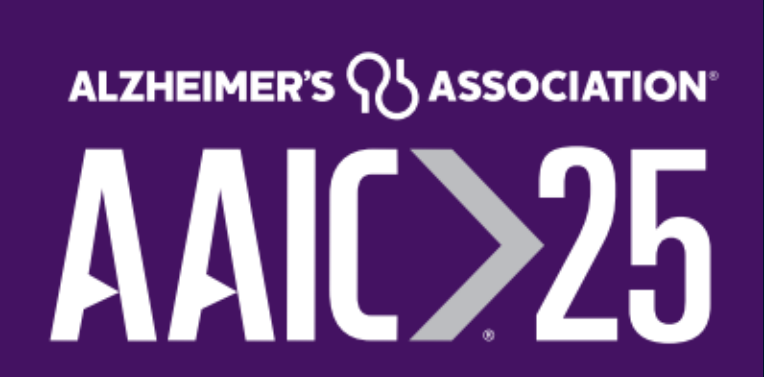


A Comparison of Fully-Automated Segmentation Pipelines: QyScore® vs Freesurfer vs FSL vs ANTs

Luca M. Villa¹, Ayoub Gueddou¹, Mathilde Borrot¹, Zi Hui Su¹, Nicolas Guizard¹, and Elizabeth Gordon¹
¹Qynapse, 2-10 rue d'Oradour-sur-Glane, 75015 Paris, France



BACKGROUND

QYNAPSE

- The gold-standard for brain segmentation is expert manual segmentation, which is both time-consuming and prone to high inter-rater variability.
- Fully automated segmentation tools have emerged, offering faster and more reproducible brain region measurements, though their accuracy and ease of implementation can differ.
- QyScore® is a fully-automated CE-Marked and FDA-cleared medical imaging platform, certified for structural neuroimaging segmentation.
- We evaluated QyScore®'s performance against other leading segmentation pipelines, being FreeSurfer, FSL, and ANTs.

METHODS

- A total of 54 T1-weighted images were used for manual and automatic segmentations of the following brain regions: whole brain grey matter ($n = 30$), whole brain white matter ($n = 30$), hippocampus ($n = 48$), amygdala ($n = 48$), brainstem, cerebellum, caudate, putamen, thalamus, globus pallidus, and lateral ventricles ($n = 49$) (Table 1).
- Automated segmentations were produced by QyScore® v1.13, FreeSurfer v7.4.1, FSL v6.0.6.2, and ANTs v2.5, using default parameters, followed by parameter optimization in the instance of preprocessing or segmentation failures. Consensus manual segmentations, created by three expert neuroradiologists, were used as ground truth.
- The Dice Similarity Coefficient (DSC) between manual gold-standard consensus and automated segmentations were used to compare the accuracy of each automated segmentation pipeline.

Table 1. Sample clinical and demographic details.

Age (SD), Range	Sex (% Female)	Diagnosis	Field Strength
49.8 (21.5), 18-90	51.9	Healthy control: 77.8% Alzheimer's Disease: 7.4% Multiple Sclerosis: 9.3% Parkinson's Disease: 5.6%	1.5T: 53.7% 3T: 46.3%

RESULTS

- QyScore® was the **best performing segmentation** approach for whole brain grey matter, whole brain white matter, amygdala, putamen, lateral ventricles and brainstem (QyScore® DSC's: 0.79-0.93; ANTs DSC's: 0.73-0.85; FreeSurfer DSC's: 0.61-0.82; FSL DSC's: 0.75-0.83), $p_{\text{bonf}} < 0.001$; **Figure 1**.
- FSL performed the best for hippocampal segmentation (DSC=0.83), followed by QyScore® (DSC=0.82), and FreeSurfer (DSC=0.67), $p_{\text{bonf}} < 0.05$. There were no significant differences in cerebellum, caudate, globus pallidus, or thalamus segmentation.
- QyScore® and ANTs successfully processed and segmented 100% of subjects without errors using default parameters, while FreeSurfer automatically processed 53.7% and FSL 88.9%.

Table 2: Dice for each of the 11 structures available for QyScore®, FSL, Freesurfer and ANTs.

Method	Whole brain Grey Matter	Whole brain White Matter	Hippocampus	Amygdala	Caudate	Putamen	Globus Pallidus	Thalamus	Brainstem	Cerebellum	Lateral Ventricles
QyScore®	0.85	0.90	0.82	0.79	0.82	0.88	0.81	0.86	0.93	0.94	0.91
FSL	0.76	0.78	0.83	0.75	0.81	0.83	0.77	0.84	0.79	NA	NA
FreeSurfer	0.65	0.72	0.67	0.61	0.79	0.74	0.69	0.85	0.80	0.85	0.81
ANTs	0.73	0.76	NA	NA	NA	NA	NA	NA	0.85	0.87	NA

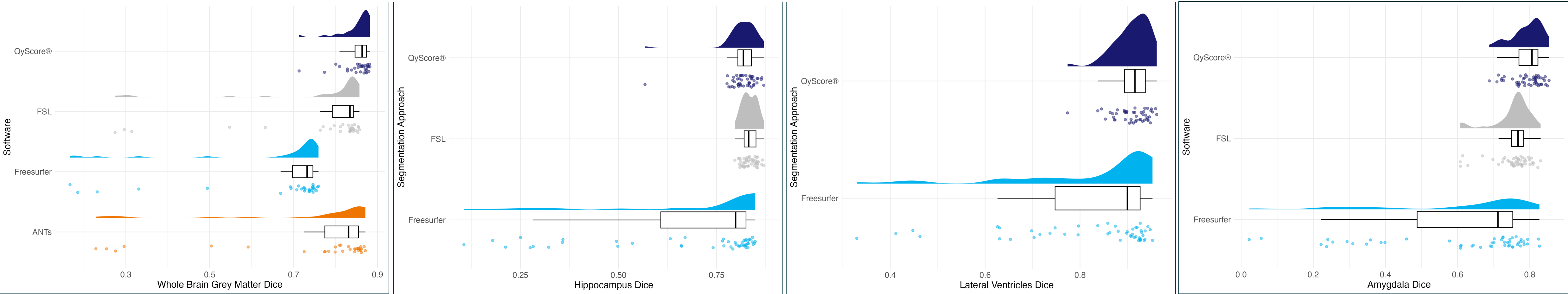


Figure 1. Dice similarity coefficients (DSC) for brain region segmentations across QyScore®, FSL, FreeSurfer, and ANTs. Higher DSC values (closer to 1) indicate greater overlap with the manual ground truth, reflecting higher segmentation accuracy.

CONCLUSION

- QyScore® was simple to implement, fully automated, requiring no expert knowledge to operate, and outperformed other pipelines in segmentation accuracy for most key regions.
- QyScore® enables objective, reproducible brain measurements to improve clinical diagnosis and monitoring of CNS diseases like Alzheimer's disease, Parkinson's disease and multiple sclerosis.
- QyScore® uniquely provides both accuracy and reliability, supporting its implementation into healthcare and clinical trial settings.