Comparing Automated Subcortical Volume Estimation Methods; Amygdala Volumes Estimated by FSL and FreeSurfer Have Poor Consistency

Sadil, Patrick & Lindquist, Martin A.

Johns Hopkins Bloomberg School of Public Health, Baltimore, MD

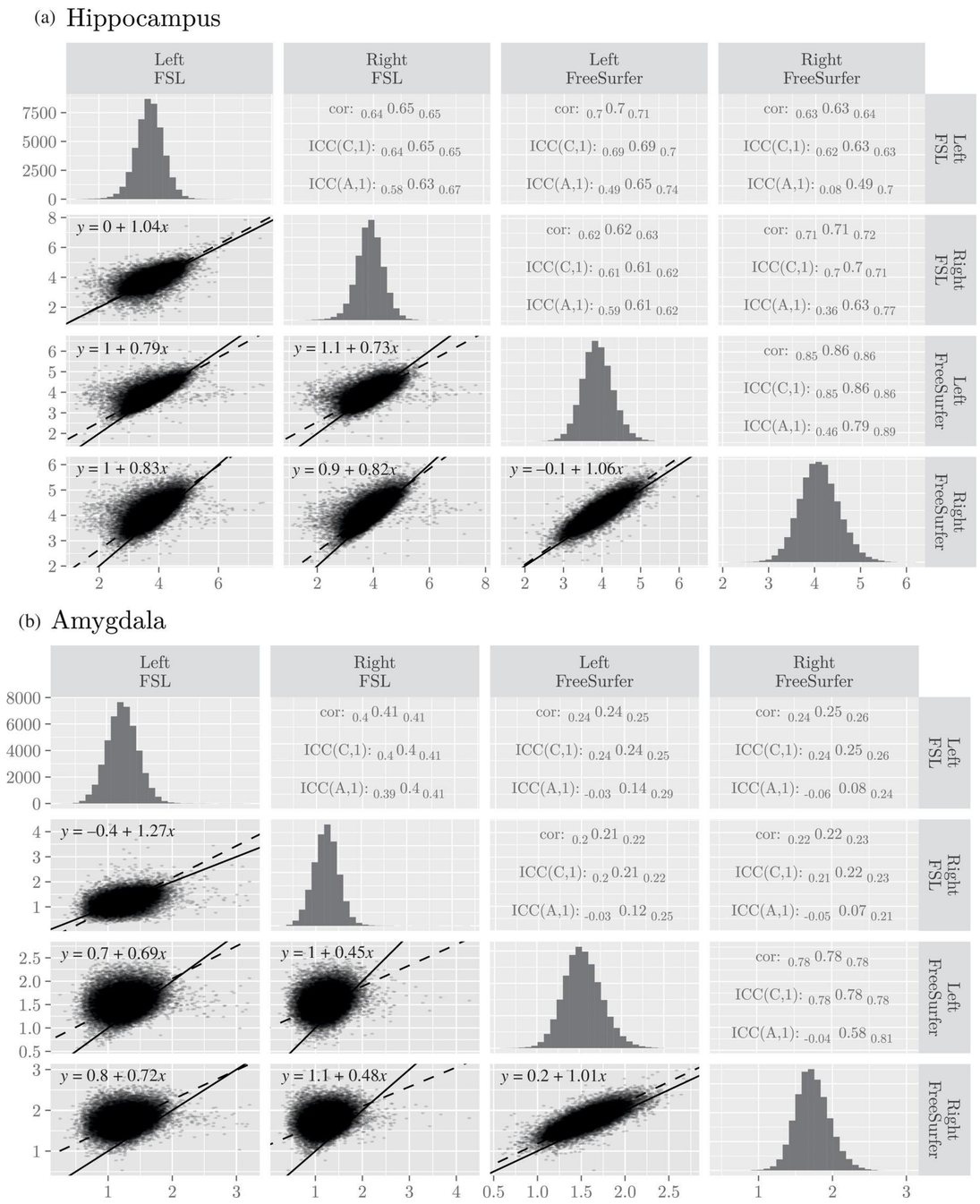
Introduction

- Two software suites have techniques for estimating subcortical volumes are popular: FSL and FreeSurfer
- Both tools are often reasonable choices
- **Aim**: Measure the reliability of methods

Methods

- Data from the UK Biobank: N=45,743
- Reliability measured using ICC
- Simulated studies to assess how often low consistency leads to opposing significance and magnitude

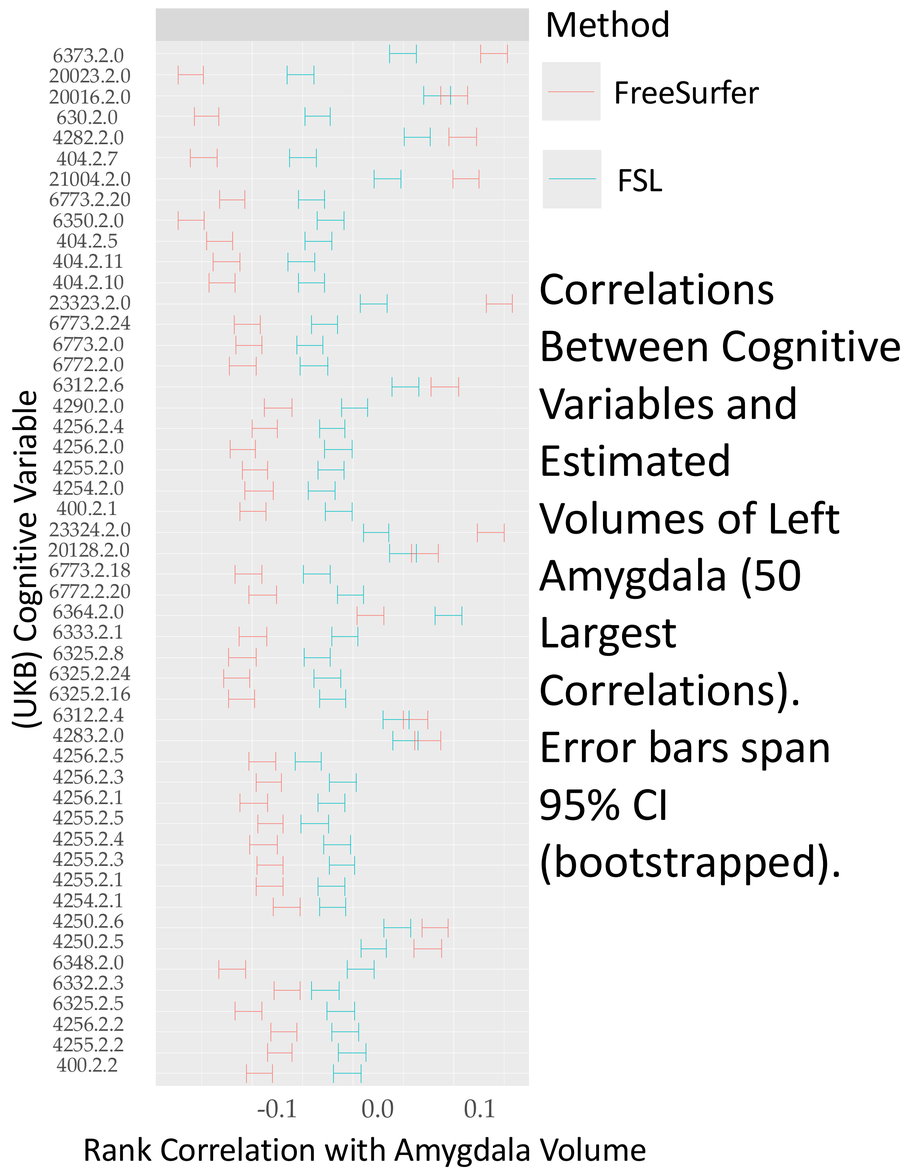
Comparison of Volumes Across Software



Structure	Hemisphere	ICC(C,1)	ICC(A,1)
Accumbens	Left	$_{0.57}0.58_{0.58}$	$_{0.05}0.44_{0.65}$
	Right	$_{0.56}0.57_{0.57}$	$_{-0.03}0.38_{0.63}$
Amygdala	Left	$_{0.24}0.24_{0.25}$	$_{-0.03}0.14_{0.29}$
	Right	$_{0.21}0.22_{0.23}$	$_{-0.05}0.07_{0.21}$
Caudate	Left	$_{0.85}0.85_{0.86}$	$_{0.75}0.83_{0.88}$
	Right	$_{0.86}0.86_{0.87}$	$_{0.62}0.82_{0.90}$
Hippocampus	Left	$_{0.69}0.69_{0.70}$	$_{0.49}0.65_{0.74}$
	Right	$_{0.70}0.70_{0.71}$	$_{0.36}0.63_{0.77}$
Pallidum	Left	$_{0.68}0.68_{0.69}$	$_{-0.09}0.41_{0.71}$
	Right	$_{0.66}0.67_{0.67}$	$_{0.04}0.51_{0.73}$
Putamen	Left	$_{0.79}0.79_{0.80}$	$_{0.54}0.74_{0.84}$
	Right	$_{0.82}0.83_{0.83}$	$_{0.56}0.77_{0.87}$
Thalamus	Left	$_{0.82}0.82_{0.83}$	$_{-0.08}0.49_{0.79}$
	Right	$_{0.83}0.83_{0.83}$	$_{-0.09}0.51_{0.81}$

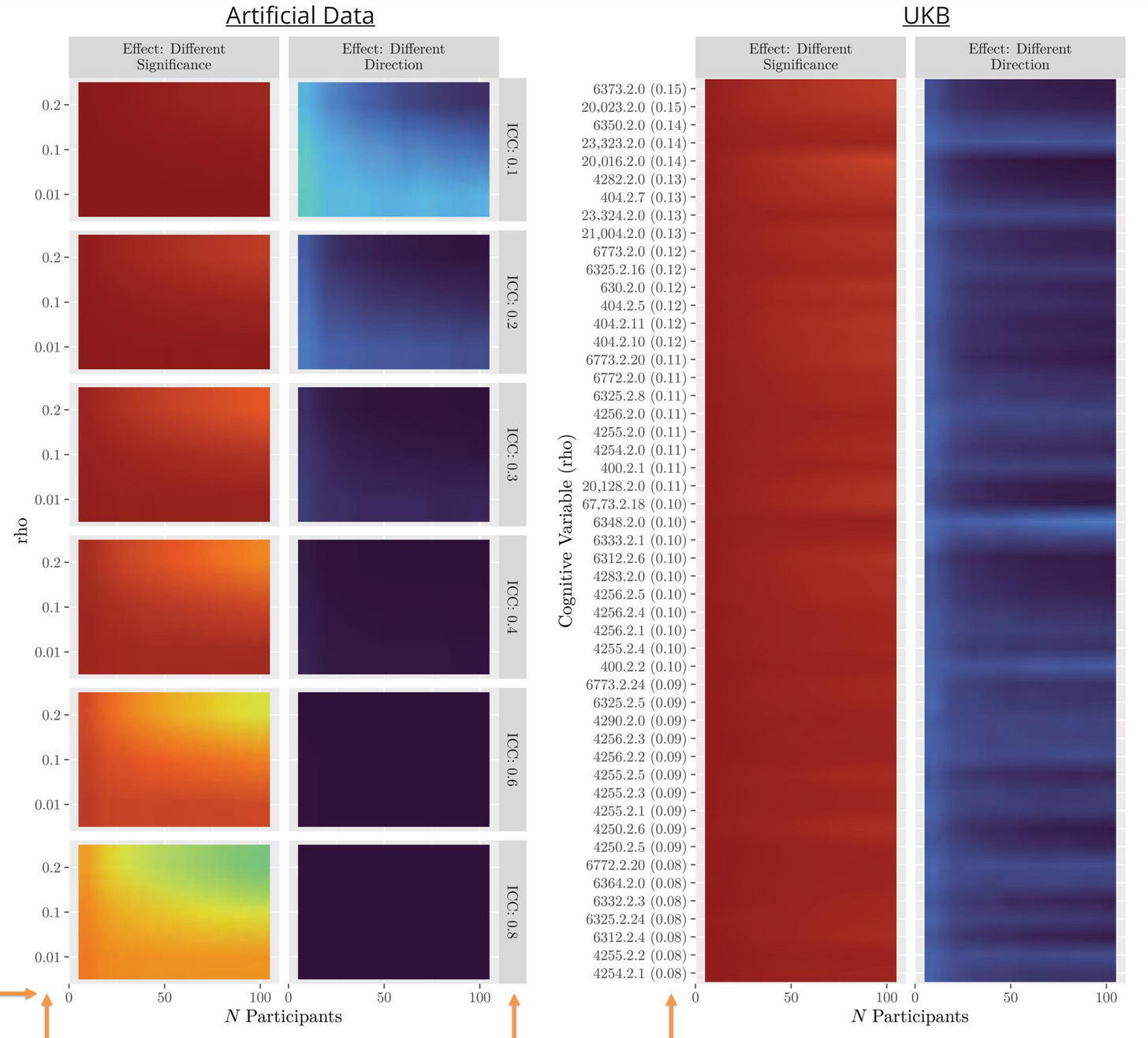
Reliability for Subcortical Structures. Subscripts indicate 95% CI.

Observed Correlations



Correlations Between Cognitive Variables and Estimated Volumes of Left Amygdala (50 Largest Correlations). Error bars span 95% CI (bootstrapped).

Impacts of Low Consistency in Simulations



Product–moment correlation with the outcome variable

Participants in simulations

Intraclass Correlation Between Measures of Volume

Field ID (Observed Correlation).

Summary

- Methods exhibit:
  - At least fair consistency for most regions
  - Poor consistency for Amygdala
- With poor consistency, methods often produce results with opposing significance, and occasionally, opposing magnitude

References

Alfaro-Almagro, F., Jenkinson, M., Bangerter, N. K., Andersson, J. L., Griffanti, L., Douaud, G., ... & Smith, S. M. (2018). Image processing and Quality Control for the first 10,000 brain imaging datasets from UK Biobank. *Neuroimage, 166*, 400-424.

Fischl, B. (2012). FreeSurfer. *Neuroimage*, 62(2), 774-781.

Patenaude, B., Smith, S. M., Kennedy, D. N., & Jenkinson, M. (2011). A Bayesian model of shape and appearance for subcortical brain segmentation. *Neuroimage, 56*(3), 907-922.