Supplemental material

Supplementary material is available at Brain online.

1 Methods

Age and disease specific cluster solutions based on bootstrapping

In this study, a bootstrapping approach was applied following a two steps procedure. First, we created bootstrap samples at the step of computation of the covariance matrices (correlating seed and target grey matter values) (Fig. 1). In a second step, bootstrapping was applied after having generated the clustering of each bootstrap samples.

In the first case, the bootstrap number was identical to the sample size generating as many bootstrap samples as participants in the sample. As bootstrapping is resampling with replacement, different versions of the original sample were generated. The covariance matrices for each bootstrap sample were computed by correlating the seed and target matrices within each generated bootstrap sample. Afterwards, we applied the clustering algorithm on these covariance matrices resulting in a solution matrix for each dataset-group (e.g. HCP_young, eNKI_old etc). This solution matrix contained, for each bootstrap sample (i.e. for each different version of the original sample), the assignment of each seed voxel to a cluster (i.e. the clustering). We merged the dataset and group specific matrices into one group-specific matrix (e.g. HCP_Young, eNKI_Young and CamCAN_Young => Young) containing for each bootstrap sample, the cluster assignment of each voxel. We again applied bootstrapping (10 000 iterations) on the group specific merged (across datasets) solution matrix to ensure further stability and to eliminate further noise. In other words, we bootstrapped the bootstrap samples containing the clustering (assignment of voxels to clusters). After this step the final matrix contained the final ("stable") cluster assignment of each seed-voxel by taking the mode for each seed-voxel across 10 000 bootstrap samples (Fig. 1).

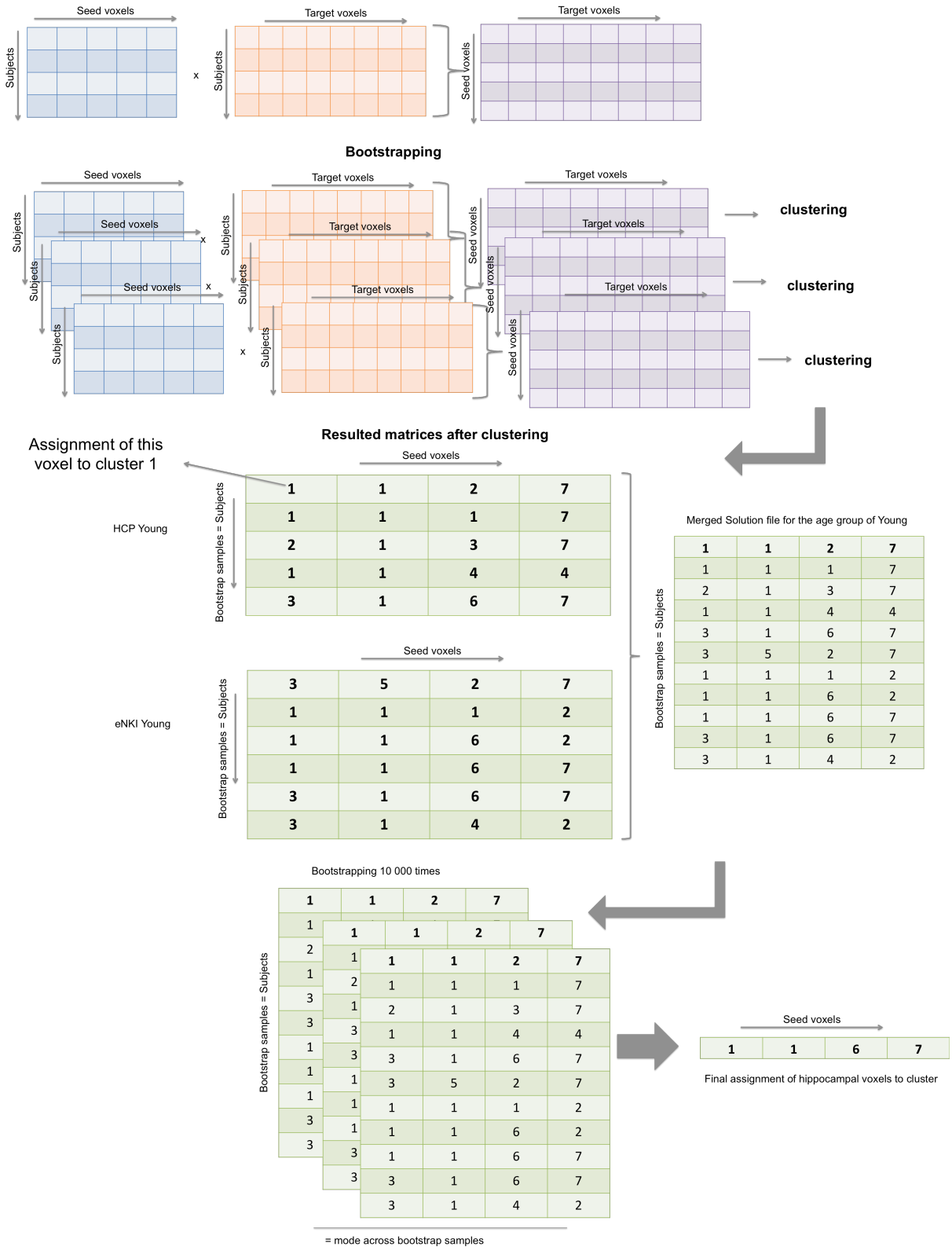


Figure 1. Illustration for the computation of a stable clustering with a bootstrapping approach.

2 Results

2.1 Clusters' stability and consistency

Simple cluster solutions (2 and 3) were more stable than partitions of higher granularity and were dependent on dataset and age/disease group. In several datasets, stability seemed to increase (or to remain stable) from 2 to 3 cluster solutions in both the right (in HCP, 1000BRAINS and eNKI) and left (in eNKI, 1000BRAINS and OASIS3) hippocampi (Fig.1). In contrast, a clear increase in stability from 3 to 4 clusters solution was only observed in the CamCAN dataset (n=94) suggesting that this pattern could be dataset specific.

We additionally examined whether stability of clusters was also dependent on age and disease group and performed a 5 (age/disease group: young, middle age, elderly, MCI, dementia) x 6 (cluster solution: 2-7) ANOVA with the aRI as dependent variable. The ANOVAs were performed separately for each hemisphere.

For the right hippocampus all main and the interaction effects were significant: cluster solution, $F(5,839970)=45388.74$, $P < 0.001$, age/disease group, $F(4,839970) = 24244.72$, $P < 0.001$, cluster solution x age/disease group, $F(20,839970) = 5406.63$, $P < 0.001$. Post-hoc Bonferroni corrected multiple comparison computations revealed that MCI was the age/disease group with the most stable cluster solutions (aRI=0.95), followed by young (aRI=0.94) and elderly (aRI = 0.94), middle aged (aRI=93), and dementia (aRI= 0.90).

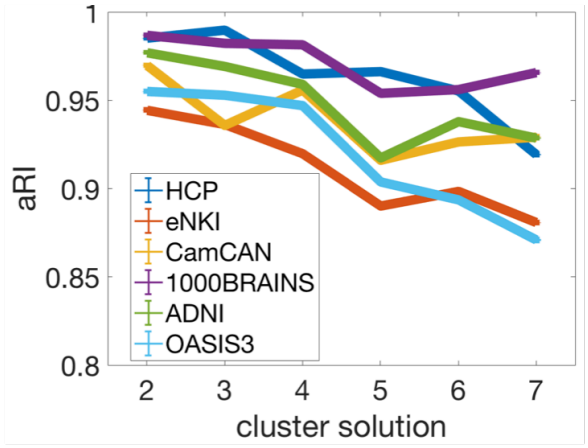
For the left hippocampus all main and the interaction effects were also significant: cluster solution, $F(5,839970) = 20243.24$, $P < 0.001$, age/disease group, $F(4,839970) = 8929.86$, $P < 0.001$, cluster solution x age/disease group, $F(20,839970) = 259.13$, $P < 0.001$. Post-hoc Bonferroni corrected multiple comparison computations revealed that elderly had the most stable partitions (aRI=0.94), followed by MCI (aRI=0.94), young (aRI = 0.93), middle aged (aRI = 0.92) and dementia (aRI = 0.93).

The most stable cluster solutions were 2 (right aRI=0.96, left aRI =0.96), 3 (right aRI=0.95, left aRI=0.95), and 4 (right aRI=0.94, left aRI =0.94) followed by 5 (right aRI=0.91, left aRI =0.92) and 6 (right aRI=0.91, left aRI = 0.92), and 7 (right aRI=0.90, left aRI = 0.93).

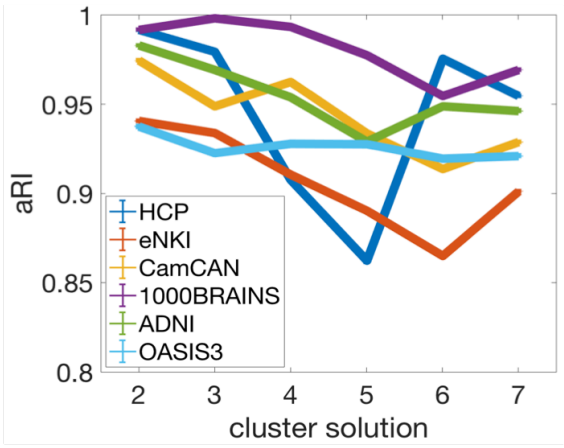
Overall, more basic cluster solutions such as 2, 3 and 4 were preferred compared to higher granularities in different age/disease groups (Fig. 2). Except for the MCI group the stability for all the other groups dropped after cluster solution 4 indicating less consistency in the differentiations.

Cluster stability dependent on dataset

Right hippocampus

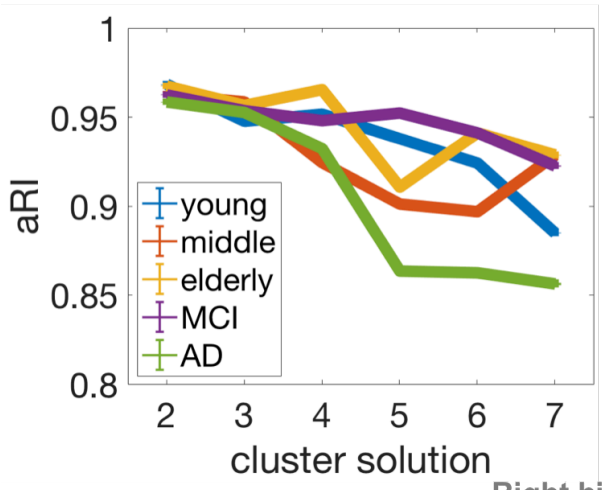


Left hippocampus

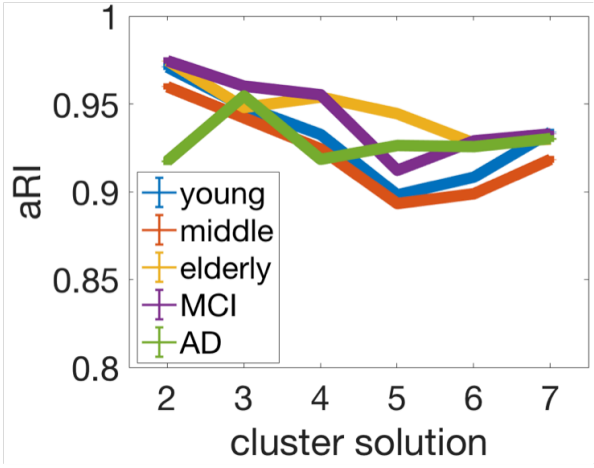


Cluster stability of age and disease groups

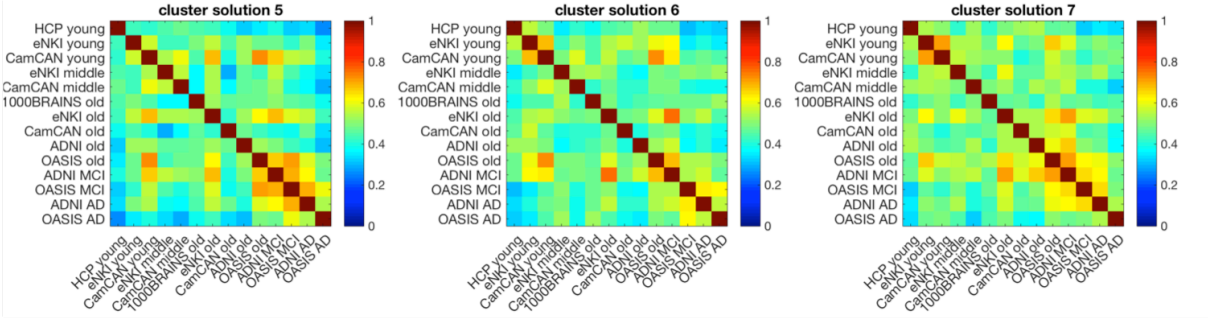
Right hippocampus



Left hippocampus



Right hippocampus



Left hippocampus

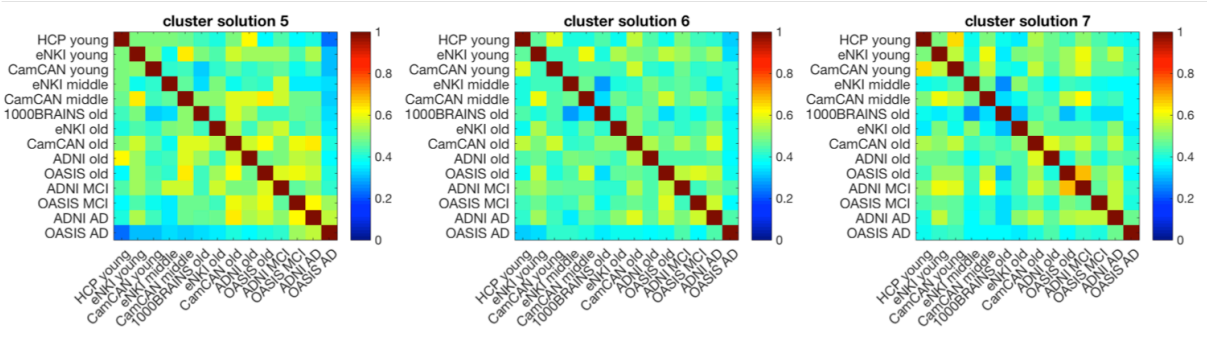


Figure 2. A) Clusters' stability dependent on dataset and age/disease groups. Lower cluster solutions (2-4) demonstrated a higher stability compared to higher cluster solutions, even though the stability was overall high > .85. B) Clusters' consistency across dataset and age/disease was low for higher cluster solutions (5-7) indicating higher heterogeneity in higher granularities possibly due to dataset intrinsic features.

2.2 Silhouette estimation of clusterings

The choice of the optimal cluster solution was guided by three criteria. To estimate the internal validity, we used stability measures estimated with split-half cross-validation and consistency measure using the mean silhouette values. As an approximation of external validity, we assessed the consistency of parcellations across datasets and age/disease specific groups.

Based on the silhouette value, our results indicated that cluster solution 2 and 3 (across all age and disease groups) provide the best data representation of voxels' differentiation in the hippocampus (Fig. 3). This result was in accordance with our previous results based on split-half cross validation and based on consistency across groups and datasets.

The silhouette value was defined in terms of both similarity and distance metric comparing cluster compactness to cluster separation. Silhouette values can range between -1 and +1. Higher positive values indicate a better fit of each individual voxel to the cluster it was assigned to. Negative values indicate a poor fit in the assignment.

We tested with an ANOVA whether silhouette values were significantly different between groups (young, middle-aged, elderly, MCI and dementia) and cluster solutions (k=2:7), revealing significant main effects of group, [$F(4,25920) = 30.26$, $P < 0.0001$ for right hippocampus; $F(4,24900)=25.73$, $P < 0.001$ for the left hippocampus], cluster solutions, [$F(5,25920)=693.6$, $P < 0.0001$ for right hippocampus; $F(5,24900)=586.27$, $P < 0.001$ for left hippocampus] and the interaction effect, [$F(20,25920) = 17.48$, $P < 0.0001$ for right hippocampus; $F(20,24900)=8.09$, $P < 0.001$ for left hippocampus].

More simplistic differentiations into 2, 3 and 4 clusters had significantly higher silhouette values compared to subdivision patterns of higher granularity for both, right and left hippocampus ($P < 0.001$, Bonferroni corrected comparisons revealed ($P < 0.001$). But there was no significant difference between cluster solution 5 and 6 ($P =0.32$ for right hippocampus and $P = 0.62$ for left hippocampus).

Post-hoc Bonferroni corrected multiple comparisons showed that parcellations in young and middle-aged participants had lower silhouette values compared to all the other groups (young vs all the other groups $P < 0.001$ for right hippocampus; young and middle aged ($P = 0.77$) compared to all the other groups $P < 0.001$ for the left hippocampus). There was no significant difference in silhouette values between the group of elderly and dementia patients ($P = 0.39$ for left hippocampus).

In sum, the silhouette metric supported cluster solution 2 and 3 as optimal subdivisions for the hippocampus across age and disease groups.

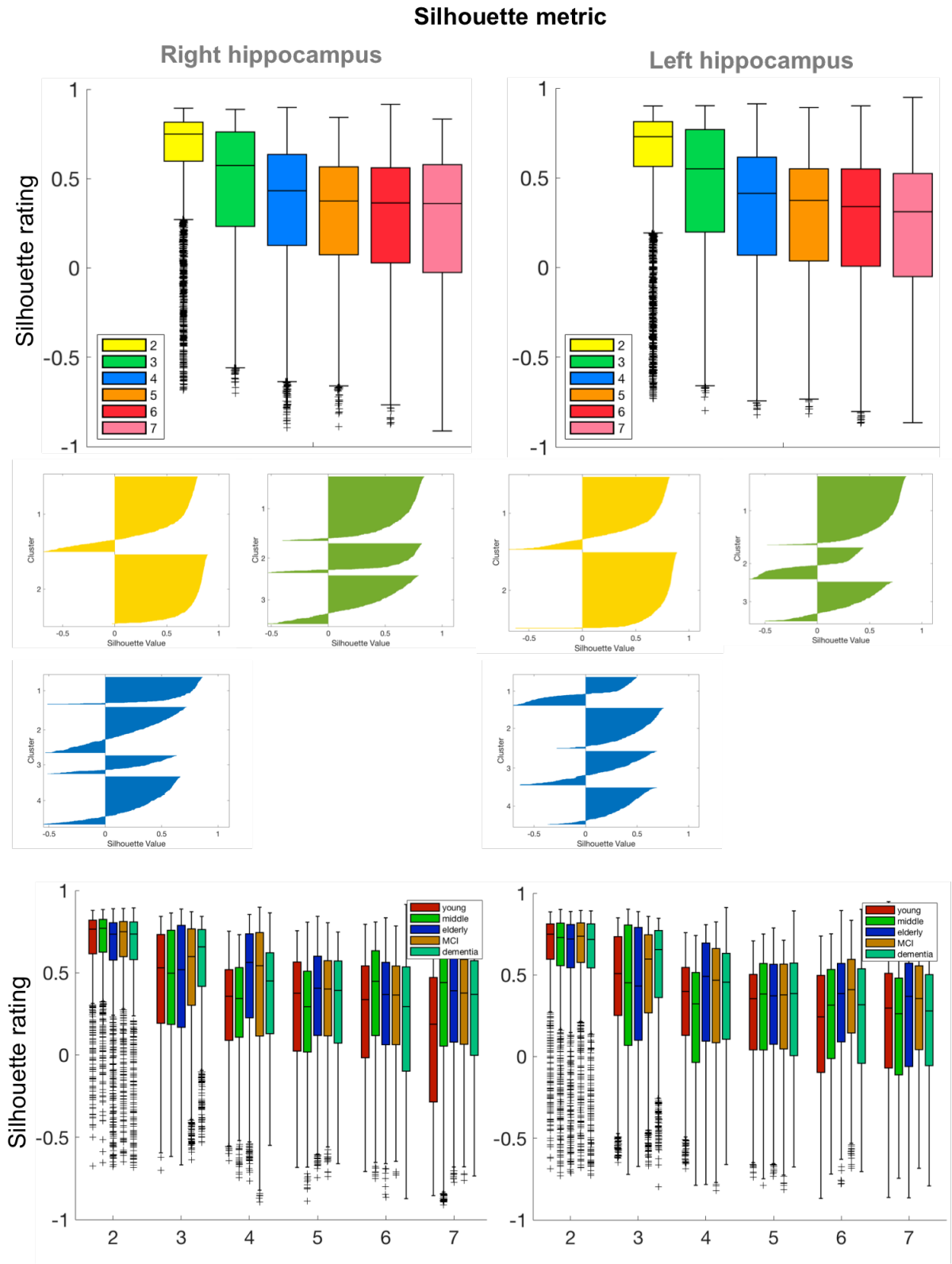


Figure 3. Silhouette measure for right and left hippocampus across age and disease groups and cluster solutions. Boxplots show the median, 1.5 interquartile range, min. Q1-1.5*IQR, max. Q3+1.5*IQR.

2.3 Dataset specific similarity of cluster solution three

Cluster solution 3 showed very high similarity across datasets and age/disease groups dividing the hippocampus in an anterior head region, lateral and medial subregions. Independent of dataset, in the group of young adulthood the tail was covered by the lateral (green) parcel, whereas with increasing age the lateral subregion decreased gradually from the tail. The parcellation in the CamCAN young sample did not follow this differentiation probably due to either a high variability within a small sample size (n=94) or dataset specific intrinsic features.

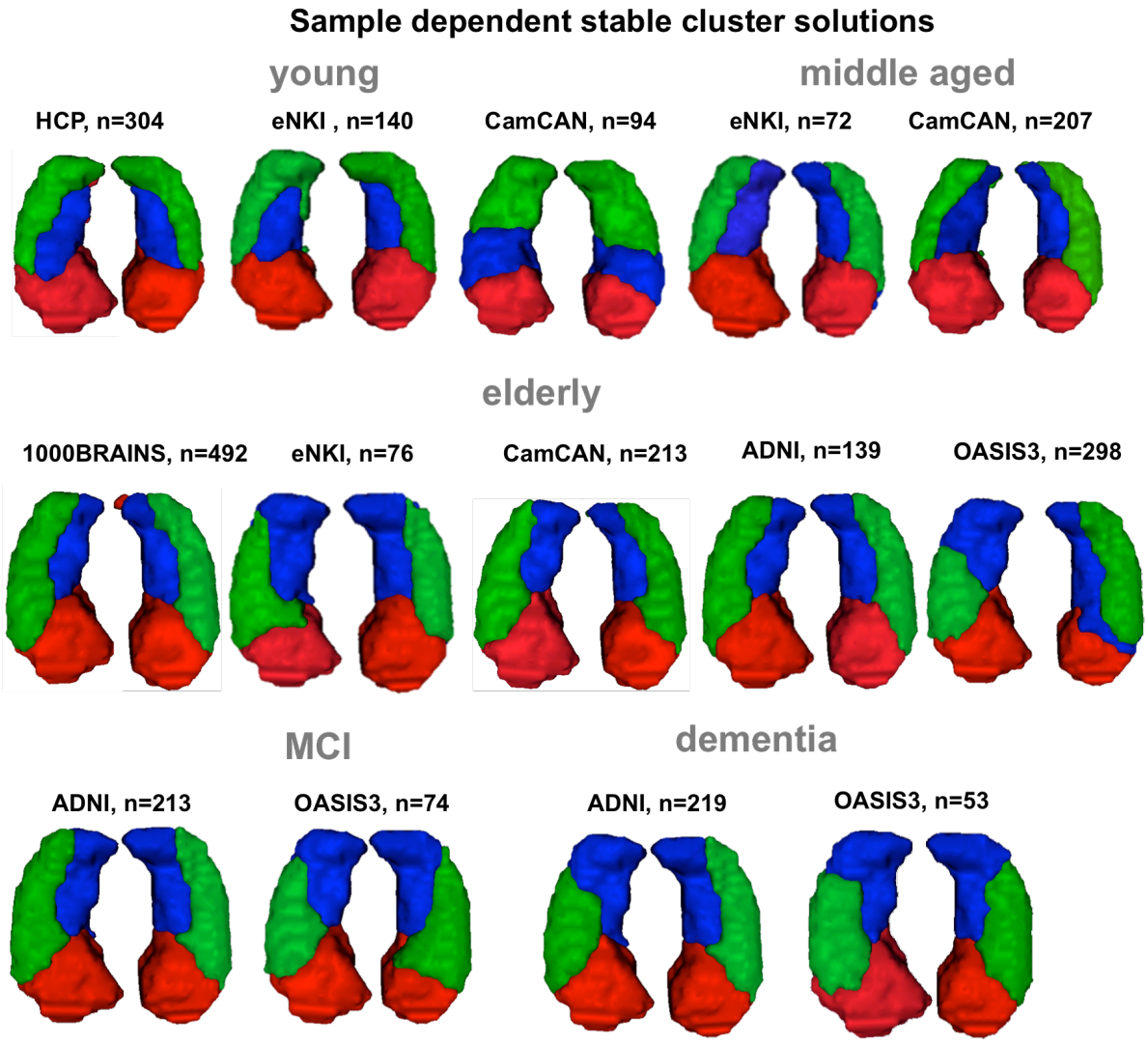


Figure 4. Sample and age/disease specific 3-cluster solution differentiation. Overall hippocampal parcellation showed a high similarity for each age/disease group even though derived from different datasets. In cases where high variability is to expect (young, middle

aged and MCI patients) clusterings diverged slightly coming from different datasets. CamCAN_young and OASIS_old slightly dropped out from the overall phenotype parcellation pattern.

2.4 Age and disease specific differentiation of cluster solution two and three

Based on stability and consistency measures we chose cluster solution 3 to study lifespan and disease related alterations, as this differentiation seemed to be more neurobiologically informative than others. First, the differentiation into 3 clusters was stable enough to ensure that we measured the same biological feature. Secondly, despite stability, it also captured age and disease dependent divergences that better reflected co-plasticity and co-atrophy than cluster solution 2. Indeed the differentiation into 2 clusters was more stable than cluster solution 3 but it was less suitable to study alterations, as this differentiation mirrored a very stable simple partition into one anterior and one posterior subregion independent of age and disease condition (Fig. 5). On the other hand, cluster solution 4 was less stable and more diverse in its qualitatively unique differentiation pattern across age/disease groups, which showed less convergence between groups (Fig. 5), and therefore challenging to study related features of aging and dementia. In the group of young and middle aged healthy adults, the subdivision into 4 parcels resulted in an additional cluster in the head hippocampus, whereas in healthy elderly and in MCI patients the posterior lateral subregion was subdivided additionally in the tail. In dementia, however, the fourth subregion emerged in the medial head-body region, illustrating high divergence between age/disease groups.

Overall, as already summarized in our analysis the composition of stability and consistency of differentiations driven by age/disease specific intrinsic characteristics are better represented in cluster solution 3 compared to 2 and 4.

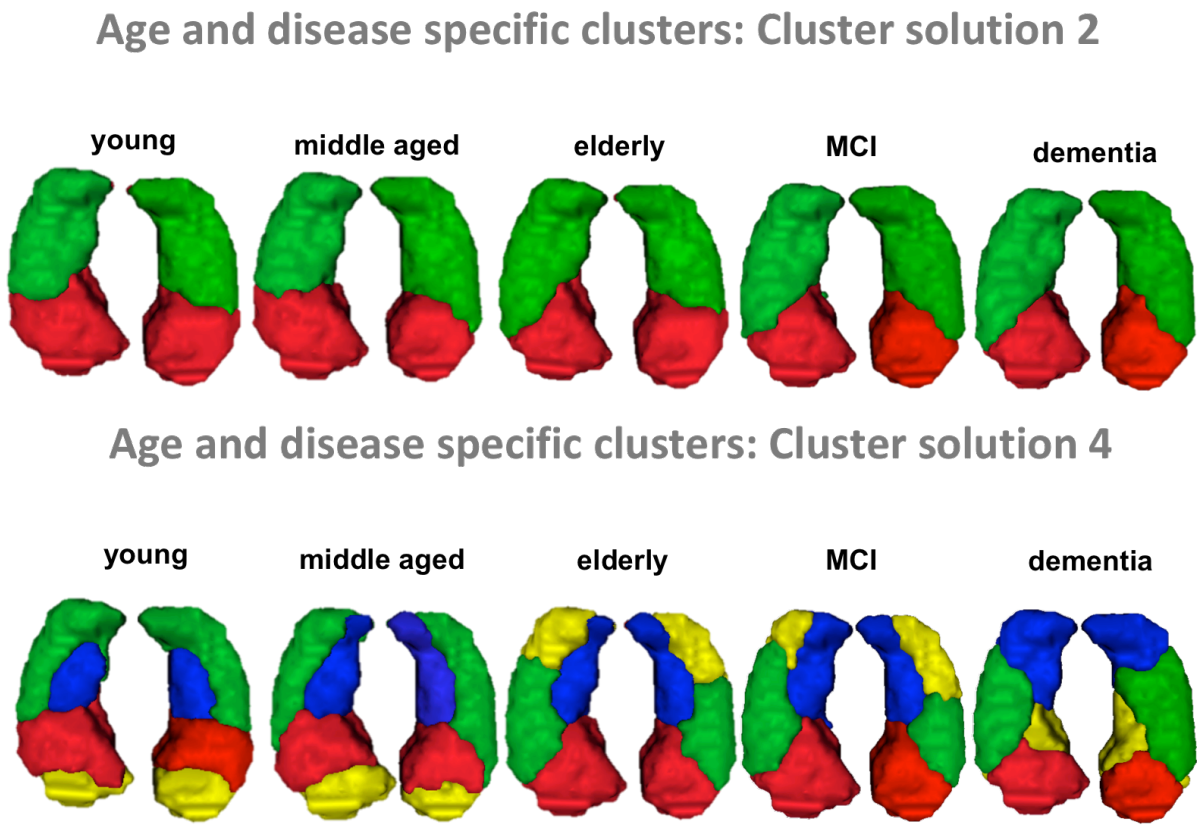


Figure 5. Age/disease specific and stable differentiations of the hippocampus into 2 and 4 parcels.

2.5 Hippocampal structural covariance networks

Uncorrected and corrected structural co-variance networks across age/disease groups.

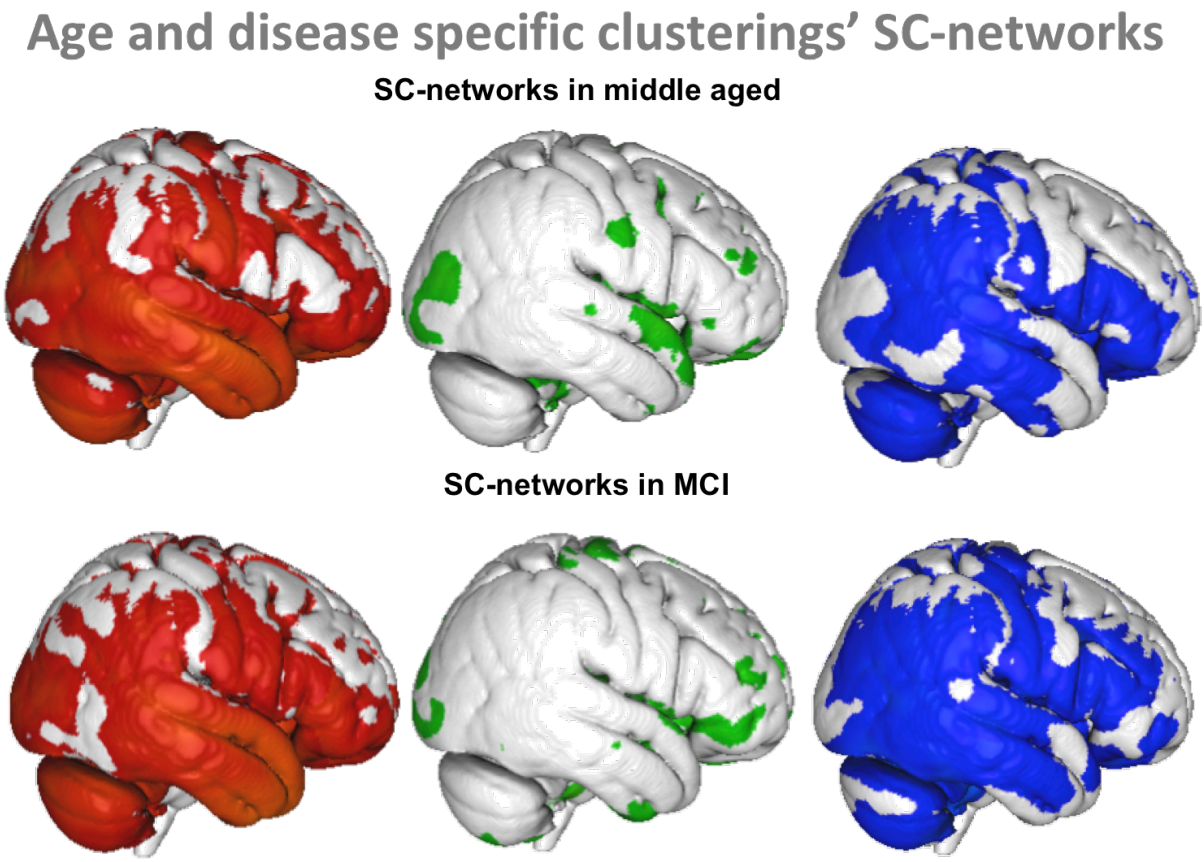


Figure 6. Uncorrected ($P < 0.001$, T=1) structural covariance networks of hippocampal clusters in middle age and MCI age/disease group.

Age and disease specific clusterings' SC-networks

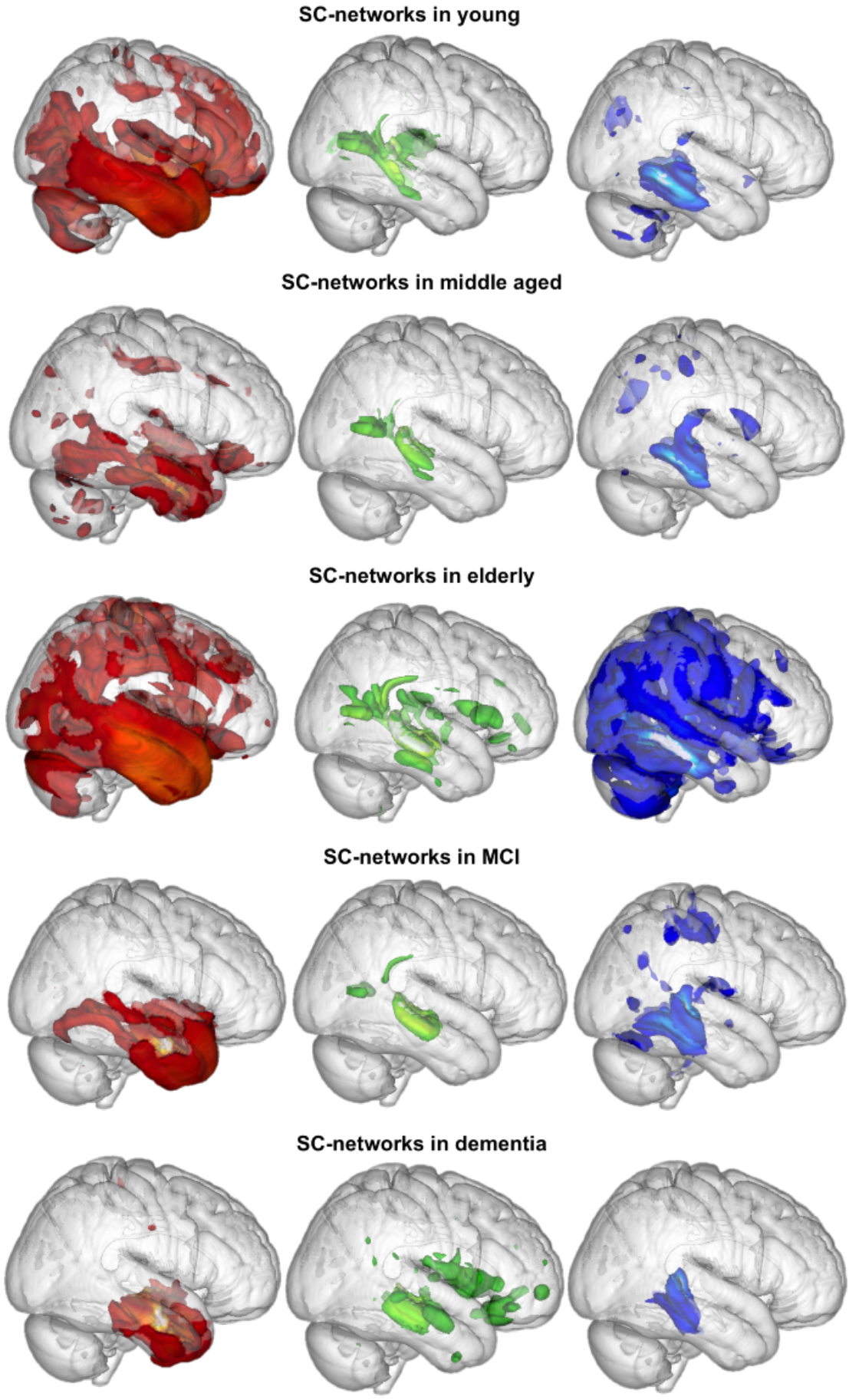


Figure 7. Corrected (FWE, $P < 0.05$, T=4.46) structural covariance networks of hippocampal clusters dependent on age/disease groups.

2.6 Harmonized hippocampal structural covariance networks

In order to account for data coming from different sites the grey matter values that were used for the general linear model in SPM to obtain underlying structural covariance networks of hippocampal subregions, were harmonized. To reduce site related noise, we harmonized the grey matter values (https://github.com/Jfortin1/ComBatHarmonization) (Fortin *et al.*, 2018) across sites n=71, before performing general linear model computations. The primary function of harmonization is to reduce unwanted, non-biological sources of variance related to MRI scanners and sites such as field strength, manufacturer and divergent scanning protocols (Fortin *et al.*, 2018). Harmonized hippocampal structural covariance networks are represented in Fig. 8 and Fig. 9 showing similar patterns compared to non-harmonized data.

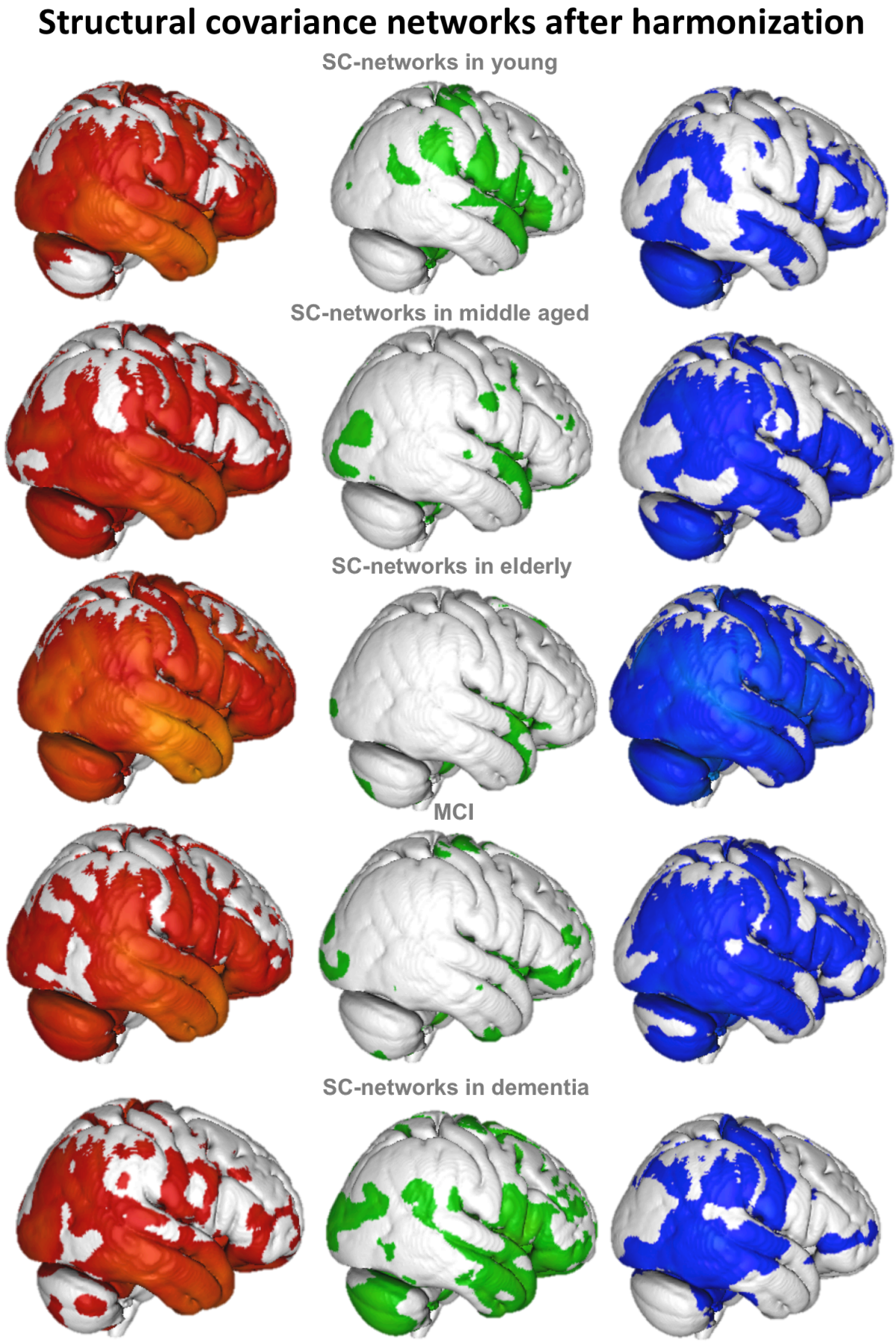
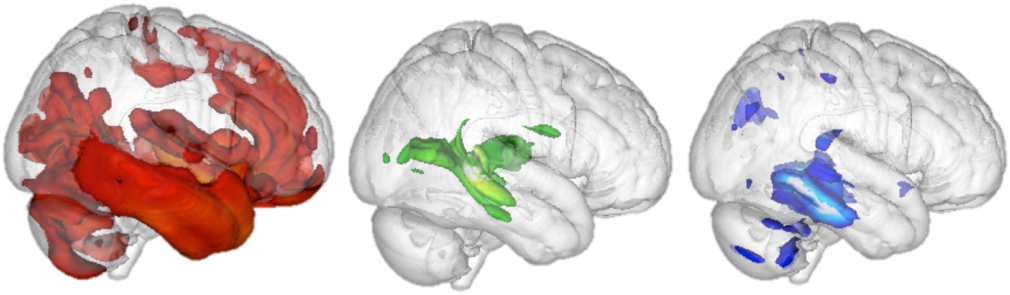


Figure 8. Uncorrected ($P < 0.001$, T=1) hippocampus associated structural covariance

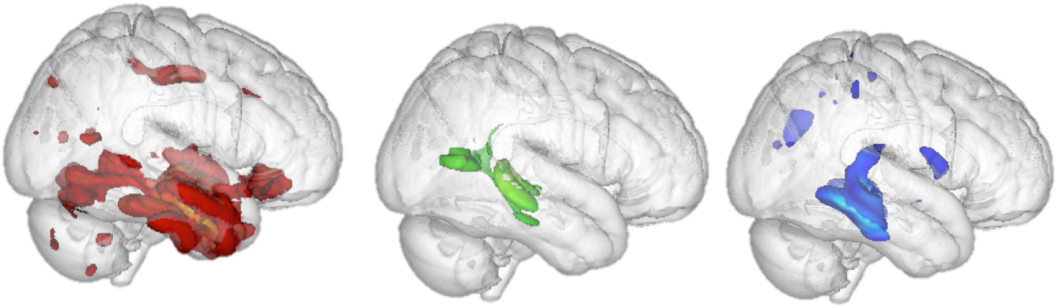
networks derived after harmonization of grey matter values.

Age and disease specific clusterings' SC-networks after harmonization

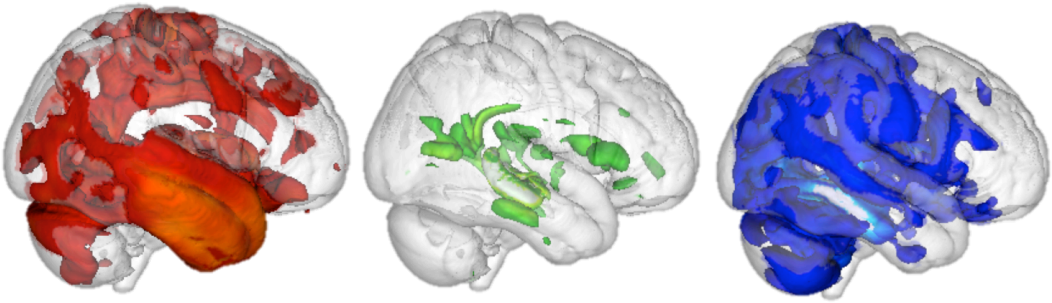
SC-networks in young



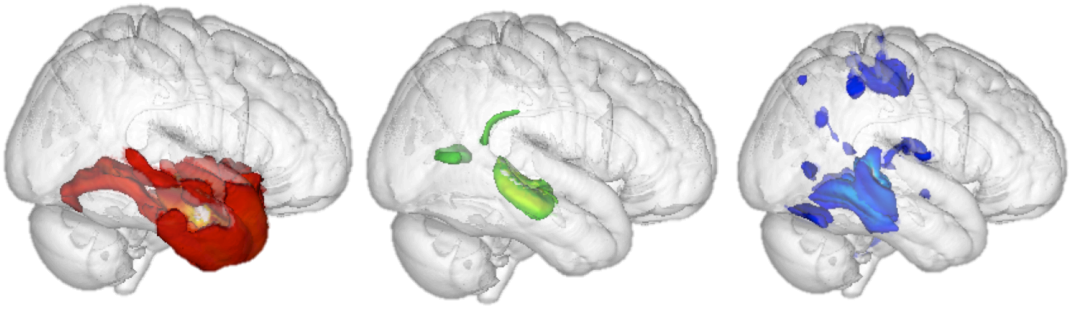
SC-networks in middle aged



SC-networks in elderly



SC-networks in MCI



SC-networks in dementia

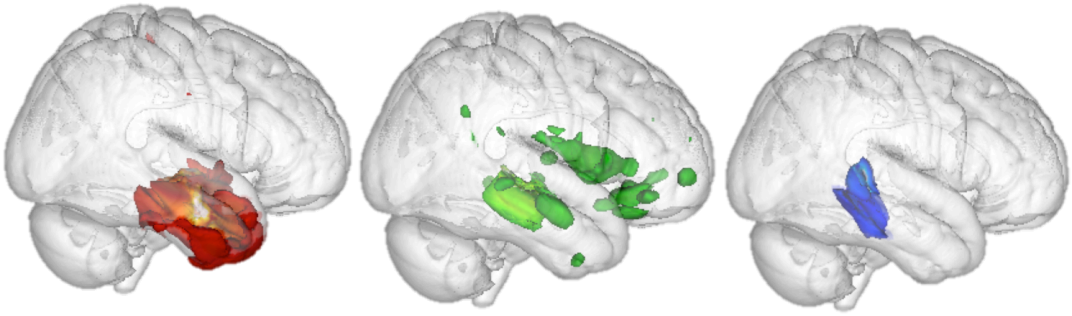


Figure 9. Corrected (FWE, $P < 0.05$, T=4.46) structural covariance networks of hippocampal clusters derived after harmonization of grey matter values.

2.7 Behavioral profiling of clusters' structural covariance networks

In the group of middle-aged participants and MCI patients the behavioral profile of the anterior and medial cluster did not change compared to other healthy age groups. The anterior cluster was involved in the perceptual-emotional-regulatory processing of information into self-relevant internal memories. The medial (blue) cluster was associated with motor exploration and orientation behavior (Fig. 10), but in MCI patients the medial (blue) cluster was additionally related to behavioral terms such as recognition, recollection and retrieval.
In both groups of middle-aged and MCI patients, the behavioral association of the lateral (green) cluster was less distinctive compared to the anterior and medial clusters. In MCI patients it was related to autobiographical memory, episodic memory and retrieval, all terms also related to either the medial or anterior parcel.

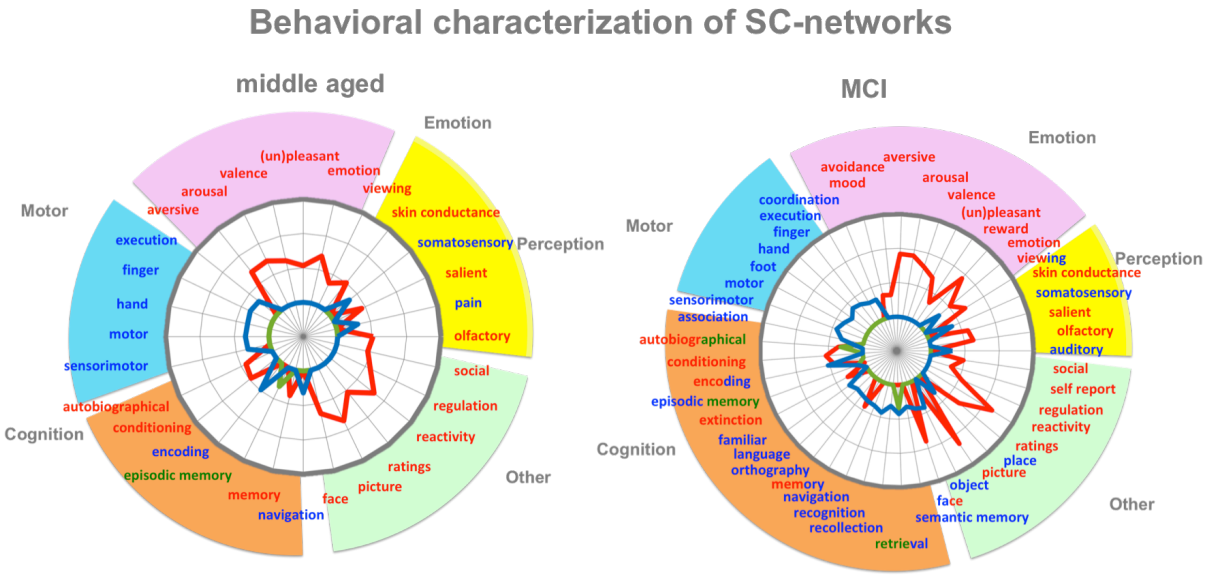


Figure 10. Behavioral characterization of structural covariance networks in the group of middle aged healthy adults and MCI patients.

3 Discussion

Structural covariance pattern in the hippocampus in MCI resemble healthy adults' pattern

In addition to our observations that the pattern of hippocampus differentiation based on structural covariance remained similar across age groups, we also found that this pattern was replicated in patients with mild cognitive impairment, despite an ongoing decrease in the tail of the lateral (green) subregion. Accordingly, the hippocampus' differentiation pattern in MCI represents a transition model between normal covariance in healthy aging and co-atrophy caused by pathology of dementia. One reason for a higher similarity with healthy elderly in this study might be the criteria of selection of the MCI patients. We here selected only patients with a strict *very mild* cognitive impairment (e.g. ADNI sample) by excluding patients with more pronounced memory deficits associated with Alzheimer's disease. In other words, we have excluded patients who were likely to be patients with Alzheimer-type pathology at the early stage of the disease (*late MCI*). We therefore hypothesize that some participants were patients at a so early stage of Alzheimer's disease that pathology hasn't affected brain structure in a way that would result in qualitatively different disease-related structural covariance patterns.

Asymmetrical differentiation pattern of the hippocampus in dementia

We found asymmetric differentiation patterns in our study for the right and left hippocampus, which were especially evident in the MCI and dementia group with the left hippocampus seemingly being more affected by disease. In MCI, this could be inferred by a higher decrease of the lateral-green cluster from the tail and, in dementia, by a higher extension of the lateral-body cluster into the medial direction. Higher left hippocampal susceptibility was already reported several times in the context of volume reductions in dementia and MCI (Lindberg et al., 2012; Müller et al., 2005; Shi, Liu, Zhou, Yu, & Jiang, 2009). It has been hypothesized that aging and disease affect more likely the left hemisphere than the right hemisphere. Nevertheless, a meta-analysis by Minkova et al. (2017) suggested a lack of support for this hypothesis, despite a tendency for the right hippocampus to be more atrophied in MCI and the left hippocampus being more affected in AD. Global lateralized atrophy as assessed by Minkova et al. (2017) might appear late in pathology or with increase in disease severity. In contrast, lateralized differentiation patterns as investigated in the present study, seem to be

more evident in the hippocampus in dementia. Future studies could reveal under which circumstances lateralized differentiation patterns and lateralized atrophy arise.

References

Fortin J-P, Cullen N, Sheline YI, Taylor WD, Aselcioglu I, Cook PA*, et al.* Harmonization of cortical thickness measurements across scanners and sites. NeuroImage 2018; 167: 104-20.
Lindberg O, Walterfang M, Looi JC, Malykhin N, Ostberg P, Zandbelt B*, et al.* Hippocampal shape analysis in Alzheimer's disease and frontotemporal lobar degeneration subtypes. Journal of Alzheimer's disease : JAD 2012; 30(2): 355-65.
Minkova L, Habich A, Peter J, Kaller CP, Eickhoff SB, Kloppel S. Gray matter asymmetries in aging and neurodegeneration: A review and meta-analysis. Human brain mapping 2017; 38(12): 5890-904.
Müller MJ, Greverus D, Dellani PR, Weibrich C, Wille PR, Scheurich A*, et al.* Functional implications of hippocampal volume and diffusivity in mild cognitive impairment. NeuroImage 2005; 28(4): 1033-42.
Shi F, Liu B, Zhou Y, Yu C, Jiang T. Hippocampal volume and asymmetry in mild cognitive impairment and Alzheimer's disease: Meta-analyses of MRI studies. Hippocampus 2009; 19(11): 1055-64.