Background

Hippocampal atrophy is a key diagnostic marker for early-preclinical Alzheimer’s disease (AD), but manual tracing on magnetic resonance (MR) images (present gold standard procedure) results in heterogeneous volumetric estimates (2 to 5.3 cm³) depending on the adopted segmentation protocol.

Objective

To define a harmonized protocol for the manual segmentation of the hippocampus.

Methods

We selected 12 most used tracing protocols in the AD literature (Figure 2). One rater carried out complete tracings on two prototypical 1.5T MR scans (0.99x0.99 mm³) (one control and one matched AD, ADNI subjects) on 1.2 mm slices, using each protocol. Individual interactive web conferences with the primary author of each protocol allowed to check or correct the execution of the tracing. We extracted the differences among the author-certified protocols, operationalized them into segmentation units (Figure 1) in order to compute their influence on total hippocampal volume, difference due to AD, and reliability measures in the manual tracing. Then, we traced and re-traced the segmentation units on 20 ADNI subjects (4 for each severity degree at the MTA scale - Scheltens et al., 1992) and, for each, we quantified their intra-rater reliability and impact on volume and differences.

Figure 1. 3D rendering of the differences among the 12 protocols, operationalized based on the certified tracings. Red=Minimum hippocampal body, common to all protocols; Yellow=alveus/fimbria; Green= different criteria to trace the medial border at the level of the subiculum; Blue= different criteria to trace the most caudal slice (tail).

Figure 2. 3D rendering of the originally examined protocols, by segmentation units assembly.

Results

The certified protocols (available at www.hippocampal-protocol.net) differed in the definition of the medial border (subiculum, green), of the last slice (tail, blue), and the inclusion of hippocampal white matter (alveus/fimbria, yellow). The impact of these differences, operationalized into segmentation units (SUs) (Figure 1), on total volume, AD difference, and tracing reliability, as computed from the 20 ADNI subjects, is reported in the Table. SUs can also be used to reconstruct the 3D renders of original protocols (Figure 2).

Conclusions

This operationalization, and the quantification of SUs features provide quantitative evidence that will assist an international panel of experts in achieving a consensus for a harmonized protocol for the manual tracing of the hippocampus.