

## Accounting for changes in data and labeling protocol: improving atlas-based hippocampal segmentation

### Submission No:

2601

### Authors:

[Courtney Bishop](#)<sup>1</sup>, [Dorit Merhof](#)<sup>2</sup>, [Jerome Declerck](#)<sup>3</sup>, [Mark Jenkinson](#)<sup>4</sup>

### Institutions:

<sup>1</sup>FMRIB Centre, University Of Oxford, Oxford, United Kingdom, <sup>2</sup>University of Konstanz, Konstanz, Germany, <sup>3</sup>Siemens Molecular Imaging, Oxford, United Kingdom, <sup>4</sup>FMRIB Centre, University of Oxford, Oxford, United Kingdom

### Introduction:

Development of fully-automated hippocampal segmentation methods is important for investigating hippocampal atrophy, particularly in Alzheimer's Disease (AD). The majority of these methods require prior knowledge gained from manually labeled data (e.g., to generate atlas priors). Substantial variation exists between datasets and labeling protocols [Konrad *et al.* 2009], so applying one prior to another dataset may limit accuracy. Hence, we seek to quantify the performance impact of differing datasets and labeling protocols on automated methods. Furthermore, we propose a method for adapting the priors to account for variations in labeling protocol, requiring a limited amount of manual labels for the new protocol. We evaluate atlas-based segmentation performance using several atlas priors generated from different manual labels and datasets, non-linearly warping these atlases to alternative labeling protocols.

### Methods:

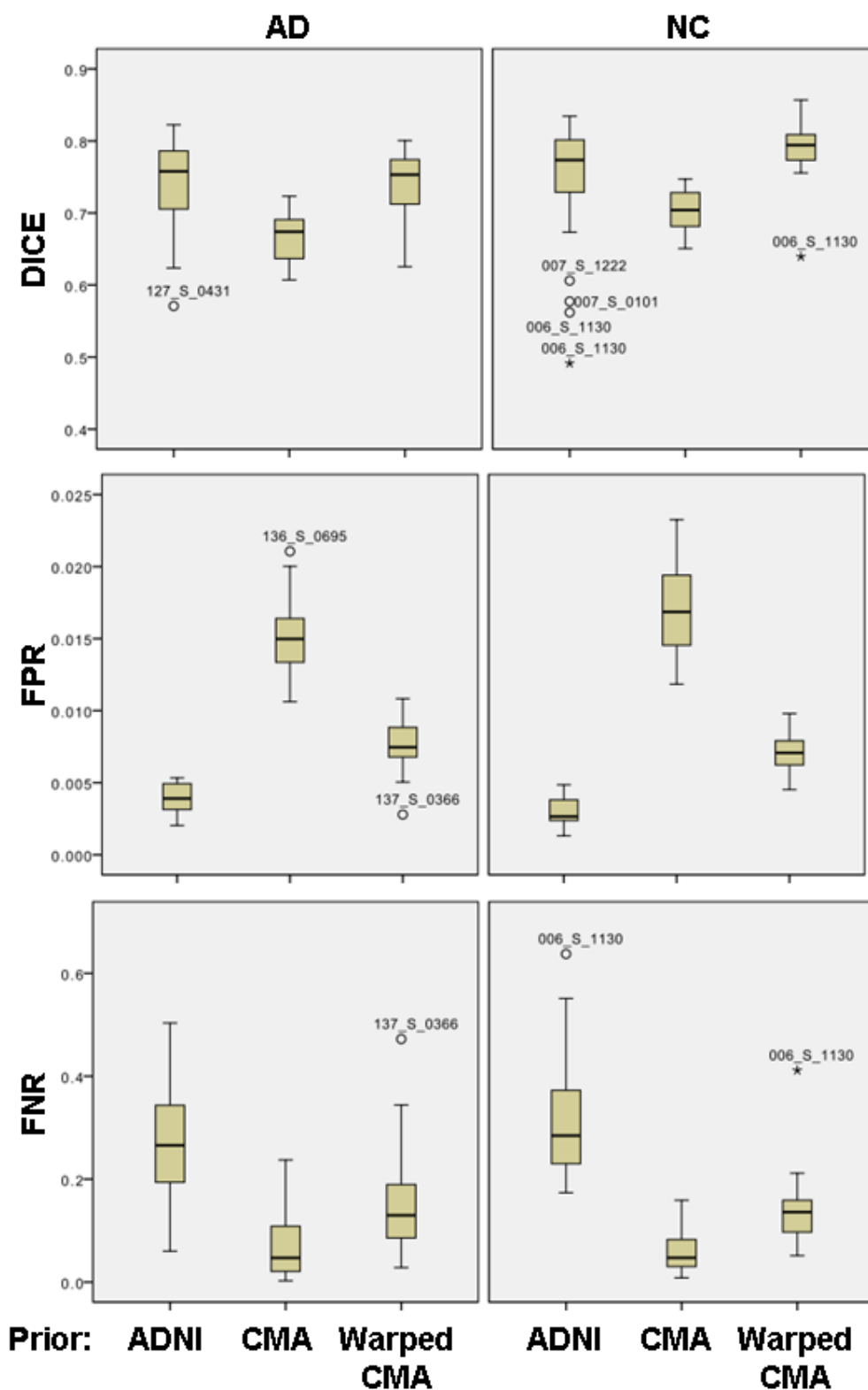
This work uses two sets of T1-weighted MR images (1.5T) with expert manual hippocampal labels, provided by the Alzheimer's Disease Neuroimaging Initiative (ADNI) and the Center for Morphometric Analysis (CMA). Priors are generated from 8 normal controls (NC) and 8 AD patients, separately for ADNI and CMA. An ADNI-warped CMA prior is created by non-linearly warping the CMA prior to the ADNI prior. For each of the priors, we evaluate the segmentation performance of a combined atlas-based and region-growing method, FMASH [Bishop *et al.* 2010], and a pure atlas-based approach (thresholded prior) on an *independent* dataset of 24 ADNI subjects (12 NC, 12 AD). Manual labels provide the gold standard to calculate Dice coefficients, false positive rate (FPR) and false negative rate (FNR), with standard-space maps showing the spatial distribution of errors.

### Results:

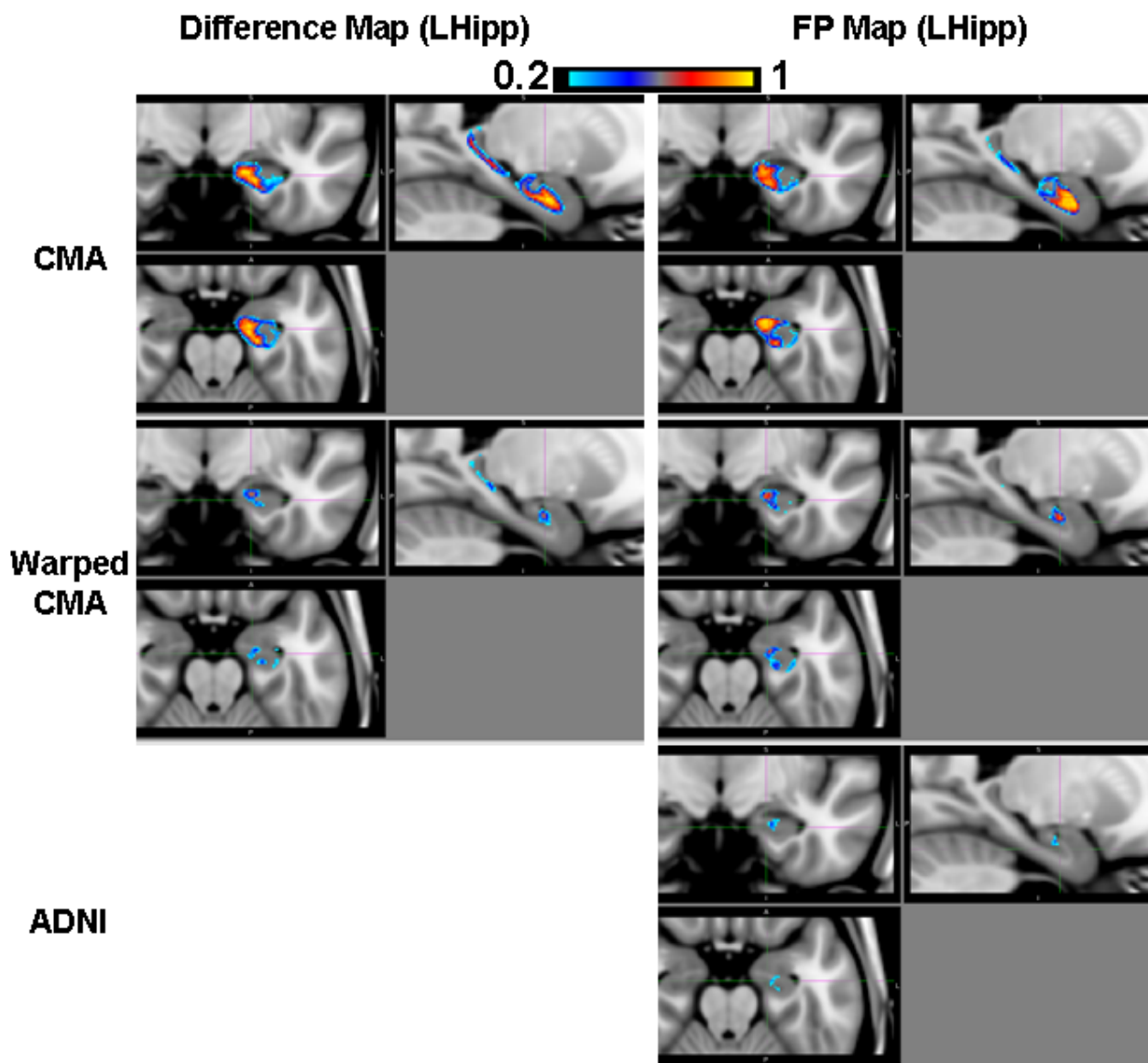
For each prior and performance measure, FMASH performs significantly better than the thresholded prior ( $p < .001$  for all except ADNI prior FNR  $p = .006$ ), so subsequent results are reported for FMASH. Figure 1 presents plots of bilateral Dice, FPR and FNR for FMASH using each hippocampal prior, whilst Figures 2 and 3 show the FMASH FP and FN maps, respectively, and the prior difference maps. As expected, FMASH with the ADNI prior gives significantly higher Dice and lower FPR than the CMA prior on the ADNI test data (Dice:  $0.74 \pm 0.08$ ,  $p < .001$ ; FPR:  $0.003 \pm 0.001$ ,  $p < .001$ ), but a surprisingly higher FNR ( $0.30 \pm 0.12$ ,  $p < .001$ ). Studying the maps in Figure 2, the CMA prior has a much more diffuse, generous labeling than the ADNI prior at medial boundaries, and in particular, anterior regions of the hippocampus head, which translates into FP findings. The warped CMA prior is able to account for most of these differences in labeling protocol, but some FP findings remain in the hippocampus head. Due to such over-estimation of this hippocampal region, both the CMA prior and the warped CMA-prior have a correspondingly low FNR. Surprisingly, the warped CMA prior gives the best overall performance, combining relatively low FPR (ADNI-prior), with low FNR (CMA-prior) and the highest average Dice coefficient ( $0.76 \pm 0.05$ ), consistent across both clinical groups.

### Conclusions:

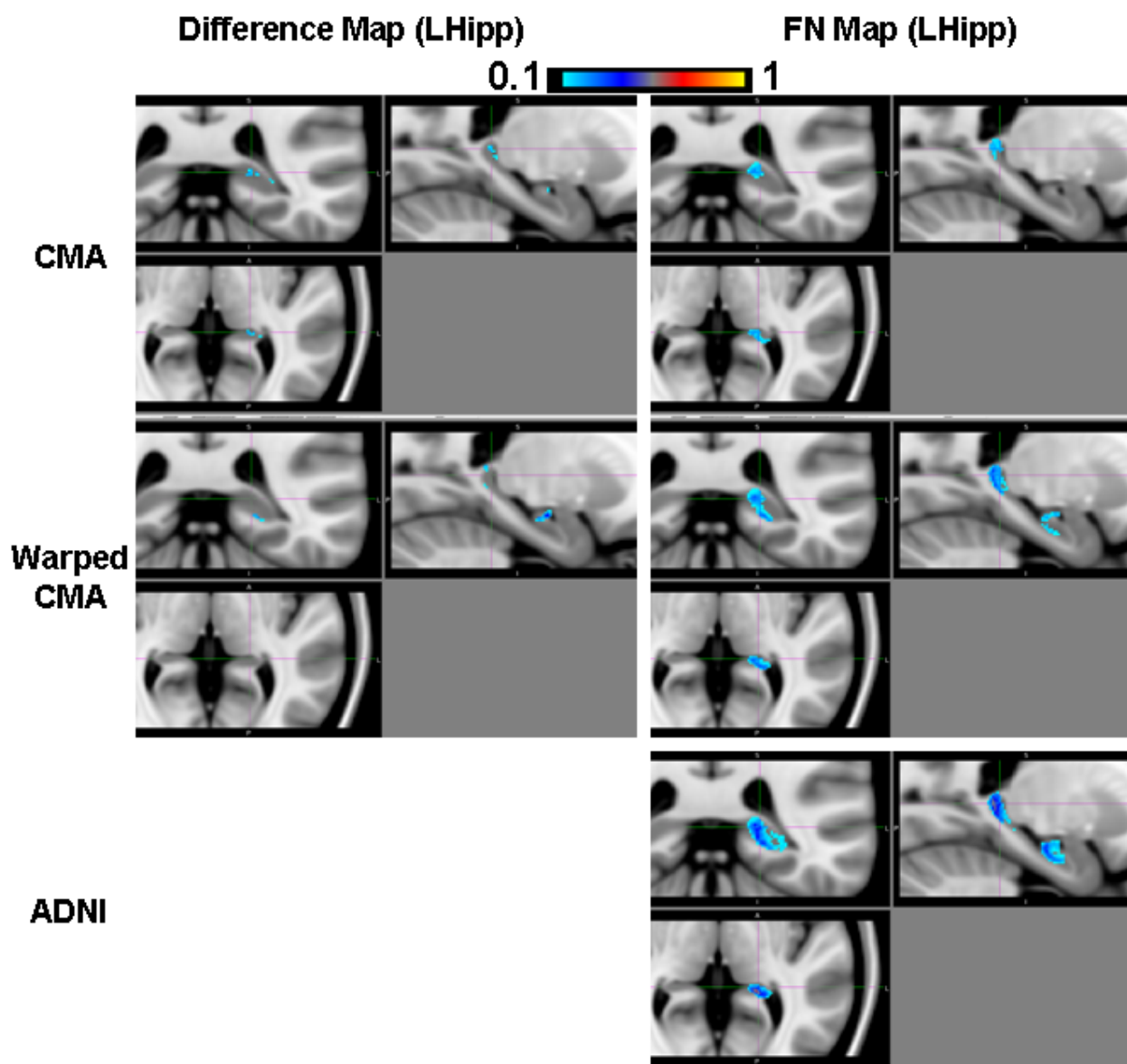
The performance of atlas-based hippocampal segmentation methods is significantly affected by the dataset and the labeling protocol used to generate the atlas prior. The combined effect resulted in Dice coefficients falling by 10% or more when using a prior from one dataset to segment another. However, by non-linearly warping the priors, it was possible to obtain accurate results using the CMA prior on the ADNI dataset. This requires an estimate of the average prior for the ADNI labeling protocol, needing labels for just a few images, potentially even from a different dataset. Thus it appears that cross-dataset priors can be used with little performance degradation as long as the effect of labeling protocol is accounted for. Future work should investigate the applicability of this across more datasets and segmentation methods.



**Figure 1:** Plots of bilateral Dice, FPR and FNR for FMASH on the ADNI test data, using each of three hippocampal priors. Results reported separately for each group: AD patients and normal controls (NC), with the subject ID of any outliers given.



**Figure 2:** Multi-view images of the prior difference maps and the FMASH FP maps for the left hippocampus (LHipp). Here, the difference maps correspond to regions where the current prior probability is greater than that of the ADNI prior.



**Figure 3:** Multi-view images of the prior difference maps and the FMASH FN maps for the left hippocampus (LHipp). Here, the difference maps correspond to regions where the current prior probability is less than that of the ADNI prior.

#### Modeling and Analysis Methods

Segmentation and Parcellation

#### Abstract Information

#### References

- Bishop, C.A. (2010), 'Novel Fast Marching for Automated Segmentation of the Hippocampus (FMASH): Method and Validation on Clinical Data', NeuroImage (Accepted Manuscript)
- Konrad, C. (2009), 'Defining the human hippocampus in cerebral magnetic resonance images - An overview of current segmentation protocols', NeuroImage, vol. 47, pp. 1185-1195