

# Independent Testing of a Publicly Available CNN Tool for Hippocampal Image Segmentation

Fraser N Sneden, Karen J Ferguson, Susana Muñoz Maniega, Maria Valdés Hernández, Joanna M Wardlaw  
Centre for Clinical Brain Sciences, University of Edinburgh

## Hippocampus

The hippocampus lies within the temporal lobe. It has an important role in the consolidation of long-term memory and other processes. Hippocampal volumes are an important biomarker of ageing and dementia. However, manual segmentation of the hippocampus is time consuming and prone to inter-rater variability, especially in images with significant atrophy.

## HippMapp3r

HippMapp3r is a CNN-based automated image segmentation tool.

### Advantages:

1. Trained on patients from various samples (MCI, AD and temporal lobe epilepsy).
2. Computationally inexpensive.
3. Performs more accurately than other comparable segmentation techniques.

### Aim

Evaluate the use of HippMapp3r in an ageing population, the Lothian Birth Cohort 1936 (LBC1936).

### Hypothesis

1. HippMapp3r will perform differently than described due to variations between the population studied.
2. Our pre-processing methods will differentially impact results.

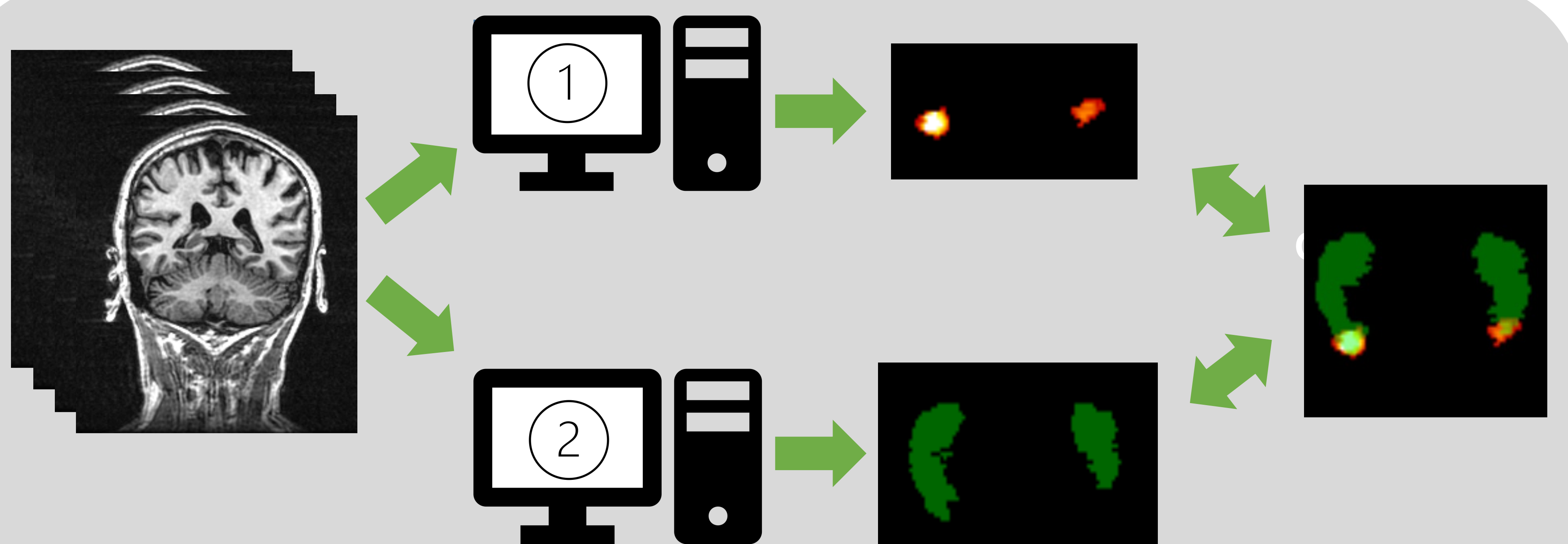
## Methods

We compared 2 different methods:

*Method 1: T1 images in native (skull-stripped) space.*

*Method 2: Linear (affine) transformation to MNI152 space with NiftyReg before segmentation.*

Comparisons are made with our ground truth measurements.



Example of mismatch between methods 1 and 2.

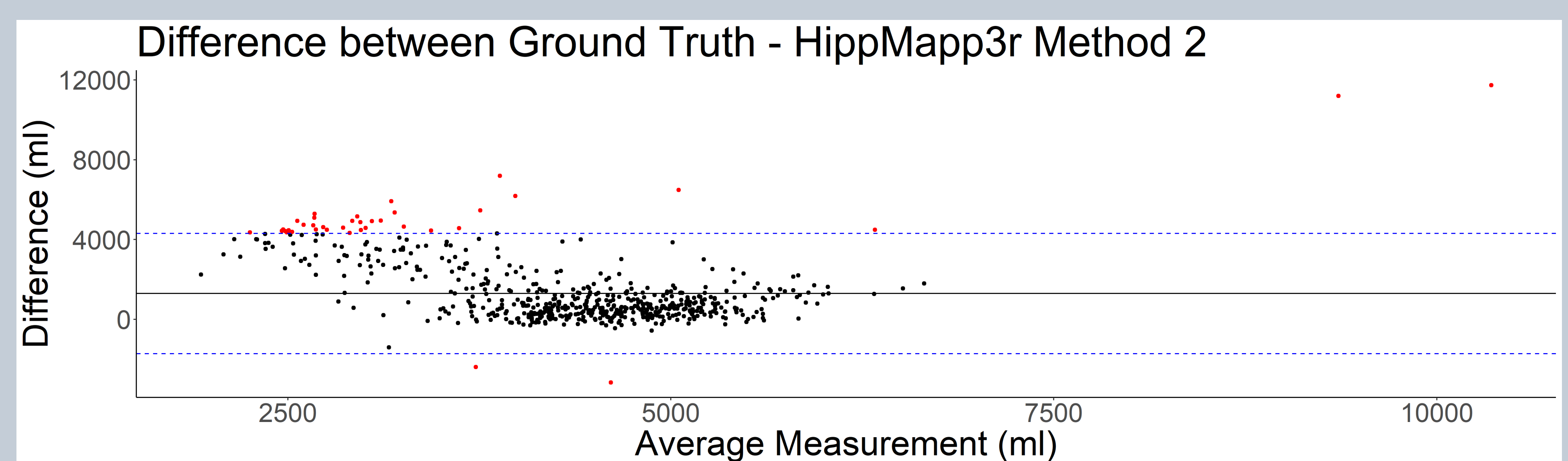
## Results

602 participants were included within our analysis, mean age = 73.

Affine transformation before hippocampal segmentation (Method 2) performed better than Method 1, with respect to Dice overlap ( $p < 0.001$ ). However, the positive predictive value of Method 2 is lower than Method 1.

Comparing Method 2 to our ground truth found that greatest variation was seen in smaller hippocampal volume.

	Method 1			Method 2			Sig.
	Mean	Median	SD	Mean	Median	SD	
Dice	0.594	0.623	0.157	0.676	0.760	0.226	<0.001
Specificity	0.999	0.999	0.000	0.998	0.998	0.002	<0.001
Sensitivity	0.502	0.502	0.173	0.838	0.838	0.153	<0.001
PPV	0.801	0.801	0.106	0.722	0.722	0.242	<0.001



## Discussion

Normalizing images to MNI152 space significantly increased performance metrics. This may be due to alterations in image resolution after co-registration.

Bland-Altman analysis showed a likely proportional bias where HippMapp3r consistently failed to assess smaller hippocampal volumes. HippMapp3r masks sometimes excluded anterior portions and detection of true hippocampal boundaries was affected by the presence of CSF in the hippocampal fissure in some participants.

## Conclusion and Next Steps

Our results reinforce the necessity for collaborative training and developing CNN-based algorithms if they are to be widely used, and multi-center testing prior to their deployment in publicly available repositories.

Further testing on longitudinal data and images of different MR strengths is required.