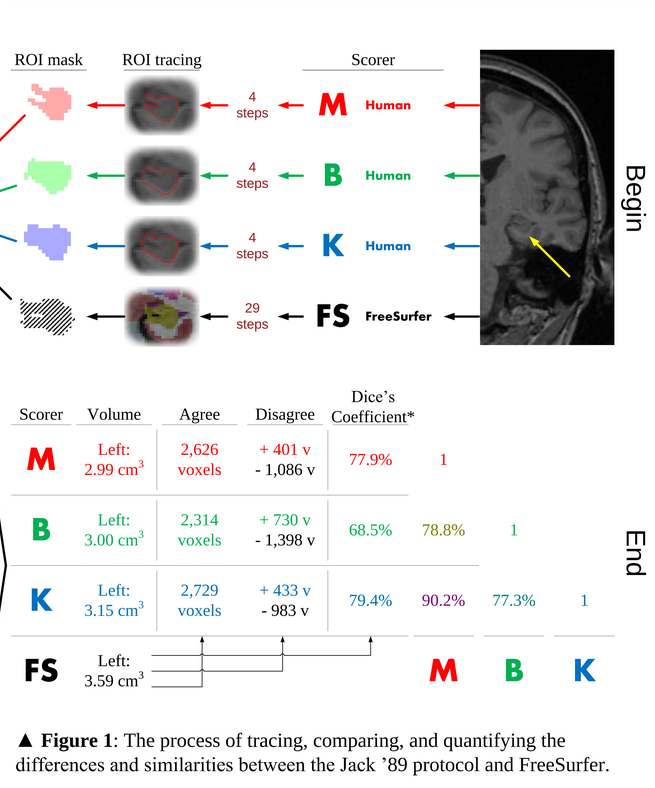
M. F. SCHMIDT[1], J. M. STORRS[2], K. B. FREEMAN[3], T. H. MOSLEY[4]

1 Program in Neuroscience, 2 Dept of Radiology, 3 Dept of Psychiatry and Human Behavior, 4 Dept of Medicine, Geriatrics
University of Mississippi Medical Center, Jackson, MS

Introduction

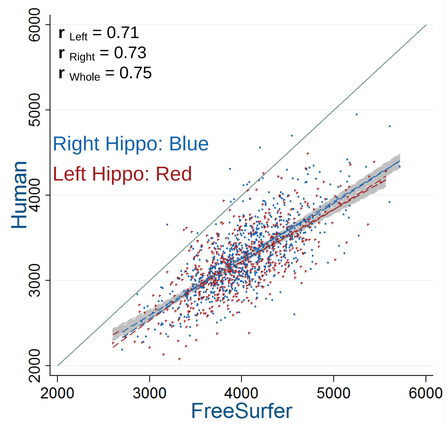
Hippocampal volume is a biomarker negatively associated with depression, schizophrenia, dementia, and aging. Thousands of studies have required hundreds of thousands of man-hours to trace hippocampal boundaries in MRI scans. Are computer algorithms good enough to replace human tracers and save us from this drudgery and expense? To answer this question, we manually traced hippocampi and ran FreeSurfer on 619 of 710 MRI scans from healthy volunteers in the Genetics in Microangiopathic Brain Injury study and compared them.

Methods

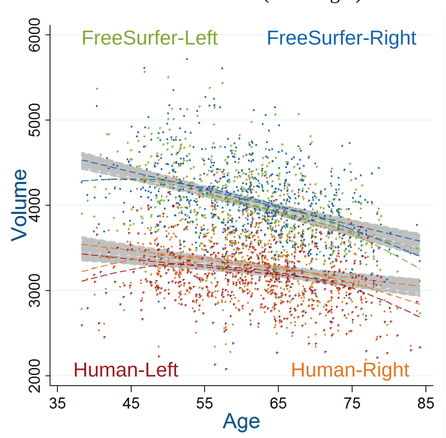


▲ **Figure 1**: The process of tracing, comparing, and quantifying the differences and similarities between the Jack '89 protocol and FreeSurfer.

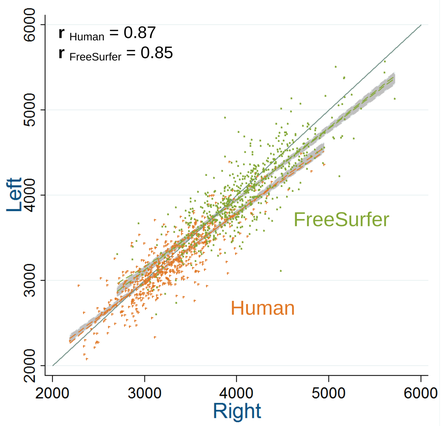
Results



▲ **Figure 2**: FreeSurfer volumes are consistently larger than manual, shown by the data points falling to the right of the parity line. The volumes are correlated at r=0.75 for whole (left + right) volumes.



▲ **Figure 3**: FreeSurfer's volumes are larger, but re-tell a similar story about atrophy over age. We expect a year-over-year atrophy rate between 0.3% and 2.0% based on the literature. Our human scoring reports 0.3% and FreeSurfer reports 0.5%.



▲ **Figure 4**: Left and Right volumes correlate highly, with no laterality bias, regardless of method.

▼ **Table 1**: Mean values from each method

	Manual (std err)	FreeSurfer (std err)
Left	3,219 (± 16) mm^3	3,997 (± 18) mm^3
Right	3,295 (± 17) mm^3	3,219 (± 19) mm^3
Combined	6,515 (± 32) mm^3	8,051 (± 36) mm^3

▼ **Table 2**: Comparisons between methods (H vs FS)

	Correlation Coefficient	Dice's Coefficient*	Jaccard Index**
Left	0.71	0.751	0.601
Right	0.73	0.758	0.610
Combined	0.75	0.755	0.606

▼ **Table 3**: Human inter-rater comparisons

	Correlation Coefficient	Dice's Coefficient*	Jaccard Index**
Left	0.88	0.813	0.687
Right	0.75	0.802	0.672
Combined	0.84	0.807	0.679

Conclusions

FreeSurfer's segmentation methodology differs from manual tracing, resulting in larger values for hippocampal volumes. This difference appears to be protocol-dependent. FreeSurfer is capable, despite the minor disagreements with human scorers, of making similar predictions and uncovering similar associations as human scoring. As algorithms evolve, however, it is important that they are continually validated not just directly against human scoring, but also indirectly against the discoveries made with human scoring.

References

Barnes J, et al. Neurobiology of Aging 2009; 30:1711-1723
Fischl B, et al. Neuron 2002; 33: 341-355
Jack CR, et al. Radiology 1989; 172: 549-554

$$* \text{ Dice} = \frac{2\,|H \cap F|}{|H| + |F|} \qquad ** \text{ Jaccard} = \frac{|H \cap F|}{|H \cup F|}$$

Thanks!

- UMMC Program in Neuroscience
- The MIND Center at UMMC
- Athinoula A Martinos Center
- UMMC Dept of Biostatistics
- Mississippi Center for Supercomputing Research
- NIH NHLBI Grant U01 HL54434
- NIH NINDS Grant R01 NS41558

Further Information

For supplemental information, the abstract, a copy of the poster, or contact information, visit http://sfn2013.mfs.ms/ or scan the code to the right.