



Δράση «Εμβληματικές δράσεις σε διαθεματικές επιστημονικές περιοχές με ειδικό ενδιαφέρον για την σύνδεση με τον παραγωγικό ιστό» ID 16618

Εθνικό δίκτυο έρευνας για την ανάδειξη της γενετικής βάσης των νευροεκφυλιστικών νόσων Alzheimer και Parkinson, την ανίχνευση αξιόπιστων βιοδεικτών και την ανάπτυξη καινοτόμων υπολογιστικών τεχνολογιών και θεραπευτικών στρατηγικών στη βάση της ιατρικής ακριβείας (BRAIN PRECISION, TAEDR-0535850)

**ΤΙΤΛΟΣ ΠΑΡΑΔΟΤΕΟΥ:** Ανάπτυξη πολυτροπικών μοντέλων ιατρικής ακριβείας για ταξινόμηση και έγκαιρη διάγνωση νευροεκφυλιστικών

**ΕΝΟΤΗΤΑ ΕΡΓΑΣΙΑΣ 5:** Ανάπτυξη νέων υπολογιστικών μοντέλων και τεχνολογιών για την έγκαιρη διάγνωση των νευροεκφυλιστικών νόσων Alzheimer και Parkinson και των πρόδρομων μορφών τους.

**ΥΠΕΥΘΥΝΗ ΕΡΕΥΝΗΤΙΚΗ ΟΜΑΔΑ (ΦΟΡΕΑΣ):** ΚΩΝΣΤΑΝΤΙΝΟΣ ΜΑΡΙΑΣ (ΙΤΕ)

## Ανάπτυξη πολυτροπικών μοντέλων ιατρικής ακριβείας για ταξινόμηση και έγκαιρη διάγνωση νευροεκφυλιστικών

Είδος Παραδοτέου: Λογισμικό

### ΠΕΡΙΓΡΑΦΗ

Συσχέτιση μοριακών και απεικονιστικών δεικτών σε προκλινικές και κλινικές μελέτες με στόχο την ανάδειξη σύνθετων προβλεπτικών υπογραφών βασισμένες σε γενετικά/μοριακά κλινικά και απεικονιστικά δεδομένα. Για τον σκοπό αυτό έχουμε προχωρήσει στην δημιουργία πολλαπλών μοντέλων μηχανικής και βαθιάς μάθησης με απώτερο στόχο την ακριβή και εξατομικευμένη λήψη αποφάσεων από την σύντηξη δεδομένων τα οποία θα προέρχονται από την ραδιομική ανάλυση την απεικονιστικής πληροφορίας για την ψηφιακή αποκωδικοποίηση των εικόνων σε ποσοτικά χαρακτηριστικά, την δημιουργία προσωποποιημένων ραδιομικών υπογραφών, την εξαγωγή χαρακτηριστικών βαθιάς μάθησης, μετά-πληροφορία από γενετικά/ μοριακά, και κλινικά δεδομένα.

Τα μοντέλα θα επιστρέφουν πιθανολογικές εκτιμήσεις πρόβλεψης οι οποίες θα αποτελέσουν τα δεδομένα εισόδου στο τελικό μοντέλο μηχανικής μάθησης για την ταξινόμηση και έγκαιρη διάγνωση των νευροεκφυλιστικών νόσων. Έχουν εξαχθεί τα πρώτα αποτελέσματα από την παραπάνω διαδικασία και είμαστε στο στάδιο αξιολόγησής τους. Σε τελικό στάδιο, το μοντέλο αυτό θα βασίζεται σε ένα σύνθετο νομόγραμμα το οποίο θα είναι υλοποιημένο σε υπολογιστικό περιβάλλον φιλικό προς το χρήστη και θα αποδίδει την τελική πρόβλεψη με τεχνικές σύντηξης πληροφορίας, επιτρέποντας την εξατομίκευση της πρόβλεψης με αυτοματοποιημένο τρόπο.

Παράλληλα, πραγματοποιήθηκε εκτεταμένη αναζήτηση σε διαδικτυακές βάσεις δεδομένων (π.χ ADNI -Alzheimer Disease Neuroimaging Initiative) και προχωράμε με την εναρμόνιση (harmonization) και επιμέλεια (curation) των δεδομένων που βρέθηκαν με σκοπό την λήψη, δημιουργία, οργάνωση και διατήρηση ενός συνόλου δεδομένων που θα μπορούσαν να χρησιμοποιηθούν, ως επιπρόσθετα δεδομένα εισόδου, με σκοπό την ενίσχυση της ακρίβειας και ευαισθησίας του τελικού μοντέλου μηχανικής μάθησης στην ταξινόμηση και έγκαιρη διάγνωση των νευροεκφυλιστικών νόσων.

### ΜΕΘΟΔΟΛΟΓΙΑ

#### Data preprocessing and denoising

Firstly, the first three volumes were discarded to allow for magnetization effects to stabilize. Slice-timing correction (corrected for Siemens-interleaved slice acquisition), re-alignment, and co-registration/normalization to standard MNI space were subsequently performed. Lastly, spatial smoothing with a 6 mm FWHM (Full Width at Half Maximum) gaussian kernel was applied to improve SNR. These steps were carried out in SPM12 ([www.fil.ion.ucl.ac.uk/spm-statistical-parametric-mapping/](http://www.fil.ion.ucl.ac.uk/spm-statistical-parametric-mapping/)) implemented in MATLAB version 9.8 (R2020a). Next, mean white matter and cerebrospinal fluid (CSF) signals (first five principal components and their first order derivatives) were regressed out of the voxel



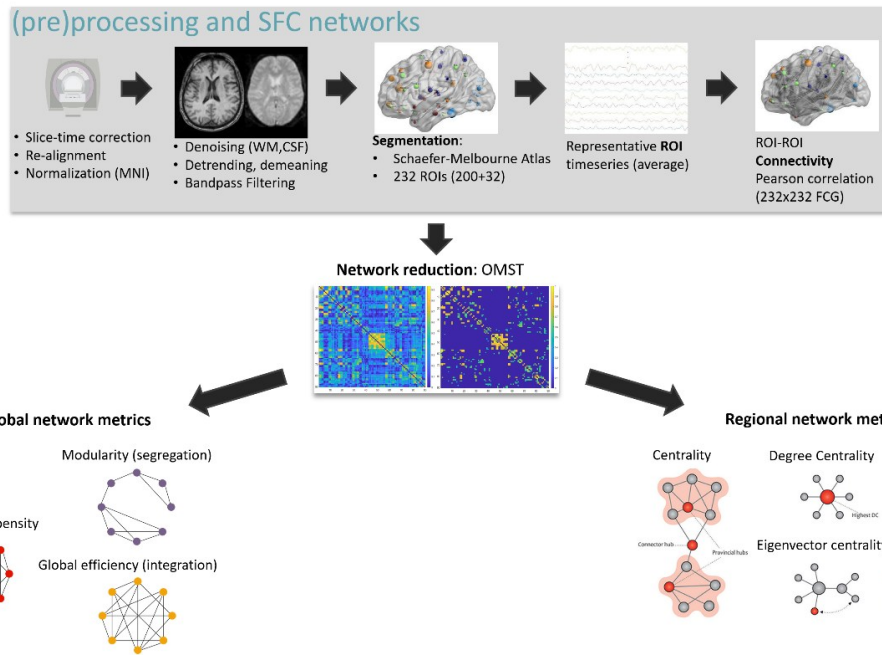
time series, using CompCor [1] included in CONN [2]. Then, voxel timeseries were detrended and bandpass filtered to 0.008–0.09 Hz to eliminate low frequency drift and high frequency noise.

### Functional connectivity estimation

#### *Static functional connectivity*

Static Functional Connectivity (SFC) networks were computed and analysed similar to previous attempts by our team [3] and according to the pipeline that has shown the most reliable, reproducible and representative networks [4], [5], [6]. In short, after preprocessing, 4D BOLD data are individually parcellated into ROI timeseries using the Schaefer(functional)/Melbourne(subcortical) atlas, consisting of 232 regions. Next, SFC graphs (SFCGs) are computed as the Pearson correlations between all pairs of ROI timeseries, resulting in a single 232x232 matrix for each subject. As fully connected networks are suboptimal, network reduction techniques are often employed to reveal underlying structures. OMST [7] was proposed for network reduction by maximizing efficient information flow. We applied OMST on fully weighted networks in order to maintain only the most meaningful connections. OMST has also been shown to produce the most robust SFC networks [5]. Next, global and nodal network metrics are computed in order to quantify various aspects of the functional connectome. This feature extraction step enables the comparison of networks exhibiting varying density and topology. Nodal (region-specific) graph measures reflect increased communication efficiency among the nodes comprising the immediate local community (local efficiency), and the node's ability to act as a go-between, facilitating the stability of a broader network (degree, betweenness or eigenvector centrality). Conversely, global, network-wide measures, reveal the integrative (global efficiency) or segregation (modularity) tendencies of the entire network, as well as indicate the presence of complex topologies, and the balance between integration-segregation and information capacity [8]. The latter, is based on the notion of small world networks [9] and can be measured by the metric of small world propensity [10] local and global measures are suitable to identify regions displaying either hypo- or hyper-connectivity associated with objective or potentially subjective cognitive and emotional disturbances related to MCI. The SFC pipeline including all steps starting from preprocessing can be seen in the following diagram.

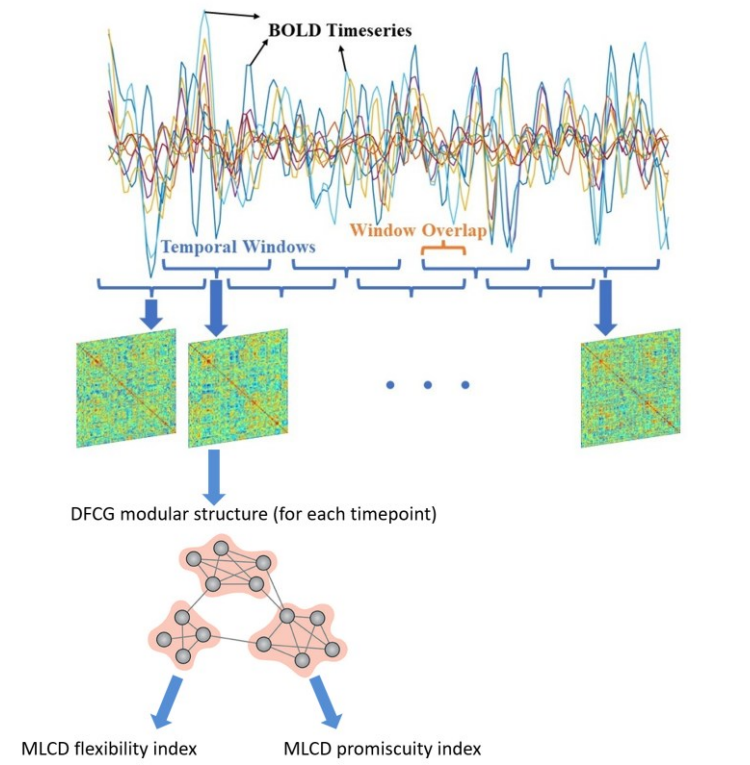




### *Dynamic functional connectivity*

In order to quantify the temporally evolving characteristics of the functional connectome, Dynamic Functional Connectivity (DFC) was estimated using a tapered maximum overlapping sliding window approach [11]. A window length of 23 TRs (53 s) was used in line with the proposed range of 30–60s [12] overlapping on a single TR. Windows were tapered, by being convolved with a Gaussian kernel of  $\sigma = 3$ TRs. In this manner, 125 windows were created, of 23 TRs length each. Pairwise ROI-ROI FC was estimated within each temporal window using the Pearson correlation coefficient, resulting in 125 time-resolved networks or DFCGs (of size  $232 \times 232$  each). Next, in order to evaluate the regional temporally dependent functional connectivity changes, we employed a dynamic functional connectivity (DFC) approach, known as multilayer community detection (MLCD). Evolution of modular structure measured through MLCD metrics has been shown to associate with relative amount of learning [13], working memory performance, planning/reasoning and amount of sleep [14]. Through MLCD we are able to quantify temporal stability/variability of functional connectivity. Communities or modules are groups of regions that are more strongly connected to each other and less or negatively connected to regions of other modules [15]. In a single instance of a functional network, i.e. an SFC network or a temporally specific DFC network, the process of identifying the modular structure, thus assigning every node/region to a particular module, is typically accomplished using the Louvain modularity optimization approach [15], [16]. The approached utilized here, is an extension of this singular regional assignment of nodes to communities to the dynamic or “multilayer” case. The generalized Louvain algorithm [17] was employed for this purpose and two measures that characterize each region’s connectivity variability in terms of affiliations across time. Flexibility [13] measures how many times a particular region has changed its module affiliation, while promiscuity index [18] measures how many modules a region has participated in throughout the entirety of the acquisition, both divided by total number of possible changes for normalization to the 0-1 range.

## Multilayer community detection (MLCD)



### Imaging and behaviour: Statistical and Machine Learning analyses

SFC regional and global metrics were combined with DFC MLCD regional metrics (1407 features total) and used as input features to a Machine Learning classifier trained to differentiate all possible combinations of cognitive and emotional disturbances. Linear and rbf kernel Support Vector Machine (SVM) and Extra Trees classifiers were evaluated in order to find the best performing model. Feature selection was performed in each repeated stratified 6-Fold cross validation iteration in order to only train models with a more conservative number of features thus alleviating the possibility of overfitting the models. Cross validation iterations were repeated 60 times, as mean and standard deviation of results stabilized over this number. Feature selection was based on feature importance consensus ranking, thus selecting only the top ranking features after fitting a preliminary model before the final training, similar to our previous work [3], [19] and the selection scheme described in [20], [21]. A multiclass classifier was trained to differentiate between all four groups. Nine different binary classification models were trained to distinguish between all available combinations using the same combination of imaging features.

### ΠΕΡΙΓΡΑΦΗ ΑΠΟΤΕΛΕΣΜΑΤΩΝ ΚΑΙ ΑΠΟΔΟΣΗΣ ΜΟΝΤΕΛΟΥ

<u>Comp.</u>	<u>Group 1</u>	<u>Group 2</u>
<u>1</u>	Both cognitive and emotional disturbances (MCI-2)	controls

<u>2</u>	Cognitive only, no emotional disturbances (MCI-1)	controls
<u>3</u>	Cognitive disturbances, regardless of emotional (MCI-1 U MCI-2)	controls
<u>4</u>	Cognitive disturbances, regardless of emotional (MCI-1 U MCI-2)	No cognitive complaints (controls U EMO)
<u>5</u>	Emotional only, no cognitive (EMO)	controls
<u>6</u>	Emotional disturbances regardless of cognitive (EMO U MCI-2)	controls
<u>7</u>	Emotional disturbances regardless of cognitive (EMO U MCI-2)	No emotional complaints (controls U MCI-1)
<u>8</u>	All patients (EMO U MCI-1 U MCI-2)	controls
<u>9</u>	Cognitive only, no emotional disturbances (MCI-1)	Emotional only, no cognitive (EMO)

Στην ενότητα των αποτελεσμάτων παρουσιάζονται το μοντέλο που εμφάνισε ικανοποιητική απόδοση (συγκρίση 3) και αφορά την ενότητα εργασίας που αναφέρθηκε παραπάνω. Οι περισσότερες άλλες συγκρίσεις δεν παρουσίασαν απόδοση σημαντικά υψηλότερη από το επίπεδο της τυχειότητας. Η προτεινόμενη διαδικασία μηχανικής μάθησης χρησιμοποιήθηκε κυρίως ως πλαίσιο επιλογής χαρακτηριστικών, με στόχο τον εντοπισμό των απεικονιστικών δεικτών (στατικής ή δυναμικής λειτουργικής συνδεσιμότητας – SFC ή DFC), καθώς και των εγκεφαλικών περιοχών ή παγκόσμιων χαρακτηριστικών, που ήταν αποτελεσματικά σε κάθε επιμέρους σύγκριση.

Σύγκριση 3: Ασθενείς με γνωστική έκπτωση (ανεξαρτήτως συναισθηματικών συμπτωμάτων· MCI-1 U MCI-2) με υγιείς

*Machine Learning classification:*

Classifier: ETC - Selector: SFM\_ETC

Number of Outer splits: 60

- Mean Accuracy 71 +- 13 %
- Mean Balanced accuracy 64 +- 14 %
- Mean Precision 76 +- 9 %
- Mean Sensitivity 84 +- 14 %
- Mean Specificity 45 +- 24 %
- Mean F1 Score 79 +- 10 %
- Mean ROC AUC 64 +- 14 %

Οι συμμετέχοντες χωρίς γνωστική έκπτωση εμφάνισαν υψηλότερες τιμές ευελιξίας

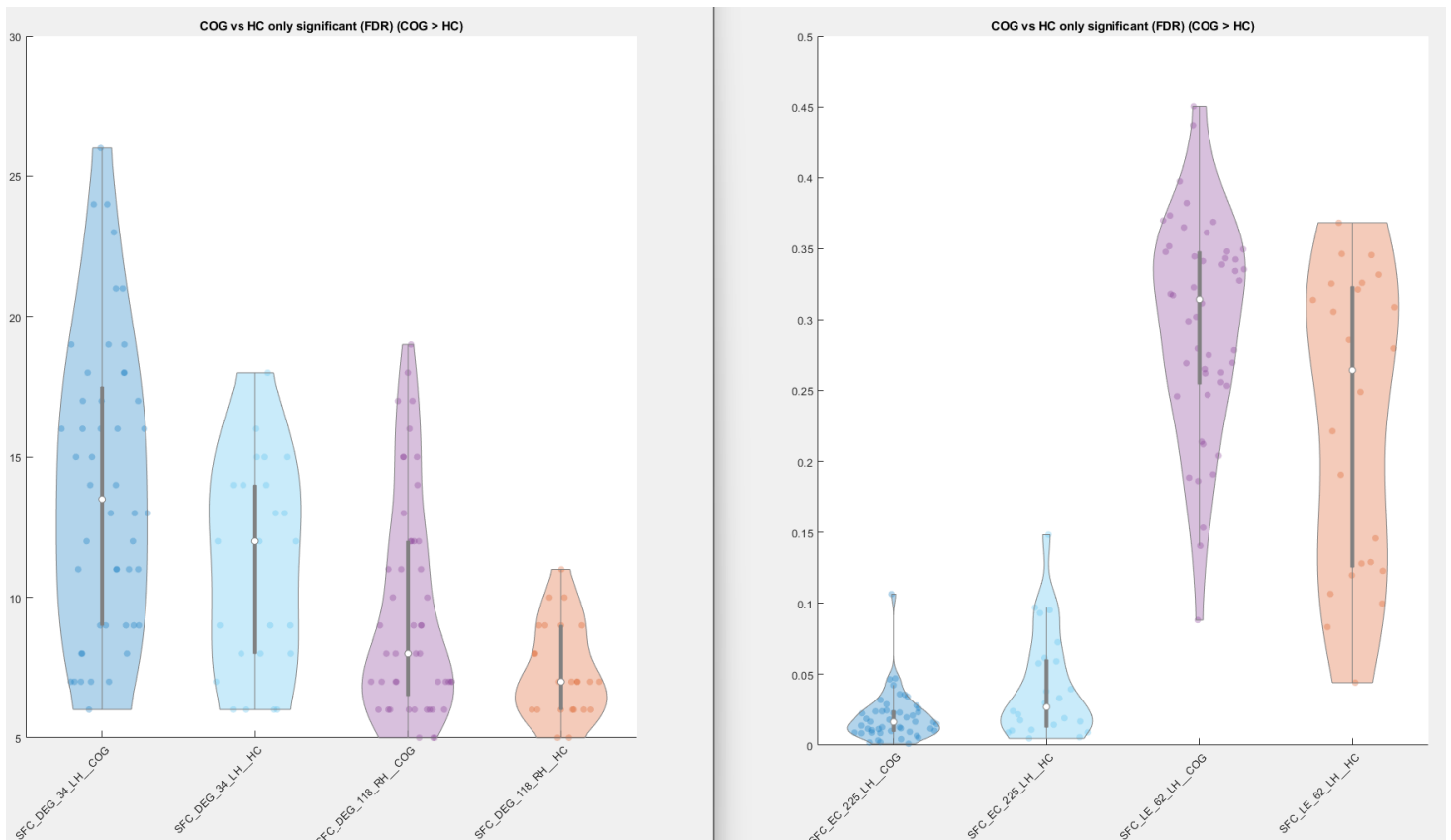
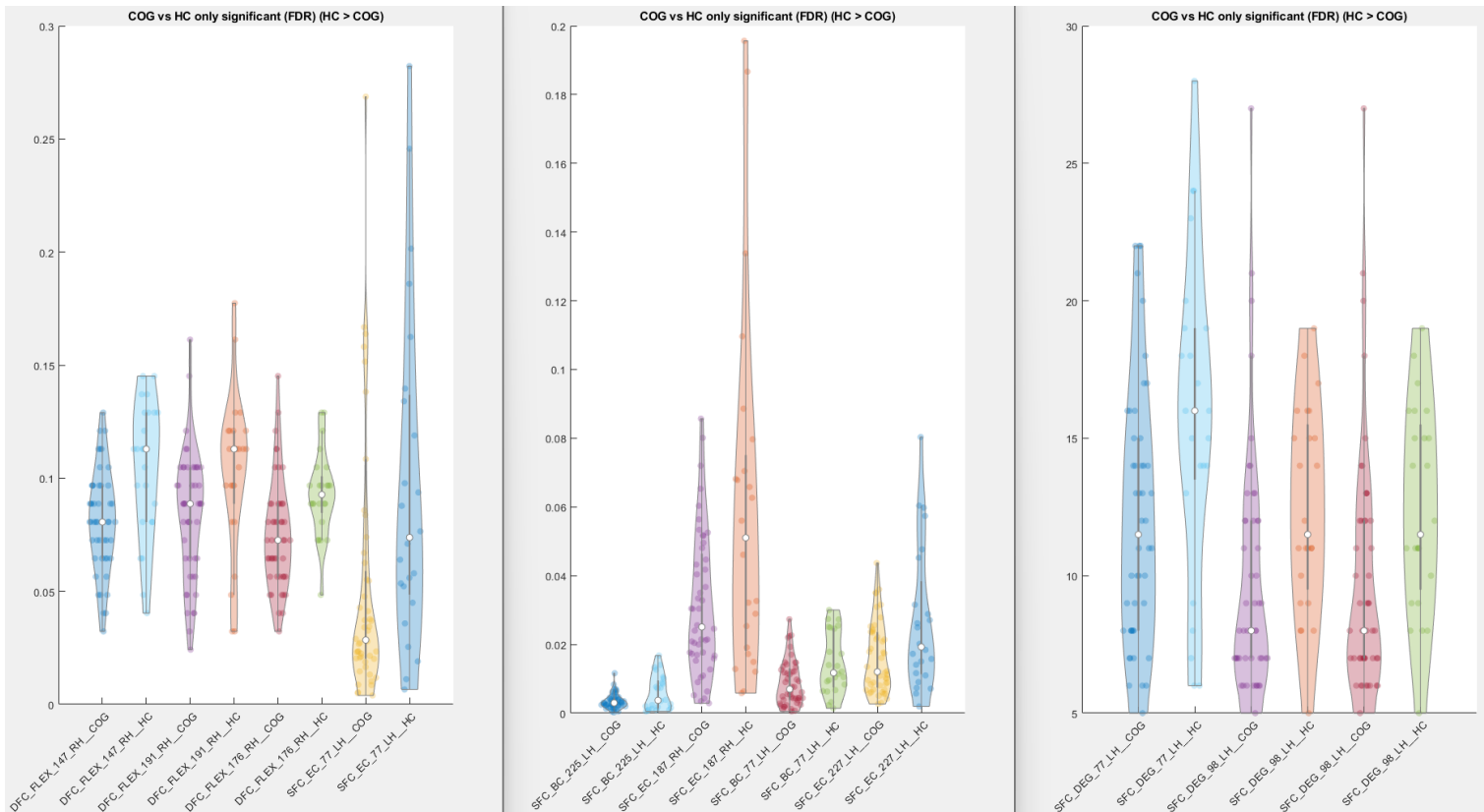
(flexibility), τοπικής αποδοτικότητας (local efficiency), βαθμού (degree), ενδιάμεσης κεντρικότητας (betweenness centrality) και ιδιοδιανυσματικής κεντρικότητας (eigenvector centrality) σε σύγκριση με τους συμμετέχοντες με αντικειμενική γνωστική έκπτωση, στις ακόλουθες εγκεφαλικές περιοχές:

- DFC\_flexibility\_147\_RH\_DorsAttn\_PrCv\_1,52,10,20
- DFC\_flexibility\_176\_RH\_Cont\_PFCI\_7,42,14,48
- DFC\_flexibility\_191\_RH\_Default\_PFCdPFCm\_1,4,36,-14
- SFC\_eigenvector\_centrality\_77\_LH\_Default\_Temp\_4,-58,-30,-4
- SFC\_degree\_centrality\_77\_LH\_Default\_Temp\_4,-58,-30,-4
- SFC\_degree\_centrality\_98\_LH\_Default\_pCunPCC\_3,-4,-30,36
- SFC\_eigenvector\_centrality\_187\_RH\_Default\_Temp\_3,54,-6,-10
- SFC\_betweenness\_centrality\_77\_LH\_Default\_Temp\_4,-58,-30,-4
- SFC\_eigenvector\_centrality\_227\_pGP-lh,-22,-8,-2

Η αντίθετη τάση παρατηρήθηκε για τον βαθμό κεντρικότητας (degree centrality) και την τοπική αποδοτικότητα (local efficiency) στις ακόλουθες εγκεφαλικές περιοχές:

- SFC\_local\_efficiency\_62\_LH\_Cont\_Par\_2,-34,-62,48
- SFC\_degree\_centrality\_118\_RH\_SomMot\_3,38,-14,14

*Στατιστικές συγκρίσεις των χαρακτηριστικών που προέκυψαν από τη μηχανική μάθηση μεταξύ των ομάδων COG και Υγιών (παρουσιάζονται μόνο τα στατιστικά σημαντικά αποτελέσματα,  $p < 0.05$ , διόρθωση FDR)*



### Ερμηνεία Αποτελεσμάτων Σύγκρισης 3:

Χρησιμοποιώντας έναν ταξινομητή (classifier) Extra Trees Classifier με επιλογή

χαρακτηριστικών (SFM\_ETC) σε 60 εξωτερικές επαναλήψεις, το μοντέλο πέτυχε μέση ακρίβεια 71% ( $\pm 13$ ), balanced accuracy 64% ( $\pm 14$ ), precision 76% ( $\pm 9$ ), ευαισθησία 84% ( $\pm 14$ ), ειδικότητα 45% ( $\pm 24$ ), F1 score 79% ( $\pm 10$ ) και ROC AUC 64% ( $\pm 14$ ).

Τα αποτελέσματα δείχνουν ότι ο ταξινομητής ήταν πιο αποτελεσματικός στην ανίχνευση ασθενών με αντικειμενική γνωστική έκπτωση (Mild Cognitive Impairment) (υψηλή ευαισθησία), παρά στη σωστή αναγνώριση συμμετεχόντων χωρίς γνωστική έκπτωση (χαμηλότερη ειδικότητα).

#### *Διαφορές στα εγκεφαλικά δίκτυα*

Οι αναλύσεις γραφηματικής θεωρίας (Graph-theoretical analyses) έδειξαν ότι οι συμμετέχοντες χωρίς γνωστική έκπτωση εμφάνισαν σημαντικά υψηλότερες τιμές σε δείκτες λειτουργικής συνδεσιμότητας όπως: flexibility, degree centrality, betweenness centrality, eigenvector centrality, and local efficiency σε αρκετές περιοχές του εγκεφάλου και συγκεκριμένα σε δίκτυα προσοχής (PrCv), ελέγχου (PFCI) και προεπιλεγμένης λειτουργίας (PFCd/PFCm). Αντίθετα, οι ασθενείς με γνωστική έκπτωση παρουσίασαν υψηλότερο degree centrality and local efficiency στον αριστερό βρεγματικό φλοιό του δικτύου ελέγχου και στον δεξιό σωματοκινητικό φλοιό. Τα ευρήματα υποδηλώνουν ότι η διατήρηση της γνωστικής λειτουργίας σχετίζεται με μεγαλύτερη ολοκλήρωση και προσαρμοστικότητα των εγκεφαλικών δικτύων σε περιοχές ελέγχου, προσοχής και του δικτύου προεπιλεγμένης λειτουργίας. Αντίθετα, οι ασθενείς με αντικειμενική γνωστική έκπτωση (MCI) εμφανίζουν τοπικές αυξημένες συνδεσιμότητες σε βρεγματικές και κινητικές περιοχές, κάτι που ενδέχεται να αντανακλά αντισταθμιστική ή δυσπροσαρμοστική αναδιοργάνωση της εγκεφαλικής λειτουργίας.

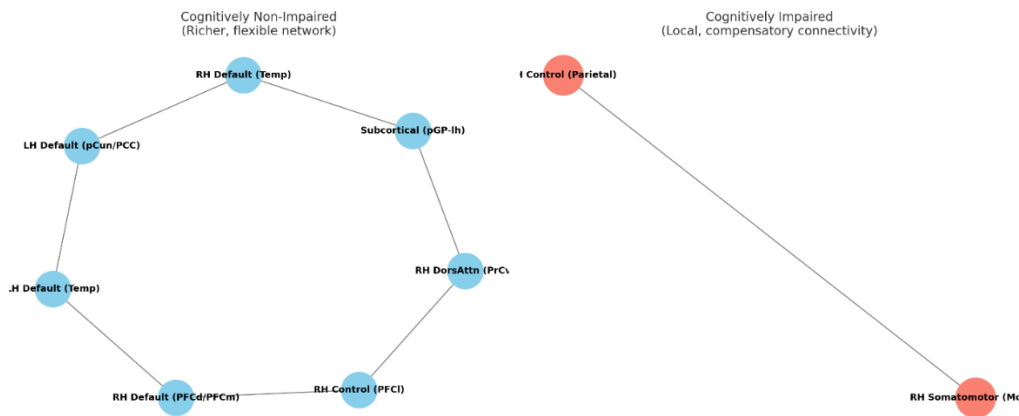
#### *Κλινική Χρήση*

Εκπαιδύσαμε ένα μοντέλο μηχανικής μάθησης για να κάνει διάκριση μεταξύ ασθενών με γνωστική έκπτωση και συμμετεχόντων χωρίς γνωστικά ελλείματα. Συνολικά, το μοντέλο απέδωσε αρκετά καλά, και ήταν ιδιαίτερα αποτελεσματικό στον εντοπισμό της γνωστικής έκπτωσης (84% ευαισθησία), αλλά λιγότερο αξιόπιστο στο να επιβεβαιώνει ποιοι ασθενείς δεν παρουσιάζουν έκπτωση (45% ειδικότητα).

Σε ό,τι αφορά τη λειτουργική συνδεσιμότητα του εγκεφάλου, τα άτομα χωρίς έκπτωση εμφάνισαν πιο υγιή και ευέλικτα εγκεφαλικά δίκτυα, ιδίως σε περιοχές που σχετίζονται με μνήμη, προσοχή και εκτελεστικό έλεγχο (δίκτυο προεπιλεγμένης λειτουργίας, δίκτυο ελέγχου και δίκτυα προσοχής). Αντίθετα, οι ασθενείς με έκπτωση παρουσίασαν αυξημένη συνδεσιμότητα σε βρεγματικές και κινητικές περιοχές, κάτι που μπορεί να αντανακλά μια αντισταθμιστική προσπάθεια του εγκεφάλου να αντιμετωπίσει τη γνωστική επιδείνωση.

#### *Σύνοψη:*

- Το ταξινομητικό μοντέλο μπορεί να συμβάλει στον πρώιμο εντοπισμό γνωστικής έκπτωσης, αν και απαιτούνται επιπλέον διαγνωστικά εργαλεία για τη βελτίωση της ειδικότητας.
- Τα ευρήματα των εγκεφαλικών δικτύων υποδηλώνουν ότι η απώλεια ευελιξίας και ολοκλήρωσης σε βασικούς εγκεφαλικούς κόμβους ίσως αποτελεί χαρακτηριστικό της πρώιμης γνωστικής έκπτωσης, ενώ η αυξημένη δραστηριότητα σε κινητικές/βρεγματικές περιοχές μπορεί να αντιπροσωπεύει αντισταθμιστικές αλλαγές.



**Figure 1.** Schematic representation of functional network differences between cognitively non-impaired and cognitively impaired participants. Left (blue): Non-impaired individuals displayed richer, more flexible, and more integrated connectivity, with central hubs across the default mode, control, and dorsal attention networks, as well as subcortical regions. Right (red): Impaired participants showed localized increases in connectivity, primarily between parietal control and somatomotor regions, suggesting compensatory or maladaptive reorganization.

## ΒΙΒΛΙΟΓΡΑΦΙΑ

- Zhao, J., Liu, J., & Chen, H. (2025). Thalamic degree centrality mediates the relationship between resilience and processing speed in healthy aging. *GeroScience*, 47(1), 101–115.
- Zhou, J., Greicius, M. D., Gennatas, E. D., et al. (2013). Divergent network connectivity changes in behavioural variant frontotemporal dementia and Alzheimer’s disease. *Brain*, 133(5), 1352–1367.
- Hwang, K., Bertolero, M. A., Liu, W. B., & D’Esposito, M. (2017). The human thalamus is an integrative hub for functional brain networks. *Journal of Neuroscience*, 37(23), 5594–5607.
- de Jong, L. W., van der Hiele, K., Veer, I. M., et al. (2008). Strongly reduced volumes of putamen and thalamus in Alzheimer’s disease: An MRI study. *Brain*, 131(12), 3277–3285.
- Serra, L., Cercignani, M., & Carlesimo, G. A. (2016). Connectivity-based parcellation of the thalamus in Alzheimer’s disease and mild cognitive impairment. *NeuroImage: Clinical*, 11, 283–290.
- Jiang, X., Shen, Y., Yao, J., et al. (2018). Increased thalamic connectivity in major depressive disorder: A resting-state fMRI study. *BMC Psychiatry*, 18(1), 62.