

Training for manual hippocampal segmentation based on the EADC-ADNI harmonized protocol

Marina Boccardi, Nicolas Robitaille, Fernando Valdivia, Martina Bocchetta, Corinna Bauer, Melanie Blair, Emma Burton, Enrica Cavado, Adam Christensen, Kristian Frederiksen, Michel Grothe, Mariangela Lanfredi, Yawu Liu, Oliver Martinez, Masami Nishikawa, Marileen Portegies, Margo Pronk, Travis Stoub, Tim Swihart, Chad Ward, Liana Apostolova, Rossana Ganzola, Gregory Preboske, Dominik Wolf, Clifford Jack, Simon Duchesne, Giovanni Frisoni

Background

A Harmonized Protocol (HP) for manual hippocampal segmentation was defined by a Delphi panel (Boccardi et al., *Neurology* 2012;78(S1):S04003) and described in a written document. Benchmark images were segmented by 5 HP-expert tracers as the reference gold standard and uploaded on a web-based platform providing a standard training system.

Methods

Seventeen “Naïve” tracers with ICC>0.80 in hippocampal segmentation based on their local protocols and no previous experience with the HP were recruited from EADC/ADNI (European Alzheimer Disease Consortium/Alzheimer’s Disease Neuroimaging Initiative) centres. They were provided with the same HP criteria and segmentation instructions through the web-platform, and with the same version of MultiTracer for manual segmentation. Training images from 10 ADNI subjects for whom benchmark labels were available were divided into three rounds (n=2, n=4, n=4), balanced by magnetic strength field and degree of hippocampal atrophy. Tracers were asked to segment both hippocampi based on the HP and upload segmented images on the platform. Visual feedback was provided showing point by point discrepancies of segmentations versus the reference in color code. Written feedback was provided slice by slice in addition to the visual feedback for the first two rounds. In subsequent rounds tracers were asked to upload the images of the previous round that had been corrected based on the feedback, and to segment 4 additional images. Dice and Jaccard overlapping indices were computed versus the mean of the HP-experts’ segmentations. A slice by slice visual quality check (QC) was carried out to match overlapping values with levels of compliance with the HP criteria.

Results

The 10 tracers who completed the training so far had very high Dice values (1.5T: median=0.90 range=0.82–0.91; 3T: median=0.91 range=0.89–0.92). Jaccard values (1.5T: median=0.81 range=0.74–0.84; 3T: median=0.83, range=0.80–0.85) better discriminated segmentations based on compliance versus the HP, with values below 0.71 denoting major segmentation mistakes, and values below 0.80 denoting incomplete compliance with the HP criteria.

Conclusions

Preliminary results show that tracers can successfully be trained a remoto to segment the hippocampus based on the HP in three training rounds. Results paired with QC will allow to set criteria for qualification within a standard web-based platform.