



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΣΥΣΤΗΜΑΤΩΝ ΜΕΤΑΔΟΣΗΣ ΠΛΗΡΟΦΟΡΙΑΣ ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ ΥΛΙΚΩΝ

Γενετική απεικόνιση και η νόσος του Αλτσχάιμερ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

του

ΠΡΙΟΒΟΥΛΟΥ ΝΙΚΟΛΑΟΥ

Επιβλέπων : Δημήτριος-Διονύσιος Κουτσούρης
Καθηγητής Ε.Μ.Π.

Αθήνα, Ιούλιος 2013



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΣΥΣΤΗΜΑΤΩΝ ΜΕΤΑΔΟΣΗΣ ΠΛΗΡΟΦΟΡΙΑΣ
ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ ΥΛΙΚΩΝ

Γενετική απεικόνιση και η νόσος του Αλτσχάιμερ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

του

Νικόλαου Πριόβουλου

Επιβλέπων : Δημήτριος-Διονύσιος Κουτσούρης
Καθηγητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 12^η Ιουλίου 2013.

.....
Δ.-Δ.Κουτσούρης
Καθηγητής Ε.Μ.Π.

.....
Γ. Ματσόπουλος
Επ. Καθηγητής Ε.Μ.Π.

.....
Δ. Φωτιάδης
Καθηγητής Παν. Ιωαννίνων

Αθήνα, Ιούλιος 2013

(Υπογραφή)

.....

ΝΙΚΟΛΑΟΣ ΠΡΙΟΒΟΥΛΟΣ

Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π.

Copyright © Νικόλαος Πριόβουλος, 2013.

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς το συγγραφέα.

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν το συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.

Περίληψη

Ο στόχος της διπλωματικής εργασίας ήταν η μελέτη των σύγχρονων βιοδεικτών της νόσου του Αλτσχάιμερ που γίνεται όλο και πιο κρίσιμη με την παγκόσμια γήρανση του πληθυσμού. Η νόσος του Αλτσχάιμερ είναι η πιο συχνή μορφή άνοιας και από τις πλέον συχνές νευροεκφυλιστικές ασθένειες.

Οι βιοδείκτες ενδιαφέροντος για τη νόσο του Αλτσχάιμερ στοχεύουν στην πρόβλεψη της νόσου, καθώς πιστεύεται ότι σε πρώιμα στάδια η νόσος είναι ευκολότερο να αντιμετωπιστεί. Ιδιαίτερο βάρος στην παρούσα εργασία δίνεται στις μεθόδους γενετικής απεικόνισης (ειδικά όσες αφορούν το σύνολο το γονιδιώματος) και στους συνδυαστικούς βιοδείκτες. Αυτές οι μέθοδοι έχουν εμφανιστεί πρόσφατα στην έρευνα για τη νόσο του Αλτσχάιμερ, ενώ παράλληλα η σχετική επιστημονική βιβλιογραφία μεγαλώνει ταχύτατα, κάτι που κάνει απαραίτητη την αξιολόγησή τους.

Η απεικόνιση στη νόσο του Αλτσχάιμερ εξετάζει είτε τη δομή του εγκεφάλου είτε τις διεργασίες του. Τα σημαντικότερα εργαλεία είναι οι διάφορες μορφές απεικόνισης πυρηνικού μαγνητικού συντονισμού και τομογραφίας εκπομπής ποζιτρονίων. Η επεξεργασία και ταξινόμηση των δεδομένων που προκύπτουν είναι επίσης μία διαδικασία με μεγάλο επιστημονικό ενδιαφέρον.

Η γενετική απεικόνιση αφορά είτε μεμονωμένα γονίδια είτε το σύνολο του γονιδιώματος και την έκφρασή τους στον εγκέφαλο. Παράλληλα εξετάζεται η συσχέτιση των γενετικών αιτιών στη νόσο του Αλτσχάιμερ με επιγενετικούς μηχανισμούς.

Τέλος οι συνδυαστικοί βιοδείκτες επιχειρούν να επεξεργαστούν το σύνολο των βιοδεικτών με μία κοινή μέθοδο, για να διευκολύνουν την ταξινόμηση και την ακόλουθη πρόβλεψη της ασθένειας ή τη διάγνωση.

Στη συγκεκριμένη διπλωματική αξιολογήθηκε κάθε ένα από τα παραπάνω τμήματα είτε ως προς το ρόλο τους στην κατανόηση της αιτιολογίας της νόσου του Αλτσχάιμερ, είτε ως προς τις διαγνωστικές δυνατότητές τους.

Λέξεις Κλειδιά: <<νόσος του Αλτσχάιμερ, απεικονιστικοί δείκτες, γενετικοί δείκτες, γενετική απεικόνιση, συστήματα υποστήριξης κλινικών αποφάσεων>>

Abstract

The target of this thesis is the study of the recent biomarkers of Alzheimer's Disease, which becomes increasingly more important with the aging of the world's population. Alzheimer's Disease is the most common kind of dementia and one of the most common neurodegenerative diseases.

The biomarkers of interest in Alzheimer's Disease are targeting the prediction of the disease, because it is widely believed that it is easier to confront the disease in its early stages. The study focuses mostly on the biomarkers of genetic neuroimaging (especially the ones concerned with the whole genome) and the multimodal biomarkers. These are methods that have only recently appeared in the scientific research on Alzheimer's Disease, while at the same time the relevant scientific bibliography is growing extremely fast, obviously necessitating their thorough study and evaluation.

The imaging of Alzheimer's Disease focuses either on the structures of the brain or on the functions of the brain. The most important tools are the various forms of magnetic resonance imaging and of positron emission tomography. The processing and classification of the acquired data is also a process that holds significant scientific interest.

Genetic imaging is concerned either with specific genes or with the whole genome and their expression on the brain. At the same time the association between the genetic etiology of Alzheimer's Disease and the epigenetic mechanisms is also under scrutiny.

Finally, the multimodal biomarkers are trying to process the whole of the biomarkers with a common method to facilitate the classification and the following prognosis or diagnosis.

In this particular thesis each of the above was studied either in relation to the understanding they offer us on the etiology of Alzheimer's Disease, or their role in the diagnosis of the disease.

Keywords: <<Alzheimer's Disease, neuroimaging markers, genetic markers, genetic neuroimaging, clinical decision support systems>>

Πρόλογος

Η παρούσα διπλωματική εργασία εκπονήθηκε στο Εργαστήριο Βιοϊατρικής Τεχνολογίας της σχολής Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών του Εθνικού Μετσόβιου Πολυτεχνείου.

Για την ολοκλήρωσή της θέλω να εκφράσω τις θερμές μου ευχαριστίες προς τον επιβλέποντα καθηγητή, κ. Κουτσούρη, για την πολύτιμη υποστήριξη που μου προσέφερε και την ευκαιρία που μου έδωσε να ασχοληθώ με το συγκεκριμένο θέμα. Επίσης θέλω να ευχαριστήσω τη Δρα. Μαρία Χαρίτου για τη συνολική της υποστήριξη στην εκπόνηση της διπλωματικής, όπως και τον υποψήφιο Δρα. Παναγιώτη Μπουντρή, για την αμέριστη βοήθεια που μου έδωσε. Παράλληλα, θέλω να ευχαριστήσω και τα υπόλοιπα μέλη της εξεταστικής επιτροπής, τον κ. Ματσόπουλο, επίκουρο καθηγητή της σχολής Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών του Ε.Μ.Π. και κ. Φωτιάδη, καθηγητή του τμήματος Μηχανικών Επιστήμης Υλικών του Πανεπιστημίου Ιωαννίνων.

Η συγκεκριμένη διπλωματική είναι μία ανασκόπηση της σύγχρονης έρευνας στη νόσο του Αλτσχάιμερ, με ιδιαίτερη βαρύτητα στη γενετική απεικόνιση. Το πρώτο κεφάλαιο συνιστά μία γενική εισαγωγή στην παθολογία της νόσου και στη σύγχρονη έρευνα σε αυτή. Το δεύτερο κεφάλαιο εξετάζει τις απεικονιστικές μεθόδους που χρησιμοποιούνται και την ακόλουθη στατιστική επεξεργασία τους. Το τρίτο κεφάλαιο ασχολείται κυρίως με τις έρευνες γενετικής απεικόνισης στη νόσο και τη σύνδεση του γονιδιώματος με εξωτερικούς παράγοντες στην εμφάνιση της νόσου. Το τέταρτο κεφάλαιο μελετά το συνδυασμό όλων των παραπάνω μεθόδων σε ένα ενιαίο σύστημα και την εισαγωγή τους στην κλινική διαδικασία. Στο πέμπτο κεφάλαιο εξάγονται τα συνολικά συμπεράσματα για την έρευνα στη νόσο του Αλτσχάιμερ, καθώς και πιθανές μελλοντικές κατευθύνσεις στην έρευνα.

Πίνακας Περιεχομένων

Περίληψη.....	6
Abstract.....	7
Πρόλογος.....	8
Πίνακας Περιεχομένων.....	9
Κατάλογος Εικόνων.....	13
Κατάλογος Πινάκων.....	14
Κατάλογος Διαγραμμάτων.....	15
Κατάλογος Συνοτομογραφιών.....	16
ΚΕΦΑΛΑΙΟ 1: Η ΝΟΣΟΣ ΤΟΥ ΑΛΤΣΧΑΪΜΕΡ.....	18
1.1 Εισαγωγή.....	18
1.2 Ιστορικά για τη νόσο του Αλτσχάιμερ.....	19
1.3 Ανατομία του ανθρώπινου εγκεφάλου.....	20
1.3 Αλλαγές στον εγκέφαλο από τη φυσιολογική γήρανση.....	21
1.4 Νευροπαθολογία της νόσου του Αλτσχάιμερ.....	22
1.5 Διάγνωση της νόσου του Αλτσχάιμερ.....	23
1.6 Ήπια Νοητική Διαταραχή.....	24
1.7 Νευροψυχολογικές δοκιμασίες και δοκιμασίες γνωσιακής λειτουργίας.....	25
1.8 Γιατί οι νευροψυχολογικές δοκιμασίες δεν αρκούν.....	29
1.9 Βιοδείκτες της νόσου του Αλτσχάιμερ: Πρώιμη ανίχνευση της νόσου.....	29
1.9.1 Νευροαπεικονιστικοί βιοδείκτες.....	29
1.9.2 Εγκεφαλονωτιαίο υγρό και πλάσμα.....	30
1.9.3 Γενετικοί βιοδείκτες.....	31
1.9.4 Συνδυαστικοί βιοδείκτες.....	31
1.10 Πρωτοβουλία νευροαπεικόνισης της νόσου του Αλτσχάιμερ.....	32
ΚΕΦΑΛΑΙΟ 2: ΕΦΑΡΜΟΓΕΣ ΝΕΥΡΟΑΠΕΙΚΟΝΙΣΤΙΚΩΝ ΒΙΟΔΕΙΚΤΩΝ.....	33
2.1 Απεικονιστικοί βιοδείκτες.....	33
2.1.1 Απεικόνιση πυρηνικού μαγνητικού συντονισμού.....	33
2.1.2 Τομογραφία εκπομπής ποζιτρονίων.....	36
2.1.3 Ηλεκτροεγκεφαλογράφημα.....	37
2.2 Επεξεργασία εγκεφαλικών εικόνων.....	40

2.2.1 Μορφομετρία του εγκεφάλου	40
2.2.2 Προεπεξεργασία.....	41
2.2.3 Κατάτμηση	41
2.2.4 Επιλογή χαρακτηριστικών– με μορφομετρία βασισμένη σε voxels	41
2.2.5 Μορφομετρία βασισμένη σε tensors	42
2.2.6 Άλλες μέθοδοι.....	42
2.3 Παραγοντοποίηση του σήματος σε πηγές – Αυτοματοποιημένη εξαγωγή και επιλογή χαρακτηριστικών	43
2.3.1 Ανάλυση σε πρωταρχικές συνιστώσες	44
2.3.2 Ανάλυση σε ανεξάρτητες συνιστώσες	45
2.3.3 Παράλληλη ανάλυση σε ανεξάρτητες συνιστώσες.....	47
2.3.4 Παλινδρόμηση	50
2.3.5 Παλινδρόμηση κορυφογραμμής	50
2.3.6 Μέθοδος least absolute shrinkage and selection operator	51
2.3.7 Ελαστικό δίκτυο.....	51
2.3.8 Παλινδρόμηση μειωμένου βαθμού	51
2.3.9 Γραμμική διακριτική ανάλυση	52
2.3.10 Επιβαρυσμένη γραμμική διακριτική ανάλυση.....	53
2.3.11 Λογιστική παλινδρόμηση	53
2.3.12 Πολυδιακριτή ανάλυση	53
2.4 Ταξινομητές.....	53
2.4.1 Μηχανές διανυσμάτων υποστήριξης.....	55
2.4.2 Αλγόριθμος ADABOOST.....	56
2.4.3 Διανύσματα υπολοίπων.....	60
2.4.4 Ταξινομητής τυχαίου δάσους	61
ΚΕΦΑΛΑΙΟ 3: ΕΦΑΡΜΟΓΕΣ ΓΕΝΕΤΙΚΩΝ ΔΕΙΚΤΩΝ	62
3.1 Νόσος του Αλτσχάιμερ και γενετική.....	62
3.2 Συνδυασμός με απεικονιστικές τεχνικές.....	62
3.3 Μελέτες υποψηφίων γονιδίων	63
3.3.1 Βιοδείκτες για την πρόωμη ανίχνευση της νόσου του Αλτσχάιμερ.....	63
3.3.2 Εφαρμογή στη νόσο του Αλτσχάιμερ πρόωμης εκδήλωσης: Πρόωμες διαφοροποιήσεις.....	63
3.3.3 Πρόωμες διαφοροποιήσεις στη νόσο του Αλτσχάιμερ όψιμης εκδήλωσης: Αλλαγές στη μικροδομή της λευκής ουσίας	64

3.3.4	Πρώιμες διαφοροποιήσεις στη νόσο του Αλτσχάιμερ όψιμης εκδήλωσης: Αλλαγές στον όγκο	66
3.4	Μελέτες συσχέτισης γονιδιώματος.....	68
3.4.1	Πρώτη εφαρμογή μελέτης συσχέτισης γονιδιώματος στη λειτουργία του εγκεφάλου: Διάγνωση σχιζοφρένειας.....	70
3.4.2	Πρώτη μεγάλη μελέτη συσχέτισης γονιδιώματος στη νόσο του Αλτσχάιμερ: Σημασία στατιστικής επεξεργασίας.....	71
3.4.3	Νέοι δείκτες σε μελέτη συσχέτισης γονιδιώματος: Florbetapir-PET.....	74
3.4.4	Πέρα από τους απεικονιστικούς δείκτες: Εκτεταμένη μελέτη συσχέτισης γονιδιώματος για δείκτες του εγκεφαλονωτιαίου υγρού.....	77
3.5	Πολυπαραγοντικές μελέτες συσχέτισης γονιδιώματος.....	79
3.5.1	Γιατί απαιτούνται πολυπαραγοντικές μελέτες συσχέτισης γονιδιώματος: Δομή του γενότυπου	79
3.5.2	Παλινδρόμηση κορυφογραμμής για την ομαδοποίηση πολυμορφισμών: Αύξηση στατιστικής ισχύος.....	80
3.5.3	Προσομοίωση με αραιή μέθοδο μειωμένου βαθμού για μελέτη συσχέτισης γονιδιώματος στη νόσο του Αλτσχάιμερ	82
3.5.4	Εφαρμογή της αραιής μεθόδου μειωμένου βαθμού σε πραγματικά δεδομένα: Πλήρης πολυπαραγοντική εφαρμογή	82
3.5.5	Πολυπαραγοντική ταξινόμηση του γονιδιώματος με πολυδιακριτή ανάλυση: Ανάπτυξη γενετικού δείκτη χωρίς απεικονιστικά δεδομένα	88
3.5.6	Εφαρμογή παράλληλης ανάλυσης σε ανεξάρτητες συνιστώσες σε πολυπαραγοντική μελέτη συσχέτισης γονιδιώματος: Αντιστοίχιση με βιολογικά δίκτυα του εγκεφάλου	91
3.6	Συνολικά για τις μελέτες συσχέτισης γονιδιώματος στη νόσο του Αλτσχάιμερ	97
3.7	Γενετικοί και περιβαλλοντικοί παράγοντες.....	98
3.7.1	Επιγενετική.....	99
3.7.2	Επιγενετικοί μηχανισμοί	100
3.7.3	Γήρανση, επιγενετική και νόσος του Αλτσχάιμερ	100
3.7.4	Επιγενετική, νόσος του Αλτσχάιμερ και θεραπεία	101
3.8	Πρωτεομική και γενετική στη νόσο του Αλτσχάιμερ.....	102
ΚΕΦΑΛΑΙΟ 4: ΣΥΝΔΥΑΣΤΙΚΕΣ ΜΕΘΟΔΟΙ ΒΙΟΔΕΙΚΤΩΝ ΓΙΑ ΤΗ ΝΟΣΟ ΤΟΥ ΑΛΤΣΧΑΪΜΕΡ		
.....		106
4.1	Συνδυαστικές μέθοδοι.....	106
4.1.1	Αξιολόγηση συστημάτων συνδυασμού πηγών δεδομένων στη νόσο του Αλτσχάιμερ: Πλατφόρμα πολυπύρηνης εκμάθησης	108

4.1.2 Αξιολόγηση συστημάτων συνδυασμού πηγών δεδομένων στη νόσο του Αλτσχάιμερ: Χρήση διανυσμάτων υποστήριξης.....	110
4.1.3 Εισαγωγή της συμπλήρωσης δεδομένων σε συστήματα συνδυασμού πηγών δεδομένων	114
4.2 Εφαρμογή στη διαγνωστική διαδικασία	121
4.2.1 Εξατομίκευση και κόστος στα συστήματα υποστήριξης κλινικών αποφάσεων.....	121
4.2.2 Συνδυασμός βιοδεικτών: Αλληλοεπικάλυψη πληροφορίας;	123
4.2.3 Η εισαγωγή ενός αγνωστικιστικού συστήματος ως προδρόμου ενός κλινικού εργαλείου ...	125
4.3 Συμπεράσματα για τα πολυτροπικά συστήματα.....	129
ΚΕΦΑΛΑΙΟ 5: ΣΥΜΠΕΡΑΣΜΑΤΑ – ΜΕΛΛΟΝΤΙΚΗ ΕΡΕΥΝΑ.....	130
5.1 Συνολικά συμπεράσματα	130
5.2 Νέες ερευνητικές προοπτικές και στόχοι.....	131
Βιβλιογραφία	134

Κατάλογος Εικόνων

Εικόνα 1: Βασικά Μέρη Εγκεφάλου [3]

Εικόνα 2: Διαφορά στο στόχο των sMRI και fMRI [16]

Εικόνα 3: Δείγμα απεικόνισης fMRI που αποτυπώνει τη διαφοροποίηση της αιμάτωσης του εγκεφάλου [17]

Εικόνα 4: Αλγόριθμος για τη σύνδεση νοητικής κατάστασης ασθενών με NA και χαρακτηριστικών δικτύου [18]

Εικόνα 5: Παράδειγμα εξαγωγής επιπέδου από χώρο με την εφαρμογή της ανάλυσης σε πρωταρχικές συνιστώσες σε τρισδιάστατα δεδομένα [20]

Εικόνα 6: Ανάλυση σε ανεξάρτητες συνιστώσες σε σχέση με το γενικό γραμμικό μοντέλο [21]

Εικόνα 7: Ομαδοποίηση δύο πηγών δεδομένων με παράλληλη ανάλυση σε ανεξάρτητες συνιστώσες [22]

Εικόνα 8: Ομαδοποίηση με sRRR και με πολλαπλή πολυπαραγοντική παλινδρόμηση [23]

Εικόνα 9: Αλγόριθμος ταξινόμησης ADABoost σε συνδυασμό με SVM και διαφορά από απλό ADABoost[27]

Εικόνα 10: Απεικόνιση των περιοχών του εγκεφάλου που σε μικρή ηλικία εμφανίζουν μεγαλύτερη μεταβολή στη συνοχή της λευκής ουσίας με κάθε επιπλέον CLU-C αλλέλη [32]

Εικόνα 11: Απεικόνιση των περιοχών του εγκεφάλου που ο όγκος μεταβάλλεται σε πρώιμη ηλικία από το γονίδιο GAB2 [33]

Εικόνα 12: Επιλογή φαινότυπων με LDA και επιβολή sRRR στα δεδομένα για εξαγωγή γενετικών συσχετίσεων. Ο πίνακας υπογραφής εκφράζει τον πίνακα απεικονιστικών βιοδεικτών που επιλέχθηκαν [42]

Εικόνα 13: Διαφορά διαδικασίας GWAS και πολυδιακριτής ανάλυσης σε συνδυασμό με GWAS [24]

Εικόνα 14: Σχηματική αναπαράσταση της εφαρμογής του paraICA στην επεξεργασία γενετικών και απεικονιστικών δεδομένων σε GWAS [19]

Εικόνα 15: Κύρια δίκτυα του εγκεφάλου που επηρεάζονται από τη ΝΑ, η δυσλειτουργία τους και τα γονίδια που εκφράζονται μέσω αυτών [19]

Εικόνα 16: Παθολογία της ΝΑ και οι επιγενετικοί μηχανισμοί που θεωρείται ότι συνδέονται με αυτή [44]

Εικόνα 17: Παραγωγή γράφου πρωτεϊνικού σήματος από γενετικά δεδομένα για ΝΑ και πρωτεϊνική βάση δεδομένων έπειτα από στατιστική επεξεργασία [53]

Εικόνα 18: Σχηματικό διάγραμμα της πολυπύρηνης πολυτροπικής επεξεργασίας και ταξινόμησης δεδομένων [58]

Εικόνα 19: Κατάλληλη δομή δεδομένων για «block-wise» συμπλήρωση δεδομένων, δηλαδή με ολόκληρα τμήματα δεδομένων να λείπουν [59]

Εικόνα 20: Τα τέσσερα διαφορετικά μοντέλα που δημιουργήθηκαν για εκτίμηση τιμών από τον iMSF αλγόριθμο [59]

Εικόνα 21: Δημιουργία μοντέλου συμπλήρωσης επίδοσης στα ελλιπή δεδομένα [59]

Εικόνα 22: Εξατομίκευση της επιλογής βιοδεικτών για ημι-αυτοματοποιημένη διάγνωση της ΝΑ με k -πιο κοντινούς γείτονες και συνυπολογισμό του κόστους και της αποτελεσματικότητας [60]

Εικόνα 23: Περιπτώσεις ταξινόμησης του συστήματος [62] ανάμεσα σε ασθενή με ΝΑ και υγιή και τα αντίστοιχα κριτήρια

Κατάλογος Πινάκων

Πίνακας 1: Σύγκριση επιδόσεων paraICA με ICA [22].

Πίνακας 2: Επιδόσεις του ADABoost στην επιλογή εικόνας [26]

Πίνακας 3: Αποτελέσματα άλλων αυτοματοποιημένων μεθόδων κατάτμησης μέχρι το 2008. Ο ADABoost έχει καλύτερες επιδόσεις [26]

Πίνακας 4: Επιδόσεις ADABoost, ADABoost με SVM και FreeSurfer. Τα μεγέθη εκφράζουν το ποσοστό της διαφοράς σε σχέση με την κατάτμηση από άνθρωπο [27]

Πίνακας 5: Χαρακτηριστικά που επέλεξε ο αλγόριθμος ADABoost για την επιλογή εικόνας [27]

Πίνακας 6: Αλλαγές στον εγκέφαλο από το presinilin-1[31]

Πίνακας 7: Τα SNPs με τη μεγαλύτερη επιρροή για CSF,t-tau και p-tau βιοδείκτες [39]

Πίνακας 8: Τα σημαντικά SNPs στο πείραμα ασθενών με ΝΑ και υγιών οργανωμένα σύμφωνα με το γονίδιο. Στη δεξιά στήλη η εκτιμώμενη στατιστική σημασία [42]

Πίνακας 9: Τα σημαντικά SNPs στο πείραμα ασθενών με ΝΑ και ασθενών με εξελισσόμενο MCI. Στη δεξιά στήλη η εκτιμώμενη στατιστική σημασία [42]

Πίνακας 10: Πίνακας των σημαντικών SNPs σύμφωνα με την έρευνα [19]. Τα αποτελέσματα είναι οργανωμένα στις γενετικές ομάδες που σχηματίστηκαν και με κατάταξη αντίστοιχη της σημασίας τους

Πίνακας 11: Αλλαγές σε επιγενετικούς μηχανισμούς σε ασθενείς με ΝΑ και οι αντίστοιχες μελέτες [44]

Πίνακας 12: Οι επιδόσεις της ταξινόμησης του συνολικού απεικονιστικού δείκτη, του συνολικού βιολογικού δείκτη, του συνολικού γνωσιακού δείκτη και του ολικού δείκτη στην ταξινόμηση ασθενών με ΝΑ και υγιών [56]

Πίνακας 13: Σύγκριση των επιδόσεων της ταξινόμησης των διαφορετικών πηγών δεδομένων και των ενοποιημένων πηγών [58]

Πίνακας 14: Σύγκριση των επιδόσεων των διαφόρων μεθόδων συμπλήρωσης δεδομένων[59]

Πίνακας 15: Σύγκριση των επιδόσεων πολυτροπικών συστήματος με το «score-completion» τρόπο συμπλήρωσης δεδομένων και τον iMSF [59]

Πίνακας 16: Επιδόσεις στην ταξινόμηση για τις διάφορες ομάδες [61]. Τα πιο σημαντικά χαρακτηριστικά είναι οι νευροψυχολογικές δοκιμασίες

Πίνακας 17: Επιδόσεις ταξινόμησης σε εξελισσόμενο MCI και στάσιμο MCI με τους διαφορετικούς αλγορίθμους [62]

Κατάλογος Διαγραμμάτων

Διάγραμμα 1: Αλλαγές στον όγκο του κροταφικού λοβού για τα SNPs με μεγαλύτερη επιρροή [38]

Διάγραμμα 2: Τα SNPs με μεγαλύτερη επιρροή από τη florbetapir-PET GWAS [5]

Διάγραμμα 3: Μεγάλη διαφοροποίηση επιρροής με εφαρμογή παλινδρόμησης κορυφογραμμής στο GWAS [41]

Διάγραμμα 4: Επίδοση των διαφόρων πηγών δεδομένων στην ταξινόμηση ασθενών με Αλτσχάιμερ και υγιών και ασθενών με MCI και υγιών [58]

Κατάλογος Συντομογραφιών

Συντομογραφία	Όρος στη διεθνή βιβλιογραφία	Όρος στην ελληνική βιβλιογραφία
ADNI	Alzheimer's Disease Neuroimaging Initiative	Πρωτοβουλία Νευροαπεικόνισης της Νόσου του Alzheimer
CDR	Clinical Dementia Rating	Κλινική Εκτίμηση της Άνοιας
CDSS	Clinical Decision Support Systems	Συστήματα Υποστήριξης Κλινικών Αποφάσεων
CNV	Copy Number Variation	Αριθμός Αντιγράφων DNA
CSF	Cerebrospinal Fluid	Εγκεφαλονωτιαίο Υγρό
DNA	Deoxyribonucleic Acid	Δεσοξυριβονουκλεϊκό Οξύ
EEG	Electroencephalography	Ηλεκτροεγκεφαλογράφημα
EOAD	Early Onset Alzheimer's Disease	ΝΑ Πρώιμης Εκδήλωσης
FDG	Fluorodeoxyglucose	Φθοριοδεοξυγλυκόζη
fMRI	functional Magnetic Resonance Imaging	Λειτουργική Απεικόνιση Πυρηνικού Μαγνητικού Συντονισμού
GWAS	Genome Wide Association Study	Μελέτη Συσχέτισης Γονιδιώματος
ICA	Independent Component Analysis	Ανάλυση σε Ανεξάρτητες Συνιστώσες
LDA	Linear Discriminant Analysis	Γραμμική Διακριτική Ανάλυση
LOAD	Late Onset Alzheimer's Disease	ΝΑ Όψιμης Εκδήλωσης
MCI	Mild Cognitive Impairment	Ήπια Νοητική Διαταραχή

MMSE	Mini Mental State Examination	Βραχεία Δοκιμασία Ελέγχου της Νοητικής Κατάστασης
MRI	Magnetic Resonance Imaging	Απεικόνιση Πυρηνικού Μαγνητικού Συντονισμού
NA		Νόσος του Alzheimer
paraICA	Parallel Independent Component Analysis	Παράλληλη Ανάλυση σε Ανεξάρτητες Συνιστώσες
PCA	Principal Component Analysis	Ανάλυση σε Πρωταρχικές Συνιστώσες
penalized LDA	Penalized Linear Discriminant Analysis	Επιβαρυσμένη Γραμμική Διακριτική Ανάλυση
PET	Positron Emission Tomography	Τομογραφία Εκπομπής Ποζιτρονίων
p-tau	Phosphorylated-tau	Φωσφορυλιωμένη-tau
sMRI	Structural Magnetic Resonance Imaging	Δομική Απεικόνιση Πυρηνικού Μαγνητικού Συντονισμού
SNP	Single-Nucleotide Polymorphism	Σημειακοί Πολυμορφισμοί
SVM	Support Vector Machines	Μηχανές Διανυσμάτων Υποστήριξης
TBM	Tensor Based Morphometry	Μορφομετρία βασισμένη σε tensors
t-tau	Total-tau	Ολική-tau
VBM	Voxel Based Morphometry	Μορφομετρία βασισμένη σε voxels

ΚΕΦΑΛΑΙΟ 1: Η ΝΟΣΟΣ ΤΟΥ ΑΛΤΣΧΑΪΜΕΡ

1.1 Εισαγωγή

Η Νόσος του Αλτσχάιμερ (NA, Alzheimer's Disease) είναι η πιο κοινή μορφή άνοιας και αποτελεί μία από τις μεγαλύτερες ιατρικές προκλήσεις καθώς πλήττει πάνω από 26 εκατομμύρια ανθρώπους σε όλο τον κόσμο. Με τη συνεχή αύξηση του αριθμού των υπερηλίκων, ο αριθμός των ανθρώπων που πάσχουν από NA αναμένεται να αυξηθεί κατακόρυφα τις επόμενες δεκαετίες και μπορεί να φθάσει μέχρι και το 1/85 του πληθυσμού το 2050 [1]. Ταυτόχρονα είναι και μία από τις πιο “ακριβές” ασθένειες, τόσο για το σύστημα υγείας όσο και για το κοινωνικό περιβάλλον του ασθενούς, δεδομένου ότι σε προχωρημένο στάδιο ο ασθενής βασίζεται σε άλλους για όλες του τις λειτουργίες.

Η νόσος εντάσσεται στις εξελισσόμενες νευροεκφυλιστικές ασθένειες και χαρακτηρίζεται από γενικευμένη μείωση της νοητικής δυνατότητας, που περιλαμβάνει σταδιακή απώλεια μνήμης, αιτίασης και γλώσσας. Εμφανίζεται συνήθως σε προχωρημένη ηλικία και προς το παρόν δεν υπάρχει ίαση ή αποτελεσματική πρόληψη. Τα φαρμακευτικά σκευάσματα που διατίθενται έχουν περιορισμένη αποτελεσματικότητα, αλλά εκτιμάται ότι όσο νωρίτερα χρησιμοποιηθούν, τόσο πιο αποτελεσματικά είναι. Ως εκ τούτου, η πρόγνωση αυτών που θα νοσήσουν από NA, η έγκαιρη διάγνωση της NA και η καθυστέρηση της εξέλιξής της είναι οι βασικοί στόχοι στην αντιμετώπιση της νόσου. Πρόσφατες έρευνες δείχνουν ότι η καθυστέρηση ακόμα και 5 χρόνων στην εμφάνιση της νόσου, θα μειώσει στο μισό τον αριθμό των ασθενών από τη NA [2].

Η NA στη συντριπτική πλειονότητα (>95%) εμφανίζεται σε ηλικίες μεγαλύτερες των 60 ετών (NA όψιμης εκδήλωσης - late onset Alzheimer, LOAD). Συμπτώματα της νόσου μπορεί να εμφανιστούν και σε πολύ μικρότερες ηλικίες (30 - 40 ετών, NA πρώιμης εκδήλωσης - early onset Alzheimer, EOAD), αλλά αυτή η μορφή της νόσου έχει ταυτιστεί ήδη από τη δεκαετία του '90 με συγκεκριμένα γονίδια (APP - 1991, presenilin1 - 1995, presenilin2 - 1995). Αντίθετα τα αίτια της νόσου στην πιο συχνή της μορφή είναι σαφώς πιο περίπλοκα: υπάρχουν συγκεκριμένα γονίδια που έχουν συνδεθεί με την εμφάνιση της NA και εξηγούν μέχρι ένα σημείο την εμφάνισή της, αλλά η αιτία της νόσου φαίνεται να είναι ένας πιο περίπλοκος συνδυασμός γενετικών και περιβαλλοντικών παραγόντων, καθώς και απόρροια συγκεκριμένων μηχανισμών του οργανισμού.

Οι πρόσφατες έρευνες με θέμα τη NA αφορούν το χειρισμό τεράστιων βάσεων δεδομένων και τη διαλεύκανση των σχετιζόμενων πληροφοριών. Αυτή ακριβώς είναι η πρώτη έκφανση της συμβολής της

βιοπληροφορικής στην αντιμετώπιση της νόσου: η ολοκλήρωση εκτεταμένων γενετικών ερευνών σε συνδυασμό με απεικονιστικές μεθόδους του ανθρώπινου εγκεφάλου, η διαλεύκανση των παραγόντων που σχετίζονται με τη νόσο μέσω συστημάτων υποστήριξης αποφάσεων (decision support systems) και η στατιστική επεξεργασία της πληροφορίας. Ένα άλλο εκτεταμένο τεχνολογικό πεδίο στην αντιμετώπιση της ΝΑ είναι η επιτυχής διάγνωσή της, που απαιτεί επίσης τη συμβολή συστημάτων υποστήριξης αποφάσεων στις νευροψυχολογικές δοκιμασίες, στην απεικονιστική επεξεργασία και στο συνδυασμό τους. Τέλος, ένα ακόμη ενδιαφέρον πεδίο βιοπληροφορικής ασχολείται με τη βελτίωση της ποιότητας ζωής των ασθενών με ΝΑ, μέσω εργαλείων που διευκολύνουν την καθημερινότητά τους. Το τελευταίο είναι ένα θέμα που δεν θα μας απασχολήσει πολύ στο πλαίσιο της συγκεκριμένης διπλωματικής.

1.2 Ιστορικά για τη νόσο του Αλτσχάιμερ

Το Νοέμβριο του 1906, στο Tübingen, στη Γερμανία ο Alois Alzheimer (1864-1915) περιέγραψε για πρώτη φορά τα ευρήματα του εργαστηρίου του όσον αφορά μια άγνωστη νευρολογική ασθένεια που εμφανίστηκε σε μία ασθενή, την Auguste D. Η Auguste D. μπήκε στο νοσοκομείο σε ηλικία 51 ετών. Πέθανε 4 χρόνια μετά από πνευμονία, εξαιτίας της εξελισσόμενης άνοιας.

Παρότι τα κλινικά χαρακτηριστικά της “ασθένειας των ηλικιωμένων” ήταν γνωστά από την αρχαιότητα, ο Alzheimer ήταν ο πρώτος που συνέδεσε την εμφάνιση και επέκταση της ασθένειας με τις γεροντικές ή αμυλοειδικές πλάκες (senile plaques) και τα νευροϊνιδιακά τολύπια (neurofibrillary tangles).

Από τότε έως σήμερα, χιλιάδες επιστημονικές μελέτες έχουν γίνει σε σχέση με τη ΝΑ. Νέα χαρακτηριστικά της νευροπαθολογίας της ΝΑ κατεγράφησαν, αλλά χωρίς να ερμηνεύσουν την ασθένεια. Η εισαγωγή των μεθόδων απεικόνισης του εγκεφάλου επέτρεψε την παρακολούθηση της εξέλιξης της ασθένειας σε ασθενείς εν ζωή.

Η αποκωδικοποίηση του ανθρώπινου DNA έδωσε νέα ώθηση στην έρευνα για τη ΝΑ. Ερμηνεύθηκε πρώτα από όλα η καθαρά κληρονομική μορφή της ΝΑ, με τον εντοπισμό των σχετικών γονιδίων. Από τότε ως σήμερα, η γενετική ιατρική προσπαθεί να επαναλάβει την επιτυχία για την πιο συνηθισμένη μορφή της ΝΑ. Είναι πλέον σαφές ότι το LOAD έχει εν μέρει γενετική ερμηνεία, εξίσου σαφές όμως είναι ότι η γενετική μόνη της δεν αρκεί για να εξηγήσει τη ΝΑ. Νέες υπολογιστικές

μέθοδοι που εφαρμόστηκαν την τελευταία δεκαετία επιτρέπουν προσεγγίσεις που μέχρι τώρα ήταν αδύνατες και δίνουν πολύ υποσχόμενα αποτελέσματα.

1.3 Ανατομία του ανθρώπινου εγκεφάλου

Προτού εξετάσουμε τη ΝΑ, είναι αναγκαίο να αναφερθούμε εν συντομία στη δομή του ανθρώπινου εγκεφάλου και στο ρόλο του κάθε μέρους του. Πρέπει να γίνει σαφές βέβαια ότι ο εγκέφαλος είναι ένα περίπλοκο όργανο και συνήθως δεν υπάρχει μία μόνο περιοχή αποκλειστικά υπεύθυνη για κάποια λειτουργία. Οι παρακάτω περιγραφές δίνονται κυρίως σαν γενικές αρχές.

Ο μεγάλος όγκος του ανθρώπινου εγκεφάλου (περίπου το 85%) αποτελείται από τα δύο ημισφαίρια. Τα ημισφαίρια περιβάλλονται από ένα εξωτερικό στρώμα, τον εγκεφαλικό φλοιό, όπου εισάγεται η πληροφορία από τις αισθήσεις και την κίνηση. Κάθε ημισφαίριο έχει 4 λοβούς:

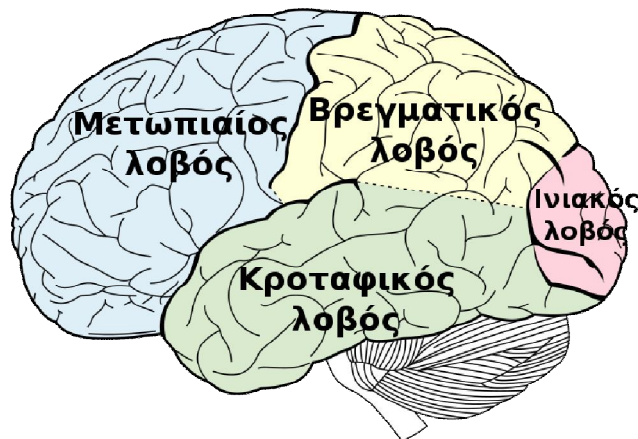
- Τον μετωπιαίο λοβό (frontal lobe), που επικεντρώνεται στην επίλυση προβλημάτων, στην οργάνωση και το σχεδιασμό.
- Τον πλευρικό ή βρεγματικό λοβό (parietal lobe), που βρίσκεται πίσω από το μετωπιαίο λοβό και συγκεντρώνει τις πληροφορίες από τις αισθήσεις και παρέχει, για παράδειγμα, την αίσθηση της προοπτικής.
- Τον ινιακό λοβό (occipital lobe), που είναι στο πίσω μέρος του εγκεφάλου και ασχολείται με τις λειτουργίες της όρασης.
- Τον κροταφικό λοβό (temporal lobe), που εντοπίζεται στο πλαϊνό μέρος του εγκεφάλου, κάτω από το μετωπιαίο και πλευρικό και καθορίζει πώς αντιλαμβανόμαστε γεύσεις, οσμές, ήχους και το σχηματισμό των μνημών.

Τα υπόλοιπα βασικά μέρη του εγκεφάλου είναι:

- Η παρεγκεφαλίδα (cerebellum) που βρίσκεται πάνω από το εγκεφαλικό στέλεχος και κάτω από τον ινιακό λοβό, καθορίζει την ισορροπία και το συντονισμό του ανθρώπου. Έχει δύο ημισφαίρια και μπορούμε να την αντιμετωπίσουμε σαν ένα κλειστό σύστημα που δέχεται πληροφορία από τις αισθήσεις και εξάγει οδηγίες στο σώμα.
- Το εγκεφαλικό στέλεχος (brain stem), στη βάση του ανθρώπινου εγκεφάλου, συνδέεται με τη σπονδυλική στήλη και ελέγχει τις αυτοματοποιημένες διαδικασίες, όπως τον καρδιακό παλμό, την αρτηριακή πίεση, την αναπνοή αλλά και τον ύπνο και τα όνειρα.

- Το λιμβικό σύστημα (limbic system) που αποτελεί ένα δίκτυο δομών μέσα στα ημισφαίρια. Συνδέει το εγκεφαλικό στέλεχος με τις ανώτερες επεξεργασίες του εγκεφαλικού φλοιού και συμμετέχει σε λειτουργίες όπως η ανάπτυξη ενστικτωδών αντιδράσεων, συναισθημάτων ή η σύνδεση οσμών με κάποια μνήμη. Τα μέρη του περιλαμβάνουν την αμυγδαλή, που επεξεργάζεται ισχυρά συναισθήματα, όπως ο φόβος, και βρίσκεται στον κροταφικό λοβό, τον ιππόκαμπο, που επίσης βρίσκεται στον κροταφικό λοβό και ελέγχει τη μάθηση και την πρόσφατη μνήμη, το θάλαμο, πάνω ακριβώς από το εγκεφαλικό στέλεχος, που επεξεργάζεται πληροφορίες των αισθήσεων και τον υποθάλαμο που καθορίζει λειτουργίες όπως τη θερμοκρασία του σώματος, το πόσο φαγητό θα φάει ο άνθρωπος και το «εσωτερικό ρολόι» του οργανισμού.

Το σύνολο του εγκεφάλου χρησιμοποιεί σαν “μαξιλάρι” το εγκεφαλονωτιαίο υγρό (cerebrospinal fluid, CSF) που παρέχει μηχανική αλλά και ανοσολογική προστασία στον εγκέφαλο.



Εικόνα 1: Βασικά Μέρη Εγκεφάλου [3]

1.3 Αλλαγές στον εγκέφαλο από τη φυσιολογική γήρανση

Η φυσιολογική γήρανση του ανθρώπου επιφέρει πολλές αλλαγές στον εγκέφαλο. Αρκετές από αυτές προσομοιάζουν την παθολογία της ΝΑ [4]. Είναι αναγκαίο συνεπώς να εξετάσουμε και να διαχωρίσουμε όσο είναι δυνατό τις δύο περιπτώσεις.

Σε πολλούς ηλικιωμένους ανθρώπους μέρη του εγκεφάλου συρρικνώνονται, ιδιαίτερα ο προμετωπιαίος φλοιός και ο ιππόκαμπος, που ελέγχει την πρόσφατη μνήμη. Επίσης μειώνεται η ικανότητα επικοινωνίας των νευρώνων, καθώς μειώνεται ή υποβαθμίζεται η λευκή ουσία του

εγκεφάλου (που αποτελείται από άξονες μυελίνης). Σε πολλές περιπτώσεις παρατηρείται μειωμένη ροή του αίματος στα αγγεία, ενώ παρατηρούνται και δομές, όπως οι αμυλοειδικές πλάκες, που ενώ είναι χαρακτηριστικά της νευροπαθολογίας της ΝΑ, σε μικρότερο βαθμό παρατηρούνται και σε ανθρώπους χωρίς προβλήματα υγείας.

Τα παραπάνω δεν υποβαθμίζουν απαραίτητα τη νοητική ικανότητα του ανθρώπινου εγκεφάλου. Σχετικές μελέτες έχουν δείξει ότι οι περισσότεροι υγιείς ηλικιωμένοι μπορούν να λύσουν τα ίδια προβλήματα με τους νεότερους, ιδιαίτερα αν τους δοθεί λίγο παραπάνω χρόνος, ακόμα και αν εμφανίζουν κάποια νευροεκφυλιστικά συμπτώματα. Ο εγκέφαλος φαίνεται να μπορεί να προσαρμοστεί στις παραπάνω βλάβες και να τις ξεπεράσει. Ιδιαίτερο αντικείμενο μελέτης είναι εξωτερικοί παράγοντες που εικάζεται ότι βοηθούν τη διαδικασία προσαρμογής του εγκεφάλου, όπως η εκπαίδευση, η οικογενειακή ζωή και η καλή διατροφή.

Από κάποιο σημείο και πέρα πάντως, η παραπάνω ισορροπία διαταράσσεται και ο εγκέφαλος εκφυλίζεται σε κάποια παθολογική κατάσταση. Αυτή είναι και η περίπτωση της ΝΑ.

1.4 Νευροπαθολογία της νόσου του Αλτσχάιμερ

Η νευροπαθολογία της ΝΑ έχει ένα κατεξοχήν χαρακτηριστικό: τη δημιουργία αμυλοειδικών πλακών και νευροϊνδιακών τολυπίων [4], δηλαδή τη συσσώρευση τοξικών επιπέδων αμυλοειδικών πεπτιδίων (tau πρωτεΐνες και αμυλοειδικές βήτα μετά από επεξεργασία από την APP πρωτεΐνη) στον εγκέφαλο που καταστρέφουν τις κοντινές περιοχές: Οι αμυλοειδικές ίνες είναι ιδιαίτερα ανθεκτικές, οπότε η συσσώρευσή τους δεν μπορεί να εμποδισθεί από τον εγκέφαλο. Φθάνουν νευροτοξικά επίπεδα όταν οι πλάκες είναι τόσο μεγάλες ώστε να πιέζουν -και άρα να καταστρέφουν- τις γειτονικές ίνες των νευρώνων. Η καταστροφή περιοχών του εγκεφάλου δημιουργεί όλες τις νοητικές διαταραχές που έχουν συνδεθεί με τη ΝΑ. Σε ακραίες περιπτώσεις οι αμυλοειδικές ίνες συσσωρεύονται γύρω από αγγεία, πράγμα που μπορεί να οδηγήσει στην εκροή αίματος στον εγκέφαλο και σε εγκεφαλικό επεισόδιο (εγκεφαλική αμυλοειδική αγγειοπάθεια).

Άλλα χαρακτηριστικά της παθολογίας της ΝΑ στον εγκέφαλο είναι η τοπική ατροφία (ιδιαίτερα στον ιππόκαμπο και στον κροταφικό λοβό), η απώλεια νευρώνων, οι αλλαγές στη μικροδομή της λευκής ουσίας και η νευρογλοιακή ενεργοποίηση. Υπάρχουν και κάποιες ακόμη αλλαγές που έχουν συνδεθεί με τη ΝΑ, αλλά δεν εμφανίζονται σε όλες τις περιπτώσεις της.

Όλα τα παραπάνω οδηγούν στα κλασικά συμπτώματα της ΝΑ: σε αρχικό ακόμη στάδιο παρατηρούνται συνήθως ελλείμματα στη μνήμη, αλλαγές της διάθεσης, απάθεια και σύγχυση. Σε τελικό στάδιο παρατηρούνται διαταραχές στην ομιλία, απραξία ή ακόμα και εγκεφαλικά επεισόδια, παράλυση περιοχών του προσώπου, και κρίσεις όμοιες με τις επιληπτικές.

1.5 Διάγνωση της νόσου του Αλτσχάιμερ

Όπως προαναφέραμε, η πρόωμη διάγνωση της ΝΑ είναι από τους κύριους στόχους στην προσπάθεια αντιμετώπισης της νόσου. Μέχρι τώρα η διάγνωση της νόσου ταυτίζεται με την επίδοση του ασθενή σε διάφορες νευροψυχολογικές δοκιμασίες. Αυτό οφείλεται στο γεγονός ότι τέτοιες δοκιμασίες είναι πολύ αξιόπιστες και σχετικά φθηνές. Ταυτόχρονα, καθώς η ασθένεια επηρεάζει τον εγκέφαλο, δεν υπάρχει εύκολη πρόσβαση στις περιοχές που θίγονται. Η ανάπτυξη απεικονιστικών τεχνικών ή δεικτών όπως η συγκέντρωση πρωτεϊνών επέτρεψαν τις τελευταίες δύο δεκαετίες να αποτιμηθεί η κατάσταση του εγκεφάλου, αλλά και πάλι οι νευροψυχολογικές δοκιμασίες είναι φθηνότερες και πιο αξιόπιστες. Εδώ να σημειώσουμε ότι η νευροπαθολογία της ΝΑ δεν εμφανίζεται μόνο σε νοσούντες από αυτή: περίπου το 30% των υπερήλικων έχουν νευροπαθολογία που παραπέμπει σε ΝΑ, χωρίς να πάσχουν από κάποια ασθένεια. Προφανώς αυτό δυσκολεύει την ανάπτυξη διαγνωστικών εργαλείων για τη ΝΑ που δεν βασίζονται σε νευροψυχολογικές δοκιμασίες.

Ωστόσο η νευροαπεικόνιση έχει εξελιχθεί τόσο, ώστε τα πρόσφατα πρότυπα διάγνωσης της ΝΑ εισάγουν και τη χρήση απεικονιστικών τεχνικών στη διαδικασία. Η χρήση δομικής απεικόνισης πυρηνικού μαγνητικού συντονισμού (structural magnetic resonance, sMRI) προτείνεται, αν και όχι σαν αποκλειστικό μέσο διάγνωσης. Αντίστοιχα νέες έρευνες εισάγουν την τομογραφία εκπομπής ποζιτρονίων (positron emission tomography, PET) με τον ραδιοϊχνηλάτη florbetapir (florbetapir-PET), κυρίως για το διαχωρισμό διαφόρων μορφών άνοιας [5].

Πάντως εδώ πρέπει να επισημανθεί ότι ο βασικός στόχος στην τρέχουσα έρευνα για τη ΝΑ είναι η επιτυχημένη πρόβλεψη ανάπτυξης της ασθένειας. Σε αυτό το πλαίσιο άλλες τεχνικές πέρα από τις νευροψυχολογικές δοκιμασίες βρίσκουν εκτεταμένη εφαρμογή.

1.6 Ήπια Νοητική Διαταραχή

Η ήπια νοητική διαταραχή (mild cognitive impairment, MCI) είναι ένα σύνδρομο της λειτουργίας του εγκεφάλου, που αφορά την ύπαρξη γνωσιακών περιορισμών που δεν δικαιολογούνται από την ηλικία ή κάποια άλλη ασθένεια. Κεντρικά χαρακτηριστικά του MCI είναι ότι τα συμπτώματά του δεν εμποδίζουν την καθημερινή λειτουργία του ασθενούς, υπάρχει κάποια απώλεια μνήμης και ο ασθενής δεν πάσχει από κάποια άλλη άνοια. Ως σύνδρομο είναι σε πολλές περιπτώσεις ο “προθάλαμος” σοβαρότερων ασθενειών, αν και αυτό δεν είναι απαραίτητο [6]. Εκδηλώνεται με πολλούς διαφορετικούς τρόπους, αλλά μία από τις πλέον κοινές μορφές του είναι το αμνησιακό MCI. Το αμνησιακό MCI θεωρείται ο πρόδρομος της ΝΑ. Αυτός είναι ο λόγος που αποσπά μεγάλο ερευνητικό ενδιαφέρον, καθώς η πρόγνωση της ΝΑ προϋποθέτει τον έγκαιρο εντοπισμό του (όταν οι ασθενείς εμφανίζουν ακόμη συμπτώματα MCI και όχι τα βαρύτερα συμπτώματα της ΝΑ) [7]. Η διάγνωση του MCI γίνεται με τον ίδιο τρόπο με τη ΝΑ, δηλαδή με νευροψυχολογικές δοκιμασίες όπως η βραχεία δοκιμασία ελέγχου της νοητικής κατάστασης. Διαφορετική βαθμολογία σε αυτές τις δοκιμασίες κατατάσσει τους ασθενείς σε πάσχοντες από MCI ή ΝΑ [8].

Ο διαχωρισμός των ασθενών με στάσιμο MCI και αυτών που από MCI θα μεταβούν στη ΝΑ (εξελισσόμενο MCI) είναι ένας από τους σημαντικότερους ερευνητικούς στόχους στην έρευνα για τη ΝΑ και θα συναντηθεί σε πολλές μελέτες. Η παθολογία του MCI παρουσιάζει αρκετά συμπτώματα που υπάρχουν και στη ΝΑ και χρησιμοποιείται

- για τον εντοπισμό των γενεσιουργών αιτιών της ΝΑ
- την παρακολούθηση του τρόπου που εξελίσσεται η ασθένεια
- την ανάπτυξη δεικτών που θα εντοπίσουν τη ΝΑ έγκαιρα, μέσω της πρόβλεψης της εξέλιξης της ασθένειας

Χαρακτηριστικά παραδείγματα βιολογικών συμπτωμάτων που εμφανίζονται στο MCI και εξελίσσονται στη ΝΑ είναι η ανάπτυξη αμυλοειδικών πλακών και νευροϊνιδιακών τολυπιών και η τοπική απώλεια όγκου στον εγκέφαλο.

Όπως και στη ΝΑ, δεν υπάρχει κάποια έγκυρη θεραπεία για το MCI. Ωστόσο πιστεύεται ότι είναι πιθανότερο να βρεθεί κάποια θεραπεία ή μέθοδος πρόληψης στο MCI παρά στη ΝΑ, καθώς η υποβάθμιση του εγκεφάλου δεν είναι ακόμη τόσο εκτεταμένη και άρα πιθανώς αναστρέψιμη. Έτσι η πρόληψη της ΝΑ περιλαμβάνει την αντιμετώπιση του MCI.

1.7 Νευροψυχολογικές δοκιμασίες και δοκιμασίες γνωσιακής λειτουργίας

Η βασική μέθοδος διάγνωσης ΝΑ ή άνοιας είναι η αποτίμηση της νοητικής κατάστασης του ατόμου. Αυτό επιτυγχάνεται με κάποιο από τα τυποποιημένα πρότυπα που απαιτούν τη διεξαγωγή νευροψυχολογικών δοκιμασιών αλλά και την παρακολούθηση του ατόμου από τον ιατρό. Προς το παρόν είναι η μέθοδος με τα μεγαλύτερα ποσοστά επιτυχίας στη διάγνωση της ΝΑ, ενώ παράλληλα έχει μικρό κόστος και δεν επιβαρύνει τον ασθενή. Είναι χαρακτηριστικό ότι πριν την (πρόσφατη) ανάπτυξη εξελιγμένων τεχνικών απεικόνισης του εγκεφάλου, η βέβαιη διάγνωση της ΝΑ γινόταν στο νεκροτομείο, ενώ ακόμη και σήμερα απαιτείται η χρήση ακριβών και περίπλοκων τεχνικών απεικόνισης. Αυτές οι δοκιμασίες συνεπώς διευκόλυναν και διευκολύνουν πολύ το έργο των ιατρών.

Γενικά, οι νευροψυχολογικές δοκιμασίες χρησιμοποιούνται κατά κόρον στην αναζήτηση πιθανών βλαβών σε περιοχές του εγκεφάλου. Οι παράγοντες που συνήθως εξετάζονται είναι η μνήμη, η εκτελεστική λειτουργία, η ευφύια και η δυνατότητα έκφρασης του ατόμου [9][10]. Οι συγκεκριμένες δοκιμασίες χορηγούνται ανά τακτά διαστήματα ακόμη και μετά τη διάγνωση, για να παρακολουθηθεί η εξέλιξη της νόσου αλλά και η επίδραση της φαρμακευτικής προσπάθειας αποκατάστασης. Η δομή που ακολουθεί μία τέτοια δοκιμασία συνήθως περιλαμβάνει την αναζήτηση του ιστορικού του ατόμου και της οικογένειάς του, προκειμένου να διαπιστωθεί αν υπήρχαν αντίστοιχες ενοχλήσεις στο παρελθόν ή έχει κάποια γενετική προδιάθεση. Μετά ακολουθεί μία δομημένη συνέντευξη με συγκεκριμένες, ανοιχτές και μη, ερωτήσεις.

Αυτές τις δοκιμασίες θα μπορούσαμε να τις χωρίσουμε στις γνωσιακές και στις κλινικές. Οι κλινικές απαιτούν την παρακολούθηση από τον ιατρό και συνήθως διαρκούν περισσότερο. Διαφοροποιούνται ελαφρώς στους στόχους τους από τις γνωσιακές δοκιμασίες, γιατί ασχολούνται και με την ικανότητα λειτουργίας του ατόμου, όχι μόνο το βαθμό άνοιας. Η πιο γνωστή κλινική δοκιμασία για τη ΝΑ είναι η κλινική εκτίμηση της άνοιας (Clinical Dementia Rating, CDR). Έχει δείχθει αξιόπιστη, αλλά βασίζεται σε αρκετά μεγάλο βαθμό στην προσωπική κρίση του ιατρού. Αυτό προφανώς σημαίνει ότι η διεξαγωγή της απαιτεί εξειδικευμένο προσωπικό και άρα δεν χρησιμοποιείται συχνά σε ημιαντοματοποιημένες μεθόδους, όπως αυτές που θα εξετάσουμε.

Αντίθετα οι γνωσιακές δοκιμασίες συνήθως είναι πιο τυποποιημένες. Τέτοια πρότυπα είναι το βραχύ ερωτηματολόγιο νοητικής κατάστασης (Short Portable Mental Status Questionnaire), η κλίμακα νευροψυχολογικής διαταραχής (Neuropsychological Impairment Scale), η αξιολόγηση του ασθενούς για τη νοητική του λειτουργία (Patient's Assessment of Own Functioning) ή η δομημένη συνέντευξη για τη διάγνωση της άνοιας (Structured Interview for the Diagnosis of Dementia). Οι πιο γνωστές

τυποποιημένες δοκιμασίες είναι ίσως η βραχεία δοκιμασία ελέγχου της νοητικής κατάστασης (mini-mental state examination, MMSE) και η γνωσιακή αξιολόγηση του Μόντρεαλ (Montreal Cognitive Assessment, MoCA - χρησιμοποιείται ευρύτατα για τη διάγνωση MCI).

Συνολικά οι παραπάνω δοκιμασίες είναι εξαιρετικά χρήσιμες. Το CDR είναι ένα κλινικό εργαλείο και είναι χρήσιμο για να εκτιμηθεί αν το άτομο χρειάζεται υποστήριξη, ενώ οι δοκιμασίες τύπου MMSE είναι χρήσιμες στον εντοπισμό της άνοιας. Οι εκτιμήσεις τους σε σχετικές έρευνες έχουν δείξει να διαφέρουν αρκετά, με τις νοητικές δοκιμασίες να αποτιμούν την κατάσταση του ατόμου ως χειρότερη [11]. Στις ακόλουθες ερευνητικές εργασίες θα συναντήσουμε κατεξοχήν το MMSE, αφενός γιατί είναι πιο τυποποιημένο και λιγότερο χρονοβόρο, αφετέρου γιατί η έρευνα της ΝΑ αφορά κυρίως στην έγκαιρη διάγνωση της άνοιας. Ακολουθεί μία MMSE δοκιμασία [12].

MINI MENTAL STATE EXAMINATION (MMSE)

Σκορ:

/30

(1 για σωστό, 0 για λάθος)

Θα σας κάνω μερικές απλές ερωτήσεις, για να δούμε πως πάει η μνήμη σας. Μερικές είναι πολύ απλές, μερικές δυσκολότερες. Σας παρακαλώ να απαντήσετε όσο μπορείτε καλύτερα.

Μπορείτε να μου πείτε

1. Τι έτος έχουμε
2. Τι εποχή
3. Τι μήνα
4. Τι ημερομηνία (πόσο έχει ο
5. Τι μέρα της εβδομάδας

Τώρα θα σας ρωτήσω για το μέρος στο οποίο

6. Πως ονομάζεται το νοσοκομείο
7. Σε ποια πόλη βρίσκεται;
8. Σε ποιο
9. Σε ποια;
10. Σε ποιον όροφο;

Θα σας πω τρεις λέξεις: παρακαλώ να μου τις επαναλάβετε και να τις θυμάστε γιατί από λίγο θα τις ρωτήσω πάλι:

11. Καρέκλα
12. Φούστα
13. Ποδήλατο

Μπορείτε να κάνετε συνεχείς αφαιρέσεις το 7 ξεκινώντας από το 100

14. 93
15. 86
16. 79
17. 72
18. 65

Μπορείτε να επαναλάβετε τις τρεις λέξεις που σας είπα

19. Καρέκλα
20. Φοίετα
21. Ποδήλατο
Δείχνουμε ένα μολύβι

22. Τι είναι αυτό;

Δείχνουμε ένα ρολόι

23. Τι είναι αυτό;

Ακούστε με προσοχή και επαναλάβετε μετά από μένα:

24. «Παιδί, κλειδί, κλαδί»

Βάζουμε ένα φύλλο χαρτί πάνω στο γραφείο, το δείχνουμε στον ασθενή και του λέμε:
Παρακαλώ ακούστε με προσεχτικά και κάντε το εξής:

25. Πάρτε αυτό το χαρτί με το δεξί χέρι

26. Διπλώστε το στα δύο

27. Και ακουμπήστε το στο πάτωμα

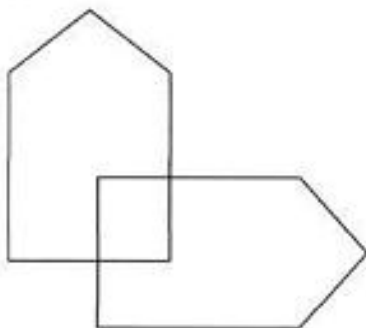
Δείχνουμε στον ασθενή ένα χαρτί που γράφει με κεφαλαία: ΚΛΕΙΣΤΕ ΤΑ ΜΑΤΙΑ

28. Παρακαλώ κάντε ότι γράφει το εδώ

Δίνουμε στον ασθενή ένα μολύβι και ένα χαρτί:

29. Θα ήθελα να μου γράψετε μια πρόταση, όποια θέλετε, αλλά να είναι ολοκληρωμένη (δεν λαμβάνουμε υπόψιν συντακτικά ή ορθογραφικά λάθη, παρά μόνο αν η πρόταση έχει υποκείμενο και ρήμα. Παρέχουμε 30")

30. Παρακαλώ αντιγράψτε μου αυτό το σχήμα.
(Πρέπει να διακρίνονται οι δέκα γωνίες εκ των οποίων οι δύο να τέμνονται. Ο τρόπος αγνοείται)



Η αποτίμηση της νοητικής κατάστασης του ατόμου συνήθως διαρκεί λιγότερο από 12 ώρες. Εννοείται ότι αυτό δεν περιλαμβάνει το χρόνο που καταναλώνει ο ιατρός για να εκτιμήσει τα δεδομένα και να καταλήξει σε διάγνωση. Η παραπάνω διαδικασία μπορεί να διαγνώσει με ακρίβεια, πέρα από τη MCI, και άλλες νευροψυχολογικές ασθένειες.

1.8 Γιατί οι νευροψυχολογικές δοκιμασίες δεν αρκούν

Το βασικό πρόβλημα με αυτήν τη μέθοδο είναι ότι δεν μπορεί να προβλέψει την ανάπτυξη της ασθένειας. Όταν η νόσος έχει προχωρήσει τόσο, ώστε να διαγνωστεί ο ασθενής με άνοια, έχουν ήδη συμβεί ανεπανόρθωτες βλάβες στον ανθρώπινο εγκέφαλο. Είναι λοιπόν κρίσιμο η πιθανή εξέλιξη της νόσου να προβλεφθεί το νωρίτερο δυνατό [13]. Το ρόλο αυτό τον έχουν αναλάβει διάφοροι βιοδείκτες. Καθώς η αιτία της νόσου δεν είναι ξεκάθαρη και πιστεύεται ότι οφείλεται σε πλήθος αλληλεπιδρώντων παραγόντων, το έργο αυτό είναι ιδιαίτερα δύσκολο. Ωστόσο κάποιοι βιοδείκτες έχουν προταθεί με σχετική επιτυχία. Αυτοί περιλαμβάνουν αλλαγές στη δομή του εγκεφάλου (πχ όγκος ιππόκαμπου ή κροταφικού λοβού), την εμφάνιση αμυλοειδικών πλακών, καθώς και αλλαγές στο μεταβολισμό του εγκεφάλου, όπως η συγκέντρωση συγκεκριμένων πρωτεϊνών. Η ανάπτυξη ενός διαγνωστικού εργαλείου, που θα περιλαμβάνει εξετάσεις για τους ανωτέρω βιοδείκτες και θα προβλέπει με ακρίβεια την πιθανή ανάπτυξη και εξέλιξη της νόσου, είναι ο κατεξοχήν στόχος της σύγχρονης προσπάθειας για τη ΝΑ.

1.9 Βιοδείκτες της νόσου του Αλτσχάιμερ: Πρώιμη ανίχνευση της νόσου

Η έρευνα για τον εντοπισμό προκλινικών συμπτωμάτων της ΝΑ βασίζεται στους βιοδείκτες της ασθένειας. Ως βιοδείκτη θα ορίζαμε μία μονάδα που η ύπαρξή της, η συγκέντρωσή της ή η δραστηριότητά της αποτελούν ένδειξη για μία ασθένεια. Τα χαρακτηριστικά που απαιτούνται από έναν ιδανικό βιοδείκτη είναι να είναι συγκεκριμένος, να προβλέπει την εξέλιξη της ασθένειας και να ανακλά την αντίδραση στη θεραπεία με ακρίβεια. Η έρευνα πάνω στους βιοδείκτες για τη ΝΑ είναι εκτεταμένη εξαιτίας των ιδιαιτεροτήτων της ασθένειας, δηλαδή της μη αποκλειστικής νευροπαθολογίας της. Οι πιο κοινοί βιοδείκτες της ΝΑ είναι αυτοί που προκύπτουν από απεικονιστικές τεχνικές, από την εξέταση του γονιδιώματος και από δείκτες του εγκεφαλονωτιαίου υγρού και του πλάσματος. Και άλλοι βιοδείκτες έχουν προταθεί που παρακολουθούν την φλεγμονή του εγκεφάλου ή την οξειδωσή του αλλά χωρίς ενθαρρυντικά αποτελέσματα.

1.9.1 Νευροαπεικονιστικοί βιοδείκτες

Οι δείκτες αυτοί βασίζονται σε κάποια από τις απεικονιστικές τεχνικές (πυρηνικός μαγνητικός συντονισμός, τομογραφία εκπομπής ποζιτρονίων, ηλεκτροεγκεφαλογράφημα). Κατεξοχήν ασχολούνται με τη δομή του εγκεφάλου με πιο διαδεδομένους δείκτες αυτούς που απορρέουν από την τεχνική του sMRI. Ωστόσο η μελέτη της φυσιολογίας του εγκεφάλου μπορεί να εξεταστεί με απεικονιστικές τεχνικές και η τεχνική του λειτουργικού πυρηνικού μαγνητικού συντονισμού (functional MRI, fMRI)

για παράδειγμα γίνεται όλο πιο διαδεδομένη. Η συμβολή τους στη ΝΑ είναι αναμφισβήτητη. Ένα απλό παράδειγμα είναι η διάγνωση της ΝΑ αντί για άλλες μορφές άνοιας όπως η μετωποκροταφική άνοια (frontotemporal dementia, FTD) και η άνοια με σωμάτια Lewy (dementia with Lewy bodies, DLB), από τα διαφορετικά μοτίβα απώλειας όγκου στον εγκέφαλο που δείχνει η MRI. Οι δείκτες αυτοί θα εξεταστούν αναλυτικότερα σε επόμενο κεφάλαιο.

1.9.2 Εγκεφαλονωτιαίο υγρό και πλάσμα

Όπως έχει ήδη αναφερθεί ένα από τα βασικά συμπτώματα της ΝΑ είναι η ύπαρξη νευροϊνιδιακών τολυπίων από υπερφωσφορυλιωμένη (hyperphosphorylated) tau πρωτεΐνη και αμυλοειδικών πλακών από Αβ40 και Αβ42 (αμυλοειδική βήτα πρωτεΐνη). Στην παρακολούθηση αυτών των συμπτωμάτων είναι σημαντική η συμβολή των βιοδεικτών του εγκεφαλονωτιαίου υγρού (cerebrospinal fluid , CSF) και του πλάσματος (εδώ να σημειώσουμε βέβαια ότι νέοι ακτινολογικοί ιχνηλάτες επιτρέπουν την αποτελεσματική παρακολούθηση του μεταβολισμού της tau πρωτεΐνης και από την PET). Το CSF έρχεται σε άμεση επαφή με τον εξωκυτταρικό χώρο του εγκεφάλου και έτσι οι δείκτες του ανακλούν καλύτερα τις βιοχημικές αλλαγές που συμβαίνουν στον εγκέφαλο σε σχέση για παράδειγμα με το πλάσμα [14]. Επίσης, η συλλογή και η αξιολόγηση του CSF είναι απλούστερη και φθηνότερη από τις περισσότερες απεικονιστικές μεθόδους, αν και επίπονη για τους ασθενείς (στις κύριες βάσεις δεδομένων για τη ΝΑ λείπουν συχνά στοιχεία CSF). Υπάρχουν τρεις κύριοι βιοδείκτες CSF με διαγνωστική αξία για τη ΝΑ: ο δείκτης που μετρά τη συγκέντρωση της Αβ42, ο δείκτης για την ολική-tau (total-tau, t-tau) και ο δείκτης φωσφορυλιωμένης-tau (phosphorylated-tau, p-tau). Ο δείκτης για την Αβ42 μόνος του δεν φαίνεται να είναι ιδιαίτερα ευαίσθητος στη ΝΑ. Οι δείκτες για την tau είναι αποτελεσματικότεροι, καθώς, όταν διαλύονται νευριτικές δομές, απελευθερώνεται στο CSF tau πρωτεΐνη. Αν και ο δείκτης για την t-tau είναι ευαίσθητος στην ανάπτυξη άνοιας και συνεπώς μπορεί να χρησιμοποιηθεί για διαγνωστικούς σκοπούς, δεν διακρίνει με ακρίβεια ανάμεσα στις διαφορετικές μορφές άνοιας. Ο p-tau δείκτης είναι ο πιο ελπιδοφόρος, καθώς φέρεται να έχει διαφορετική συγκέντρωση σε ασθενείς με ΝΑ και σε ασθενείς με κάποια άλλη άνοια, δείχνοντας ότι η p-tau δεν έχει μόνο νευροεκφυλιστικό ρόλο στη ΝΑ, αλλά συμμετέχει στην εξέλιξη της νόσου. Ο συνδυασμός των παραπάνω δεικτών έχει μεγαλύτερη διαγνωστική ισχύ από κάθε έναν μόνο του και έτσι θα συναντηθούν συνήθως.

Αντίθετα, οι δείκτες για το πλάσμα έχουν μικρότερη αξία. Παρότι εξακολουθούν να χρησιμοποιούνται σε αρκετές μελέτες, γιατί η εξαγωγή και η εξέταση είναι εύκολη, δεν έχει βρεθεί κάποιος ιδιαίτερα αποτελεσματικός βιοδείκτης. Υπάρχουν πρωτεΐνες του πλάσματος που έχουν βρεθεί σε διαφορετικά επίπεδα στους ασθενείς με ΝΑ από τους υγιείς, αλλά δεν έχουν χρησιμοποιηθεί εκτεταμένα μέχρι τώρα γιατί τα αποτελέσματα δεν έχουν επαναληφθεί. Βέβαια η έρευνα μέχρι τώρα

είχε επικεντρωθεί στην Αβ πρωτεΐνη που μπορεί μεν να ανιχνευθεί στο πλάσμα, αλλά εξαιτίας της μεγάλης διάχυσης και της υδροφοβικής φύσης της δεν είναι πολύ κατάλληλη. Υπάρχουν, όμως, πρωτεΐνες που αφορούν τις φλεγμονές που πιθανόν να είναι πιο κατάλληλες ως υποψήφιοι βιοδείκτες.

1.9.3 Γενετικοί βιοδείκτες

Η αποκρυπτογράφηση του ανθρώπινου γονιδιώματος προσέφερε νέες προοπτικές στη διάγνωση της νόσου. Αν και είναι σαφές πλέον ότι δεν υπάρχει κάποιο γονίδιο που να είναι αποκλειστικά υπεύθυνο για την εξέλιξη της νόσου, έχει καταστεί εξίσου σαφές ότι γενετικοί παράγοντες επηρεάζουν την εμφάνιση της νόσου. Συγκεκριμένοι σημειακοί πολυμορφισμοί (Single-Nucleotide Polymorphisms, SNPs) γονιδίων έχουν συνδεθεί στατιστικά με την εμφάνιση της νόσου. Τα SNPs είναι παραλλαγές σε μια βασική ακολουθία του DNA, όπου ένα συγκεκριμένο τμήμα του γονιδιώματος ενός οργανισμού διαφοροποιείται από το αντίστοιχο τμήμα ενός άλλου κατά ένα μόνο νουκλεοτίδιο. Η άλλη παραλλαγή που θα συναντηθεί είναι ο αριθμός αντιγράφων DNA (Copy Number Variations, CNVs) που εκφράζει δομικές παραλλαγές στο γονιδίωμα, όπου τμήματα του γενότυπου υπάρχουν σε μη κανονικό αριθμό αντιγράφων, δηλαδή το ίδιο τμήμα επαναλαμβάνεται ή διαγράφεται. Παράλληλα γίνονται πολλές προσπάθειες στην ταύτιση SNPs και CNVs με μορφολογίες και λειτουργίες του εγκεφάλου. Αυτό το επιστημονικό πεδίο συνδυάζει τη γενετική ανάλυση με απεικονιστικές κατεξοχήν τεχνικές, προκειμένου να αποκρυπτογραφηθεί η γενετική αιτιολογία της ΝΑ και μέχρι τώρα έχει αποδώσει ενθαρρυντικά αποτελέσματα. Το τεράστιο μέγεθος δεδομένων και οι εξελιγμένες στατιστικές μέθοδοι που απαιτούνται καθιστούσαν απαγορευτική μέχρι πρόσφατα την εκταταμένη έρευνα στο πεδίο, ενώ παράλληλα οι περιβαλλοντικοί παράγοντες περιπλέκουν ακόμη περισσότερο την έρευνα. Ωστόσο η συνδυαστική έρευνα του ανθρώπινου γονιδιώματος και της δομής και των λειτουργιών του ανθρώπινου εγκεφάλου έχει τη δυναμική να δια φωτίσει την αιτιολογία της νόσου.

1.9.4 Συνδυαστικοί βιοδείκτες

Αν και οι γενετικοί δείκτες είναι ένα ενδιαφέρον εργαλείο στην έρευνα για τη ΝΑ, για να εξηγηθεί η περίπλοκη παθολογία της ασθένειας πρέπει να ληφθούν υπόψη και άλλοι παράγοντες, δηλαδή περιβαλλοντικοί παράγοντες και επιδημιολογικά χαρακτηριστικά. Παράλληλα, καθαρά ως υποψήφιο διαγνωστικό εργαλείο, οι γενετικοί δείκτες δεν αρκούν, εκτός αν αφορούν τετριμμένες μορφές της ΝΑ (ΝΑ πρώιμης εκδήλωσης). Ο συνδυασμός όλων των παραπάνω δεικτών για τη ΝΑ προβλέπει με μεγαλύτερη ακρίβεια την ανάπτυξη της ασθένειας. Ωστόσο η ανάπτυξη ενός "παγκόσμιου" δείκτη για τη ΝΑ, που θα περιλαμβάνει όλη τη διαθέσιμη πληροφορία, είναι ακόμη ένα ανοιχτό ερευνητικό πεδίο.

Στη χρήση και στο συνδυασμό των διάφορων δεικτών για τη ΝΑ, τεράστια ώθηση έχει δώσει η πρωτοβουλία νευροαπεικόνισης της ΝΑ.

1.10 Πρωτοβουλία νευροαπεικόνισης της νόσου του Αλτσχάιμερ

Η πρωτοβουλία νευροαπεικόνισης της ΝΑ (Alzheimer's Disease Neuroimaging Initiative, ADNI) είναι μία από τις πιο δραστήριες οργανώσεις στην έρευνα για τη ΝΑ. Ιδρύθηκε το 2003 από τις National Institute of Aging, National Institute of Biomedical Imaging and Bioengineering, Food and Drug Administration και διάφορες φαρμακευτικές εταιρίες και μη κερδοσκοπικούς οργανισμούς. Ο κύριος στόχος της ADNI είναι να διαπιστώσει κατά πόσον οι απεικονίσεις MRI, PET και άλλοι δείκτες μπορούν να αποκρυπτογραφήσουν την εξέλιξη της ΝΑ και του MCI [15]. Για το σκοπό αυτό προσέλαβαν πάνω από 800 άτομα από 55 έως 90 ετών με ή χωρίς ΝΑ και MCI και τους υπέβαλαν σε εξετάσεις MRI, PET, νευροψυχολογικές δοκιμασίες και άλλα ανά τακτά χρονικά διαστήματα. Η βάση δεδομένων της ADNI είναι από τις μεγαλύτερες και πιο καλά ορισμένες στο χώρο της ΝΑ, με ασθενείς από περισσότερα από 50 διαφορετικά ιδρύματα στη Βόρεια Αμερική. Ως εκ τούτου έχει χρησιμοποιηθεί στις περισσότερες πρόσφατες έρευνες που γίνονται πάνω στη νόσο.

Ο λόγος που η βάση δεδομένων της ADNI θεωρείται πρότυπη όσον αφορά τη ΝΑ, είναι ότι οι συμμετέχοντες υποβλήθηκαν σε μία σειρά επαναλαμβανόμενων εξετάσεων σε διάστημα 5ετίας. Έτσι πολλοί ασθενείς προχώρησαν για παράδειγμα από MCI σε ΝΑ και η περίπτωσή τους μπορεί να αξιοποιηθεί για την παρακολούθηση της εξέλιξης και την πρόγνωση της νόσου. Πιο συγκεκριμένα, όλα τα διαγνωστικά συστήματα που θα εξετάσουμε πρακτικά απαιτούν την ταξινόμηση των δειγμάτων σε τρεις ή τέσσερις ομάδες: τους πάσχοντες από ΝΑ, τους πάσχοντες από MCI αλλά με σταθερή προοπτική (στάσιμο MCI), τους πάσχοντες από MCI αλλά με προοπτική ανάπτυξης ΝΑ (εξελισσόμενο MCI) και την ομάδα ελέγχου, τους υγιείς. Η ύπαρξη μιας βάσης δεδομένων, που καλύπτει κάθε μία από τις παραπάνω ομάδες με τους περισσότερους από τους έγκυρους βιοδείκτες για τη ΝΑ, έλυσε το πρόβλημα του δείγματος για τους ερευνητές. Η επιτυχία της οργάνωσης ήταν τέτοια που συνεχίζει το έργο της παρότι πέρασε η προβλεπόμενη πενταετία λειτουργίας της, συμπεριλαμβάνοντας μάλιστα και άλλους βιοδείκτες στις πιο πρόσφατες εργασίες της.

ΚΕΦΑΛΑΙΟ 2: ΕΦΑΡΜΟΓΕΣ ΝΕΥΡΟΑΠΕΙΚΟΝΙΣΤΙΚΩΝ ΒΙΟΔΕΙΚΤΩΝ

2.1 Απεικονιστικοί βιοδείκτες

Όπως είναι γνωστό η απεικόνιση εσωτερικών μερών του σώματος έφερε επανάσταση στην ιατρική. Η σύγχρονη ιατρική επιστήμη δεν θα υπήρχε χωρίς τη συμβολή των διαφόρων μεθόδων απεικόνισης σε όλους τους τομείς - πρόληψη, διάγνωση και θεραπεία.

Ανάλογη είναι και η περίπτωση της ΝΑ. Οι MRI εξετάσεις προτείνονται από τα νεώτερα πρότυπα για τη διάγνωση της πορείας της νόσου, πέρα από τις παραδοσιακές νευροψυχολογικές εξετάσεις. Παράλληλα όμως συνιστούν και έναν από τους εγκυρότερους βιοδείκτες για την πρόβλεψη της πορείας της νόσου και, ως εκ τούτου, πρωτοστατούν στην ιατρική έρευνα. Όμοια είναι και η συμβολή της PET, ενώ η fMRI επιτρέπει τη μελέτη των αλλαγών που προκαλεί η ΝΑ στις διαδικασίες του εγκεφάλου. Συμβολή έχει και το ηλεκτροεγκεφαλογράφημα στον τρόπο ενεργοποίησης των περιοχών του εγκεφάλου.

Τα παραπάνω θα μελετηθούν συνοπτικά στο παρόν κεφάλαιο, ενώ η χρήση τους θα συναντηθεί σχεδόν στο σύνολο των ερευνών που θα εξεταστούν. Τέλος, η συνδυασμένη χρήση τους ως εργαλείου πρόγνωσης της ασθένειας θα ερευνηθεί στο κεφάλαιο που ασχολείται με τα συστήματα απόφασης για τη ΝΑ.

2.1.1 Απεικόνιση πυρηνικού μαγνητικού συντονισμού

Απεικόνιση πυρηνικού μαγνητικού συντονισμού (Magnetic Resonance Imaging, MRI) είναι μια ιατρική απεικονιστική τεχνική που χρησιμοποιείται κυρίως για την αναπαράσταση των εσωτερικών δομών του σώματος. Η τεχνική χρησιμοποιεί τον πυρηνικό μαγνητικό συντονισμό για να αποτυπώσει τους πυρήνες των ατόμων μέσα στο σώμα.

Ένας MRI σαρωτής είναι στην πράξη ένας μεγάλος μαγνήτης μέσα στον οποίο τοποθετείται ο ασθενής και το μαγνητικό πεδίο μαγνητίζει τους ατομικούς πυρήνες. Η κατεύθυνση της μαγνήτισης μεταβάλλεται μέσω της εφαρμογής μαγνητικών πεδίων και έτσι οι πυρήνες παράγουν ένα περιστρεφόμενο πεδίο που καταγράφεται από το σαρωτή. Μέσω της τροποποίησης των επιβαλλόμενων μαγνητικών πεδίων, οι πυρήνες σε διαφορετικές περιοχές μετακινούνται με άλλες ταχύτητες. Έτσι εξάγεται χωρική πληροφορία με την κατάλληλη επεξεργασία του σήματος (πχ με αντιστροφή του

μετασχηματισμού φουριέρ (Fourier). Επίσης μπορεί να εξαχθεί τρισδιάστατη εικόνα με την προσθήκη και άλλων μαγνητικών πεδίων.

Ένα μεγάλο πλεονέκτημα της MRI είναι ότι δεν χρησιμοποιεί ακτινοβολία ιονισμού σε αντίθεση με την αξονική τομογραφία ή την ακτινογραφία και άρα δεν θεωρείται επιβλαβής για τον ασθενή. Επίσης, η MRI τεχνική παράγει εικόνες με καλή αντίθεση όσον αφορά τους μαλακούς ιστούς του σώματος, σε σχέση για παράδειγμα με τις ακτινογραφίες, επιτρέποντας έτσι την ευκολότερη καταγραφή του εγκεφάλου. Γι' αυτό η MRI χρησιμοποιείται εκτεταμένα στην αναγνώριση βλαβών του εγκεφάλου, όπως στο διαχωρισμό ενός όγκου από τον υπόλοιπο ιστό.

Οι εικόνες της MRI χαρακτηρίζονται από δύο σταθερές: το χρόνο χαλάρωσης spin-πλέγματος (spin-lattice relaxation time) που αντιπροσωπεύεται από μία σταθερά με την ονομασία T1, και το χρόνο χαλάρωσης spin-spin (spin-spin relaxation time) που αντιπροσωπεύεται από μία σταθερά με την ονομασία T2. Η χαλάρωση spin-πλέγματος αφορά το μηχανισμό με τον οποίο ο z-παράγοντας του πίνακα μαγνήτισης έρχεται σε θερμοδυναμική ισορροπία με το περιβάλλον του (το πλέγμα) και ο T1 χρόνος είναι ο ρυθμός με τον οποίο γίνεται αυτή η διαδικασία. Έτσι ο T1 πρακτικά είναι ο χρόνος που ένας πυρήνας θα μείνει σε μία ανώτερη κατάσταση λειτουργίας και διαφοροποιείται για διαφορετικά υλικά. Αυτός είναι ο τρόπος που παράγεται η απαραίτητη αντίθεση στις εικόνες MRI στις T1-σταθμισμένες εικόνες. Ο χρόνος T1 διαφοροποιείται αρκετά ανάμεσα στην φαιά ουσία και στη λευκή ουσία, και άρα είναι η καταλληλότερη μέθοδος για την απεικόνιση του εγκεφάλου. Οι εργασίες που θα μελετηθούν εδώ χρησιμοποιούν κυρίως T1 σταθμισμένες εικόνες.

Αρκετά κοινές είναι και οι T2 σταθμισμένες εικόνες που αφορούν τον spin-spin χρόνο χαλάρωσης και χρησιμοποιούνται κατεξοχήν για την αποτύπωση οιδημάτων. Παράλληλα όμως είναι κατάλληλες και για τη μελέτη του εγκεφαλονωτιαίου υγρού (CSF) και σε αυτό το πλαίσιο θα τις συναντήσουμε εδώ.

Οι δύο κύριες κατηγορίες MRI που αφορούν τη μελέτη της ΝΑ είναι η δομική MRI (structural MRI, sMRI) και η λειτουργική MRI (functional MRI, fMRI). Η sMRI είναι η κλασική μορφή της MRI όπως ορίστηκε πιο πάνω. Βασικός της στόχος είναι η αποκάλυψη της δομής του εγκεφάλου και είτε αποτυπώνει μία δισδιάστατη τομή του εγκεφάλου είτε την τρισδιάστατη απεικόνισή της μέσω της επεξεργασίας μιας σειράς δισδιάστατων τομών.

Δομική MRI

Η δομική MRI αποκαλύπτει την ανατομία



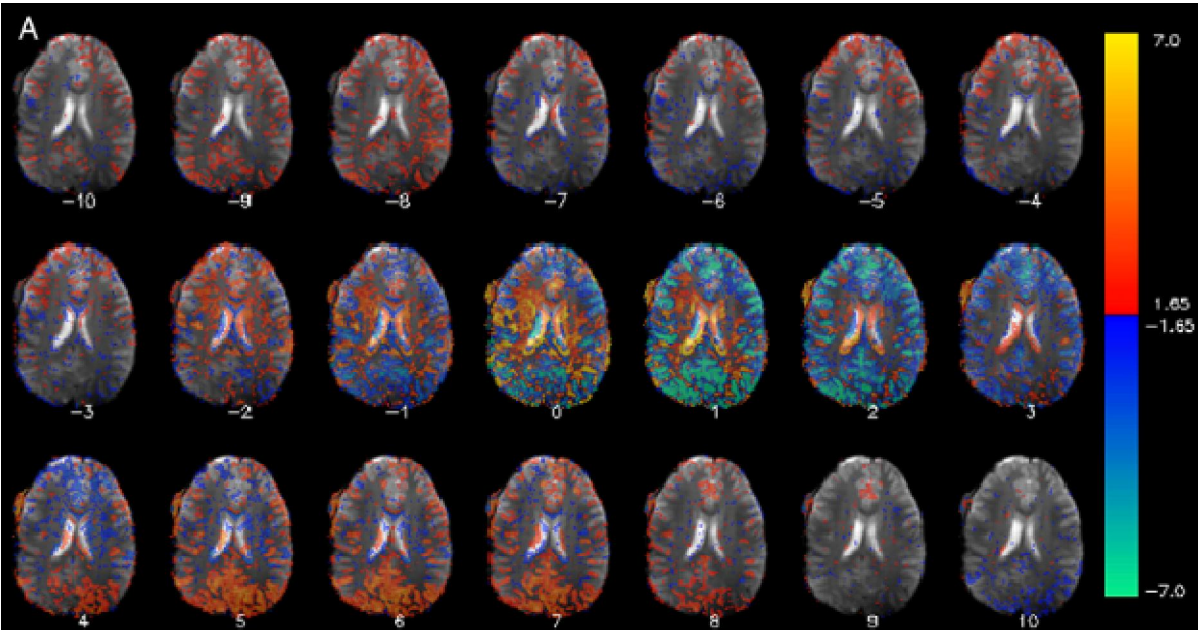
Λειτουργική MRI

Η λειτουργική MRI τη λειτουργία



Εικόνα 2: Διαφορά μεταξύ των *sMRI* και *fMRI* [16]

Η *fMRI* αντίθετα μελετά τη λειτουργία του εγκεφάλου. Χρησιμοποιεί τον κλασικό σαρωτή της *MRI* με την προσθήκη όμως ειδικού λογισμικού και υλικού. Αφορά μια σειρά εικόνων *MRI* που έχουν ληφθεί σε κάποιο χρονικό διάστημα. Η *fMRI* καταγράφει την ενεργοποίηση του νευρικού συστήματος με την ακόλουθη διαδικασία: το δείγμα εκτίθεται σε κάποιο ερέθισμα που ενεργοποιεί μία περιοχή του εγκεφάλου. Η τοπική αύξηση της δραστηριότητας των νευρώνων οδηγεί στη χαλάρωση των μικρών αρτηριών και στην εισροή επιπλέον αίματος σε αυτή την περιοχή. Η διαφορά στη μαγνήτιση ανάμεσα στο πλούσιο σε οξυγόνο αίμα και σε αίμα με λιγότερο οξυγόνο είναι το σήμα που καταγράφεται. Έτσι ταυτίζονται οι περιοχές του εγκεφάλου με τις λειτουργίες του και εξετάζεται η ανθρώπινη νόηση.



Εικόνα 3: Δείγμα απεικόνισης fMRI που αποτυπώνει τη διαφοροποίηση της αιμάτωσης του εγκεφάλου [17]

Και οι δύο παραπάνω προσεγγίσεις χρησιμοποιούνται στην αντιμετώπιση της ΝΑ. Η sMRI αποτελεί έναν καθιερωμένο δείκτη για την πορεία της νόσου και είναι ενταγμένη σε πολλά πρότυπα διάγνωσης της ΝΑ. Η fMRI παρουσιάζει ερευνητικό ενδιαφέρον καθώς μπορεί να αξιολογήσει τον τρόπο που η νόσος επηρεάζει τις λειτουργίες του εγκεφάλου. Παραδείγματα χρήσης και των δύο θα αναφερθούν στην πορεία.

2.1.2 Τομογραφία εκπομπής ποζιτρονίων

Η τομογραφία εκπομπής ποζιτρονίων (PET) είναι μία ακόμα απεικονιστική τεχνική πυρηνικής ιατρικής. Παράγει μία τρισδιάστατη εικόνα των διαδικασιών του σώματος. Το σύστημα λειτουργεί με την καταγραφή των ακτίνων γάμμα που εκπέμπονται εμμέσως από την εκπομπή ποζιτρονίων ενός ραδιοουκλιδίου (tracer), που έχει εισαχθεί νωρίτερα στο σώμα πάνω σε ένα βιολογικά ενεργό μόριο. Η τρισδιάστατη εικόνα παράγεται έπειτα από υπολογιστική επεξεργασία της συγκέντρωσης του ραδιοουκλιδίου στο σώμα. Η εικόνα PET περιέχει πληροφορία κυρίως για τη μεταβολική διαδικασία του σώματος και πολλές φορές συνδυάζεται με άλλες απεικονιστικές τεχνικές που περιλαμβάνουν ανατομική πληροφορία.

Η φθοριοδεοξυγλυκόζη-PET είναι η PET εξέταση όπου το βιολογικά ενεργό μόριο που εισάγεται στο σώμα είναι ένα ανάλογο της γλυκόζης (fluorodeoxyglucose, FDG). Αποτελεί μια πολύ συνηθισμένη περίπτωση PET εξέτασης καθώς χρησιμοποιείται για την εύρεση πιθανών καρκινικών μεταστάσεων.

Προκειμένου να μειωθεί η ακτινοβολία που απορροφά ο ασθενής χρησιμοποιούνται ραδιονουκλίδια με μικρό χρόνο ημιζωής, που προϋποθέτει όμως ότι θα χορηγηθούν στον εξεταζόμενο γρήγορα. Αυτό αυξάνει το κόστος της PET εξέτασης και προϋποθέτει τον αυστηρό προγραμματισμό των εξετάσεων, αλλά η χρησιμότητά της στη διάγνωση των μεταστάσεων του καρκίνου και στην παρακολούθηση της πορείας των καρκινοπαθών είναι τόσο μεγάλη που ξεπερνούν τα όποια μειονεκτήματα.

Στην περίπτωση της ΝΑ, η FDG-PET αποτελεί έναν διαδεδομένο βιοδείκτη δεδομένου ότι η ασθένεια μειώνει κατακόρυφα το μεταβολισμό της γλυκόζης στον εγκέφαλο, σε αντίθεση με άλλες μορφές άνοιας. Έτσι η FDG-PET είναι ένα ενδιαφέρον διαγνωστικό εργαλείο, ενώ χρησιμοποιείται εκτεταμένα και σε επίπεδο έρευνας στην αναγνώριση της σχέσης διαφόρων νευροψυχολογικών λειτουργιών ή διαταραχών και στη δραστηριότητα του εγκεφάλου.

Πρόσφατα (εντός του 2012) εγκρίθηκε από την Food and Drug Administration στις Ηνωμένες Πολιτείες η χρήση ενός νέου ραδιονουκλιδίου για την PET, του florbetapir F-18. Αυτός ο ιχνηλάτης συνδέεται στις αμυλοειδικές δομές του εγκεφάλου, επιτρέποντας έτσι την εκτίμηση της πυκνότητας των β-αμυλοειδικών πλακών. Δεδομένου ότι έχει και τα υπόλοιπα πλεονεκτήματα της PET, η florbetapir-PET έχει μεγάλο ερευνητικό ενδιαφέρον. Είναι ένας υποψήφιος αρνητικός δείκτης για τη ΝΑ (με την έννοια ότι αν δεν υπάρχουν αμυλοειδικές πλάκες, πιθανότατα δεν υπάρχει ΝΑ), ενώ η προσφορά του σε μελέτες δεν έχει αποτιμηθεί ακόμη. Μερικές από τις λίγες, προς το παρόν, μελέτες που αξιοποιούν το συγκεκριμένο δείκτη θα δούμε στην πορεία.

2.1.3 Ηλεκτροεγκεφαλογράφημα

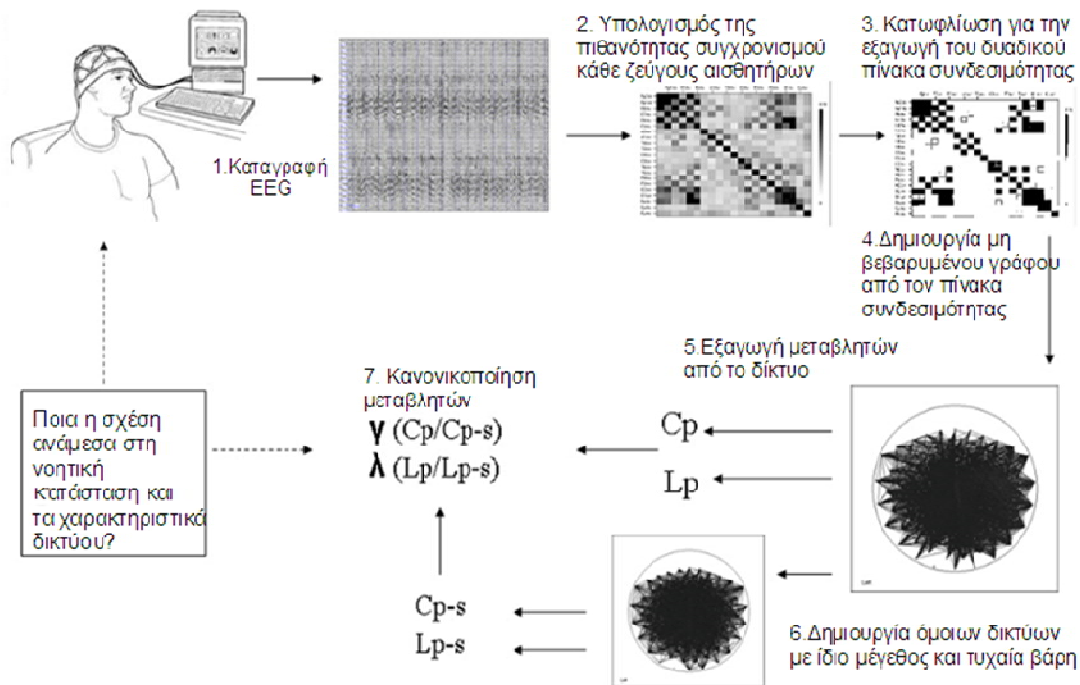
Το ηλεκτροεγκεφαλογράφημα (electroencephalography, EEG) είναι η καταγραφή της ηλεκτρικής δραστηριότητας στην εξωτερική επιφάνεια του κρανίου για κάποιο χρονικό διάστημα, συνήθως 20-40 λεπτά, μέσω ηλεκτροδίων που τοποθετούνται στο δέρμα. Το σήμα του εγκεφαλογραφήματος προκύπτει από τη διαφορά δυναμικού εξαιτίας της ροής ιόντων στους νευρώνες του εγκεφάλου.

Η κύρια χρήση του εγκεφαλογραφήματος είναι η αναγνώριση εγκεφαλικής δραστηριότητας, δηλαδή να διαπιστωθεί αν ο ασθενής είναι εγκεφαλικά νεκρός, καθώς και η καταγραφή επιληπτικών κρίσεων. Η εξέλιξη των MRI τεχνικών έχει περιορίσει τη διαγνωστική χρήση του σε άλλους τομείς,

αλλά εξακολουθεί να είναι ένα χρήσιμο εργαλείο στην έρευνα της λειτουργίας του εγκεφάλου, καθώς μπορεί να καταγράψει την ενεργοποίηση συγκεκριμένων περιοχών του εγκεφάλου σε συνάρτηση με κάποια ερεθίσματα. Σε αυτό το πλαίσιο έχει κάποιο, σχετικά περιορισμένο, ρόλο στην έρευνα για τη ΝΑ. Ο ρόλος του EEG είναι περιορισμένος γιατί αντίστοιχη λειτουργία και με μεγαλύτερη χωρική ανάλυση προσφέρουν οι fMRI και FDG-PET, αλλά το μικρό κόστος του εξοπλισμού και ο μικρός όγκος του, το ότι δεν επηρεάζεται από την κίνηση του ασθενούς και η εξαιρετική χρονική ανάλυση, καθιστούν το EEG μία χρηστική επιλογή κάποιες φορές. Εξαιτίας της χρονικής ανάλυσης του εγκεφαλογραφήματος έχει προταθεί η χρήση του σε συνδυασμό με άλλες τεχνικές που προσφέρουν υψηλή χωρική ανάλυση και βαθύτερη διεξόδυση στους νευρώνες του εγκεφάλου, όπως η fMRI, προκειμένου να εντοπιστούν αλλαγές στους νοητικούς μηχανισμούς του ανθρώπου.

Μία χαρακτηριστική περίπτωση χρήσης του EEG ως βιοδείκτη για τη ΝΑ είναι η ακόλουθη εργασία [18], όπου οι μελετητές αξιοποίησαν EEG σήματα νοσούντων και μη από ΝΑ, ώστε να εκτιμηθούν οι επιπτώσεις της νόσου στη συνδεσιμότητα του ανθρώπινου εγκεφάλου. Οι περιπτώσεις που μέχρι τώρα εξετάσαμε, εντοπίζουν κυρίως τις ανατομικές αλλαγές στον ανθρώπινο εγκέφαλο. Ωστόσο, ο εγκέφαλος αποτελεί ένα περίπλοκο όργανο που ενεργοποιεί πολλά κέντρα ταυτόχρονα για τη λειτουργία του και άρα η συγκέντρωση στις ανατομικές αλλαγές στο σημαντικότερο χώρο δεν καλύπτει όλη την παθολογία της ΝΑ. Καθώς οι διαδικασίες της νόησης διανέμονται σε πολλά κέντρα, η αξιολόγηση της μεταξύ τους επικοινωνίας είναι σημαντική.

Προηγούμενες έρευνες έχουν όντως δείξει ότι η ΝΑ μεταβάλλει τη συνδεσιμότητα των μερών του εγκεφάλου. Οι ερευνητές προσέγγισαν τον εγκέφαλο ως νευρωνικό δίκτυο και εφήρμοσαν αρχές της θεωρίας γράφων, για να εντοπίσουν τις πιθανές αλλαγές που προκαλεί η ΝΑ. Για να το πετύχουν αυτό, πήραν τα εγκεφαλογραφήματα 20 ασθενών με ΝΑ και 23 φυσιολογικών δειγμάτων. Για κάθε ζευγάρι ηλεκτροδίων μέτρησαν τη λειτουργική συνδεσιμότητά τους για διάφορες συχνότητες και αποθήκευσαν τα δεδομένα σε πίνακες. Οι πίνακες μετατράπηκαν σε γράφους, που αξιολογήθηκαν με κλασικές έννοιες της θεωρίας γράφων: συντελεστής ομαδοποίησης (τοπική συνδεσιμότητα), χαρακτηριστικό μήκος μονοπατιού (ολική συνδεσιμότητα) και συσχέτιση βαθμού (ο βαθμός που οι κόμβοι του δικτύου συγκεντρώνονται με ομοίους τους). Τα αποτελέσματα συγκρίθηκαν με τυχαία δίκτυα σαν ομάδα ελέγχου.



Εικόνα 4: Αλγόριθμος για τη σύνδεση νοητικής κατάστασης νοσούντων από NA και χαρακτηριστικών δικτύου [18]

Τα συμπεράσματα ήταν ενδιαφέροντα: οι γράφοι των ασθενών με NA είχαν μικρότερο συντελεστή ομαδοποίησης, μικρότερο χαρακτηριστικό μήκος μονοπατιού και μειωμένη συσχέτιση βαθμού σε σχέση με τα γραφήματα του δείγματος ελέγχου. Εμφανίζεται δηλαδή μειωμένη τοπική και ολική συνδεσιμότητα στους εγκεφάλους με NA και η συνολική οργάνωση του εγκεφάλου με NA μεταβάλλεται από έναν ιδανικό “μικρό κόσμο” (όπως έχει διαπιστωθεί σε αρκετές έρευνες ο φυσιολογικός εγκέφαλος έχει όντως χαρακτηριστικά “μικρού κόσμου” (small world), δηλαδή καλή ισορροπία ανάμεσα στην τοπική εξειδίκευση και στη συνολική λειτουργία) σε έναν πιο τυχαίο, με λιγότερο αποτελεσματική μεταφορά πληροφορίας. Αν από τα παραπάνω μπορεί να προκύψει κάποιος αποτελεσματικός τρόπος ταξινόμησης για τη NA είναι ανοιχτό ερώτημα.

Τα παραπάνω προσφέρουν μια πρακτική προσέγγιση στη χρήση του εγκεφαλογραφήματος ως υποψήφιου βιοδείκτη για τη NA. Διαφορετικές νευροεκφυλιστικές ασθένειες πιθανότατα έχουν άλλα αποτελέσματα στον ανθρώπινο εγκέφαλο, που ίσως επιτρέπει το διαχωρισμό της NA. Παράλληλα όμως, δίνει και μία εικόνα της NA ως “συνδρόμου αποσύνδεσης” που μπορεί να δια φωτίσει την περαιτέρω έρευνα στην παθοφυσιολογία της NA. Εδώ οφείλουμε να σημειώσουμε ότι αντίστοιχες έρευνες

μπορούν να γίνουν και με άλλες μεθόδους, όπως με την fMRI, και όντως έχουν γίνει σχετικές προσπάθειες, με όμοια συμπεράσματα. Συνολικά το EEG δεν συγκεντρώνει μεγάλο ερευνητικό ενδιαφέρον στη ΝΑ.

2.2 Επεξεργασία εγκεφαλικών εικόνων

2.2.1 Μορφομετρία του εγκεφάλου

Η μορφομετρία του εγκεφάλου ασχολείται με τη μέτρηση των εγκεφαλικών δομών και τις αλλαγές που προκύπτουν στα διαφορετικά στάδια της ζωής. Δεδομένου ότι αντίστοιχα στοιχεία είναι δύσκολα να προκύψουν με επεμβατικές μεθόδους, παρά μόνο στο πλαίσιο της αυτοψίας, ο κλάδος της μορφομετρίας αξιοποιεί εκτεταμένα τις απεικονιστικές μεθόδους που είδαμε παραπάνω και κυρίως την MRI. Η ποσοτικοποίηση των δεδομένων επιτυγχάνεται με διάφορες στατιστικές μεθόδους και παράγει σημαντικά στοιχεία για την κατάσταση του εγκεφάλου, που επιτρέπουν την αξιολόγηση της λειτουργίας του. Χαρακτηριστικό παράδειγμα είναι ο όγκος του ιππόκαμπου που έχει δείχθει καλός βιοδείκτης για τη ΝΑ.

Η μορφομετρία έχει δύο διαφορετικά σκέλη: τον καθεαυτό υπολογισμό των μορφών και την ακόλουθη στατιστική σύγκρισή τους. Μία μελέτη που περιλαμβάνει μορφομετρική εφαρμογή μπορεί να αφορά είτε τη μελέτη των τμημάτων ενός εγκεφάλου και τις μεταξύ τους συνδέσεις, είτε τη σύγκριση των μερών πολλών εγκεφάλων. Είναι σημαντική άρα η διαδικασία με την οποία οι ερευνητές επιλέγουν τα χαρακτηριστικά του εγκεφάλου που τους ενδιαφέρει να ποσοτικοποιήσουν. Επίσης παραπάνω από μία απεικονιστικές μέθοδοι μπορούν να χρησιμοποιηθούν είτε για την εξαγωγή διαφορετικών δεδομένων, είτε για επιβεβαίωση των αποτελεσμάτων. Τέλος, η ποικιλία των διαθέσιμων αλγορίθμων για την επεξεργασία των δεδομένων αποτελεί ένα γιγαντιαίο πεδίο από μόνο του.

Γενικά, τα νευροαπεικονιστικά δεδομένα αποθηκεύονται σε πίνακες από voxels. Το στοιχείο voxel είναι το αντίστοιχο του pixel στον τρισδιάστατο χώρο, δηλαδή αντιπροσωπεύει τη μονάδα απεικόνισης στις τρεις διαστάσεις. Έτσι και η πιο γνωστή μορφομετρική μέθοδος είναι η μορφομετρία βασισμένη σε voxel (Voxel Based Morphometry). Όμως, καθώς ένας πίνακας voxel δεν βολεύει απαραίτητα για την απεικόνιση μίας βιολογικής δομής, γίνονται και προσπάθειες προς διαφορετική οργάνωση των δεδομένων.

2.2.2 Προεπεξεργασία

Το πρώτο στάδιο οποιασδήποτε επεξεργασίας απεικονιστικών δεδομένων είναι η προεπεξεργασία των εικόνων. Αρχικά πολλές εικόνες περιέχουν θόρυβο, όπως για παράδειγμα στην MRI από την κίνηση του ασθενούς ή από ατέλειες του σαρωτή, όπως κάποια μικρή ανομοιογένεια των πεδίων. Δεδομένου ότι ο θόρυβος που προκύπτει από την κίνηση είναι τυχαίος, είναι δύσκολο να μοντελοποιηθεί για να εξουδετερωθεί. Συνεπώς, σε όλες τις μελέτες υφίσταται κάποια διαδικασία απόρριψης των μη-κατάλληλων εικόνων.

2.2.3 Κατάτμηση

Στη συνέχεια ακολουθεί η κατάτμηση (segmentation) της εικόνας. Ο στόχος της κατάτμησης είναι να διαιρέσει την εικόνα σε διαφορετικές ομάδες, προκειμένου να απλοποιηθεί η δομή της ή να γίνει πιο διαχειρίσιμη. Για παράδειγμα η κατάτμηση μπορεί να αφορά την εύρεση κάποιας γραμμής, ώστε να διαχωριστεί μία δομή από την υπόλοιπη εικόνα. Η κατάτμηση μπορεί να γίνει με πολλά κριτήρια, όπως χρώμα και υφή (πχ κατωφλίωση) και η εφαρμογή της συνιστά ένα ευρύ πεδίο εφαρμογών. Περιπτώσεις αυτοματοποιημένης κατάτμησης θα εξετάσουμε στην πορεία.

2.2.4 Επιλογή χαρακτηριστικών– με μορφομετρία βασισμένη σε voxels

Η φυσική συνέχεια της κατάτμησης είναι να επιλεγεί κάποια μέθοδος για την καταγραφή των όγκων. Όπως προαναφέραμε, μία από τις πιο κοινές είναι η μορφομετρία βασισμένη σε voxels (Voxel Based Morphometry, VBM). Σε αυτή τη διαδικασία οι εικόνες κανονικοποιούνται χωρικά μέσω της αναγωγής σε έναν πρότυπο εγκέφαλο και διορθώνονται αλλαγές στον όγκο, ώστε για παράδειγμα η πιθανή διαφορά του μεγέθους του εγκεφάλου ανάμεσα στα δείγματα να μην συμμετέχει στην ακόλουθη στατιστική επεξεργασία. Επίσης οι εικόνες ομαλοποιούνται, ώστε το κάθε voxel να περιέχει το μέσο όρο του εαυτού του και των γειτόνων του.

Από αυτά τα στοιχεία γίνεται η επιλογή των χαρακτηριστικών (feature selection) και η ακόλουθη στατιστική επεξεργασία των δεδομένων. Κλασικό παράδειγμα είναι το γενικό γραμμικό μοντέλο (general linear model), που ουσιαστικά είναι ο συνδυασμός πολλών γραμμικών παλινδρομήσεων για να αντιμετωπισθούν οι πολλές ανεξάρτητες μεταβλητές. Την παλινδρόμηση σαν αλγόριθμο στατιστικής επεξεργασίας θα εξετάσουμε στην πορεία.

Γενικά η VBM βελτίωσε την παραδοσιακή μέθοδο, που επέλεγε περιοχές ενδιαφέροντος (regions of interest, ROIs) και υπολόγιζε όγκους, με την προβολή των εικόνων σε έναν κοινό παρονομαστή που ομαλοποιεί τα δεδομένα και επιτρέπει ακριβέστερα αποτελέσματα. Η εφαρμογή της εξακολουθεί να

είναι πολύ συνηθισμένη και υφίσταται σε πολλά προγράμματα νευροαπεικονιστικής επεξεργασίας. Η στατιστική ευρωστία των αποτελεσμάτων όμως αμφισβητείται. Άλλες μέθοδοι ανεξαρτητοποιούνται από τα voxels και προσφέρουν μη γραμμικές προσεγγίσεις με πιο σταθερά αποτελέσματα.

2.2.5 Μορφομετρία βασισμένη σε tensors

Η πιο συχνή περίπτωση που θα συναντήσουμε, μετά την VBM, είναι η μορφομετρία βασισμένη σε tensors (tensor based morphometry, TBM). Η TBM αποτελεί μία παραλλαγή του ευρύτερου κλάδου της βασισμένης στην παραμόρφωση μορφομετρίας (deformation based morphometry), όπου χρησιμοποιούνται μη γραμμικοί αλγόριθμοι για να δημιουργήσουν πεδία παραμόρφωσης για την αντιστοίχιση με το πρότυπο. Η ακόλουθη στατιστική επεξεργασία δεν εφαρμόζεται στα voxels που προέκυψαν, αλλά στα πεδία παραμόρφωσης ή στις βαθμωτές ιδιότητες των πινάκων με τα voxels. Η TBM συγκεκριμένα αξιοποιεί την ιακωβιανή ορίζουσα του πίνακα παραμόρφωσης.

Γενικά υπάρχουν πολλές επιλογές όσον αφορά τον τρόπο παραμόρφωσης των δεδομένων, που γίνονται με βάση τις ανάγκες για τοπική ακρίβεια σε κάποιο σημείο και την προσαρμογή του μοντέλου στο σύνολο του εγκεφάλου. Οι συγκεκριμένοι αλγόριθμοι είναι ακριβότεροι υπολογιστικά από την απλούστερη γραμμική επεξεργασία της VBM, όμως μπορούν να εντοπίσουν με μεγαλύτερη ακρίβεια μικρές αλλαγές, ιδιαίτερα σε μακροχρόνιες μελέτες. Εδώ πρέπει να σημειώσουμε ότι αυτό είναι ακόμη ένα ανοιχτό ερευνητικό πεδίο και υπάρχουν πολλές αντικρουόμενες προτάσεις, αν και η TBM κερδίζει έδαφος και εφαρμόζεται σε όλο και περισσότερες μελέτες.

2.2.6 Άλλες μέθοδοι

Πέρα από τα παραπάνω υφίστανται και άλλες προσεγγίσεις στο πρόβλημα της μορφομετρίας, όπως η μορφομετρία που βασίζεται σε μοτίβα (pattern based morphometry), που επεκτείνει τα παραπάνω μοντέλα εντοπίζοντας μοτίβα στη συνολική εικόνα και δεν περιορίζεται σε συγκεκριμένα voxels. Έτσι μπορεί να επιτευχθεί καλύτερος διαχωρισμός.

Συνολικά η μορφομετρία του εγκεφάλου περιέχει αρκετά διαφορετικά μοντέλα, που καλύπτουν διαφορετικές ανάγκες. Όσον αφορά τη NA, η φυσιοπαθολογία του μπορεί να είναι εκτεταμένη, αλλά οι κοινά αποδεκτοί ανατομικοί βιοδείκτες είναι περιορισμένοι. Συνεπώς μέθοδοι όπως η TBM, που ειδικεύεται στην παρακολούθηση περιορισμένων σημείων του εγκεφάλου είναι ιδιαίτερα κατάλληλες. Η εγκεφαλική μορφομετρία δίνει και τα εγκυρότερα ποσοστά πρόβλεψης της πορείας της NA, όπως θα δούμε σε μελέτες που συγκρίνουν τους διάφορους βιοδείκτες, οπότε γίνεται αντιληπτή η σημασία της για την ασθένεια. Η αυτοματοποίηση και βελτιστοποίηση της παραπάνω διαδικασίας απασχολεί την

επιστημονική κοινότητα, που έχει ήδη αρκετές επιλογές για τον αυτόματο εντοπισμό του ιππόκαμπου για παράδειγμα. Αυτές οι επιλογές, αλλά και η στατιστική επεξεργασία των δεδομένων, θα μας απασχολήσουν στην πορεία.

2.3 Παραγοντοποίηση του σήματος σε πηγές – Αυτοματοποιημένη εξαγωγή και επιλογή χαρακτηριστικών

Η στατιστική επεξεργασία των εικόνων αλλά και γενικότερα των δεδομένων με αυτοματοποιημένο τρόπο προϋποθέτει πώς θα εξαχθεί κάποιο μοντέλο από τα δεδομένα. Κάποιοι από τους παρακάτω αλγορίθμους προσπαθούν να δημιουργήσουν ένα μοντέλο που εξάγεται από τα δεδομένα (γραμμική παλινδρόμηση, πολλαπλή γραμμική παλινδρόμηση). Στις πιο ενδιαφέρουσες μεθόδους, ο κατεξοχήν στόχος είναι η ομαδοποίηση ή διάσπαση των δεδομένων σε (κατά το δυνατόν) ανεξάρτητους ή σημαντικούς σχηματισμούς (ανάλυση σε πρωταρχικές συνιστώσες, γραμμική διακριτική ανάλυση, πολυδιακριτή ανάλυση) ώστε να εξαχθούν χαρακτηριστικά για τα δεδομένα (feature extraction). Η αναγνώριση και εξαγωγή χαρακτηριστικών μπορεί να γίνει με διάφορους τρόπους. Στην περίπτωση των απεικονιστικών δεδομένων συνήθως αφορά τον εντοπισμό ακμών, περιοχών ενδιαφέροντος (regions of interest, ROIs) και κορυφογραμμών (ridges). Η εξαγωγή χαρακτηριστικών περιλαμβάνει τη μετάβαση από το χώρο της εικόνας στο χώρο των χαρακτηριστικών (feature space) που περιλαμβάνει τη μείωση του αριθμού των διαστάσεων της πληροφορίας.

Από εκεί και πέρα, εφαρμόζονται βάρη για την αφαίρεση των μη σημαντικών παραγόντων ή για την αξιοποίηση συσχετίσεων στα δεδομένα, είτε μέσα στην ίδια πηγή δεδομένων, είτε σε διαφορετικές (παράλληλη ανάλυση σε ανεξάρτητες συνιστώσες, παλινδρόμηση κορυφογραμμής, LASSO μέθοδος). Αυτές οι μέθοδοι δηλαδή σκοπεύουν και στην επιλογή των χαρακτηριστικών που μας ενδιαφέρουν από το σύνολο των διαθέσιμων χαρακτηριστικών (feature selection). Η επιλογή των χαρακτηριστικών είναι σχεδόν πάντα απαραίτητη στην επεξεργασία ιατρικής εικόνας, γιατί οι εικόνες MRI και PET είναι γενικά πολύ μεγάλες για αποτελεσματική επεξεργασία και δεν χρειάζονται στο σύνολό τους. Η προσεκτική επιλογή των χαρακτηριστικών μειώνει τον όγκο των δεδομένων, χωρίς να χάνει χρήσιμα δεδομένα. Παράλληλα, τα ακόλουθα μοντέλα επεξεργασίας και ταξινόμησης είναι πιο έγκυρα, δεδομένου ότι δεν χρειάζεται να προσαρμοστούν σε έναν υπερβολικά μεγάλο αριθμό μεταβλητών, ενώ μειώνεται και ο χρόνος εκπαίδευσης του συστήματος. Στις έρευνες που θα συναντήσουμε συνήθως γίνεται κάποια προεπιλογή απεικονιστικών χαρακτηριστικών με την έννοια ότι εξετάζονται μόνο συγκεκριμένες περιοχές του εγκεφάλου που γνωρίζουμε ότι σχετίζονται με την εμφάνιση ΝΑ. Η

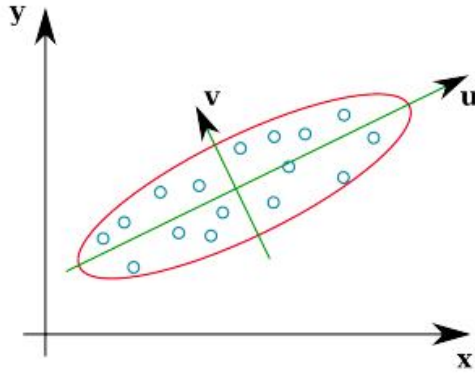
αυτοματοποιημένη εξαγωγή και επιλογή χαρακτηριστικών έπεται αυτής της επιλογής. Στα πολυτροπικά συστήματα βέβαια, όπου τα μοντέλα εξάγουν πληροφορία από πολλές πηγές δεδομένων, η επιλογή χαρακτηριστικών γίνεται ακόμη πιο κρίσιμη. Θα δούμε στην πορεία ότι ο αριθμός των δεικτών και ο συνδυασμός τους ώστε η ίδια πληροφορία να μην καλύπτεται από πολλές πηγές είναι ένας από τους κύριους στόχους για τη βελτιστοποίηση των συστημάτων.

2.3.1 Ανάλυση σε πρωταρχικές συνιστώσες

Η ανάλυση σε πρωταρχικές συνιστώσες (Principal Component Analysis, PCA) είναι μία μαθηματική μέθοδος που μετατρέπει μία ομάδα δεδομένων (συσχετισμένων ή όχι) σε μία ομάδα γραμμικά μη συσχετισμένων διανυσμάτων. Αυτό εξασφαλίζεται με την απαίτηση τα διανύσματα να είναι κάθετα μεταξύ τους (ορθογωνική μετατροπή). Ο αριθμός των διανυσμάτων (principal components) είναι μικρότερος ή ίσος του αρχικού αριθμού των μεταβλητών. Τα διανύσματα παράγονται με τέτοιο τρόπο, ώστε κάθε διάνυσμα να έχει τη μέγιστη δυνατή διακύμανση με την προϋπόθεση να είναι κάθετο στα προηγούμενα.

Ο παραπάνω αλγόριθμος χρησιμοποιείται εκτεταμένα στην ανάλυση πολυπαραγοντικών δεδομένων και στα μοντέλα πρόβλεψης. Στην πράξη ο στόχος είναι η αποκάλυψη της εσωτερικής δομής των δεδομένων, προκειμένου να εξασφαλιστεί η μέγιστη διακύμανση. Έτσι, η μείωση του αριθμού των διαστάσεων στις πρώτες n κατευθύνσεις, παρέχει στο χρήστη την εικόνα που εξηγεί ένα αντίστοιχο ποσοστό των δεδομένων [19].

Ένα απλό παράδειγμα είναι μία ομάδα τρισδιάστατων δεδομένων που εξαρτάται ως επί το πλείστον από δύο παράγοντες. Από την εφαρμογή του PCA θα προκύψει ένα επίπεδο εντός του χώρου, που θα πρέπει να περιλαμβάνει τις περισσότερες παρατηρήσεις. Το επίπεδο θα προκύψει από τις δύο κατευθύνσεις που υπάρχει μεγαλύτερη διακύμανση των δεδομένων.



Εικόνα 5: Παράδειγμα εξαγωγής επιπέδου από χώρο με την εφαρμογή της ανάλυσης σε πρωταρχικές συνιστώσες σε τρισδιάστατα δεδομένα [20]

Ο αλγόριθμος του PCA παρέχει εύρωστες λύσεις, υπό τον όρο ότι τα δεδομένα είναι καλά κανονικοποιημένα και το ισχυρότερο κριτήριο ανάλυσής τους είναι η διακύμανση των δεδομένων. Αυτή είναι μια εύλογη προσέγγιση για ένα μεγάλο μέρος δεδομένων, αλλά μη γκαουσιανές κατανομές ή ακόμη και ένα ζεύγος στενών και πλατιών γκαουσιανών κατανομών δεν αναλύονται ικανοποιητικά από τον PCA αλγόριθμο.

Αυτό το πρόβλημα επιλύουν αλγόριθμοι όπως η ανάλυση σε ανεξάρτητες συνιστώσες που θα δούμε στη συνέχεια, που χρησιμοποιούν άλλα κριτήρια για να επιλέξουν κατευθύνσεις.

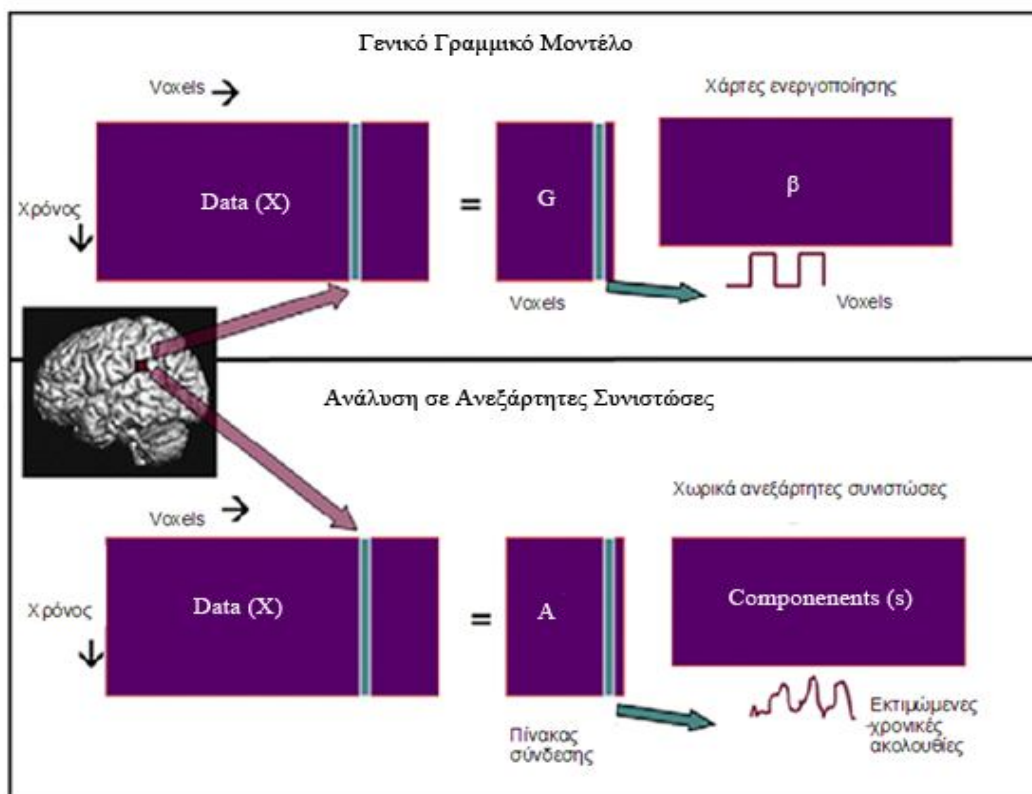
2.3.2 Ανάλυση σε ανεξάρτητες συνιστώσες

Η ανάλυση σε ανεξάρτητες συνιστώσες (Independent Component Analysis, ICA) είναι μία υπολογιστική μέθοδος που αποσκοπεί στη διάσπαση ενός πολυπαραγοντικού σήματος, υποθέτοντας ότι το σήμα αποτελείται από το γραμμικό συνδυασμό στατιστικά ανεξάρτητων πηγών. Εμπίπτει στην κατηγορία του τυφλού διαχωρισμού πηγών, καθώς η μέθοδος δεν προϋποθέτει πληροφορία όσον αφορά τις πηγές ή τη διαδικασία ανάμιξης των πηγών. Εφόσον ισχύει ότι το σήμα είναι γραμμικός συνδυασμός ανεξάρτητων, μη γκαουσιανών πηγών, η τεχνική ICA είναι σε θέση να δώσει πολύ καλά αποτελέσματα. Σε σύγκριση με την τεχνική PCA, που επίσης αποσκοπεί στην αποκάλυψη κρυμμένων τμημάτων πληροφορίας, η τεχνική ICA δεν αποσυσχετίζει απλώς τα δεδομένα, αλλά επιτυγχάνει ανεξαρτησία των πηγών. Η τεχνική ICA χρησιμοποιείται εκτεταμένα στην επεξεργασία εικόνας, αλλά παραλλαγές της χρησιμοποιούνται και για την ενοποίηση δεδομένων. Σε αυτές τις δύο μορφές έχει μεγάλο ενδιαφέρον στην έρευνα της ΝΑ [21].

Η κεντρική ιδέα του απλού ICA περιγράφεται από τις ακόλουθες εξισώσεις:

$$X = A \times S, \hat{S} = W \times X, W = A^{-1}$$

Στην προκειμένη περίπτωση, X είναι το σήμα μας, που αποτελεί συνδυασμό των πηγών S και του A , που εκπροσωπεί τη διαδικασία μίξης των πηγών. Ο πίνακας W αντιπροσωπεύει την αντίστροφη διαδικασία του A . Ουσιαστικά ο αλγόριθμος ICA προσπαθεί να βρει τον W , ώστε ο \hat{S} να προσεγγίζει όσο το δυνατόν καλύτερα τον S .



Εικόνα 6: Ανάλυση σε ανεξάρτητες συνιστώσες σε σχέση με το γενικό γραμμικό μοντέλο [21]

Η αρχή πίσω από τη λειτουργία του αλγορίθμου ICA είναι ότι η εύρεση ανεξάρτητων πηγών μπορεί να προκύψει από την εύρεση των κατευθύνσεων των μέγιστα μη γκαουσιανών κατανομών στο χώρο δεδομένων. Παράλληλα, και σε αντίθεση με τον PCA αλγόριθμο, δεν απαιτεί αυτές οι κατευθύνσεις να είναι ορθογώνιες.

Ο πλέον κλασικός αλγόριθμος εύρεσης του W και εξασφάλισης της ανεξαρτησίας των πηγών είναι ο αλγόριθμος μεγιστοποίησης της πληροφορίας (Infomax), που πρακτικά προσπαθεί να μεγιστοποιήσει την εντροπία, όπως εκφράζεται από τις παρακάτω εξισώσεις

$$\max\{H(Y)\} = -E[\ln f_y(Y)]$$

$$Y = \frac{1}{1 + e^{-U}}, U = WxX + W_0$$

όπου f_y είναι η πυκνότητα της συνάρτησης πιθανότητας Y , ενώ E είναι η συνάρτηση αναμονής και H είναι η συνάρτηση της εντροπίας.

Μία κοινή χρήση του αλγορίθμου ICA είναι στην επεξεργασία δεδομένων από fMRI. Η χρονική ανταπόκριση του εγκεφάλου είναι δύσκολο να εκτιμηθεί σε πολλά πειράματα που απαιτούν fMRI, καθώς το μοντέλο ενεργοποίησης του εγκεφάλου δεν είναι απόλυτα κατανοητό. Επειδή για να λειτουργήσει ο ICA αλγόριθμος δεν χρειάζεται να γνωρίζει τη διαδικασία μίξης των πηγών στο τελικό σήμα, μπορεί να διαχωρίσει τις περιοχές του εγκεφάλου χωρίς να γνωρίζει τη χρονική απόκριση κάθε περιοχής. Ο ICA είναι όμως σε θέση να διακρίνει και τους χρόνους ενεργοποίησης των διαφόρων περιοχών του εγκεφάλου, επεξεργάζεται δηλαδή πληροφορία χρονικά και χωρικά.

Το προφανές πρόβλημα στη χρήση του ICA για ανάλυση fMRI είναι ότι οι περιοχές του εγκεφάλου δεν παράγουν απαραίτητα ανεξάρτητο σήμα. Έχει παρατηρηθεί ότι τα περισσότερα πειράματα περιλαμβάνουν την ενεργοποίηση διαφορετικών περιοχών του εγκεφάλου, συνεπώς θεωρείται ότι μπορούν να προσεγγιστούν ικανοποιητικά από τον ICA. Έχουν επίσης προταθεί αρκετές μέθοδοι που μεγιστοποιούν τη στατιστική ανεξαρτησία των πηγών με ικανοποιητικά αποτελέσματα. Ωστόσο για να αντιμετωπισθεί το πρόβλημα της μη-ανεξαρτησίας των πηγών έχει προταθεί και η παράλληλη ανάλυση σε ανεξάρτητες συνιστώσες.

2.3.3 Παράλληλη ανάλυση σε ανεξάρτητες συνιστώσες

Η χρήση του ICA αλγορίθμου αποδείχθηκε τόσο πετυχημένη, ώστε πλέον χρησιμοποιείται εκτεταμένα και στην ενοποίηση και διαχείριση δεδομένων από εντελώς διαφορετικές εξετάσεις. Είναι γεγονός ότι η φιλοσοφία του ICA, όπως την προσεγγίσαμε μέχρι τώρα, δεν προσφέρεται για την εξέταση δεδομένων σε πολλές μορφές, καθώς κάνει πολύ λίγες παραδοχές. Ωστόσο υπάρχουν αρκετοί τρόποι για να παρακαμφθεί αυτή η δυσκολία, κυρίως με την αποδοχή κάποιας επιπλέον συνθήκης, όπως ότι όλα τα δεδομένα είναι συμφασικά στο πεδίο του χρόνου, ή ότι εκφράζονται σε κοινό χωρικό χάρτη. Στο πλαίσιο, όμως, των πολυτροπικών (multimodal) συστημάτων για τη ΝΑ, μεγαλύτερο ενδιαφέρον παρουσιάζει η παραλλαγή του ICA γνωστή ως παράλληλη ανάλυση σε ανεξάρτητες συνιστώσες (Parallel Independent Component Analysis, paraICA).

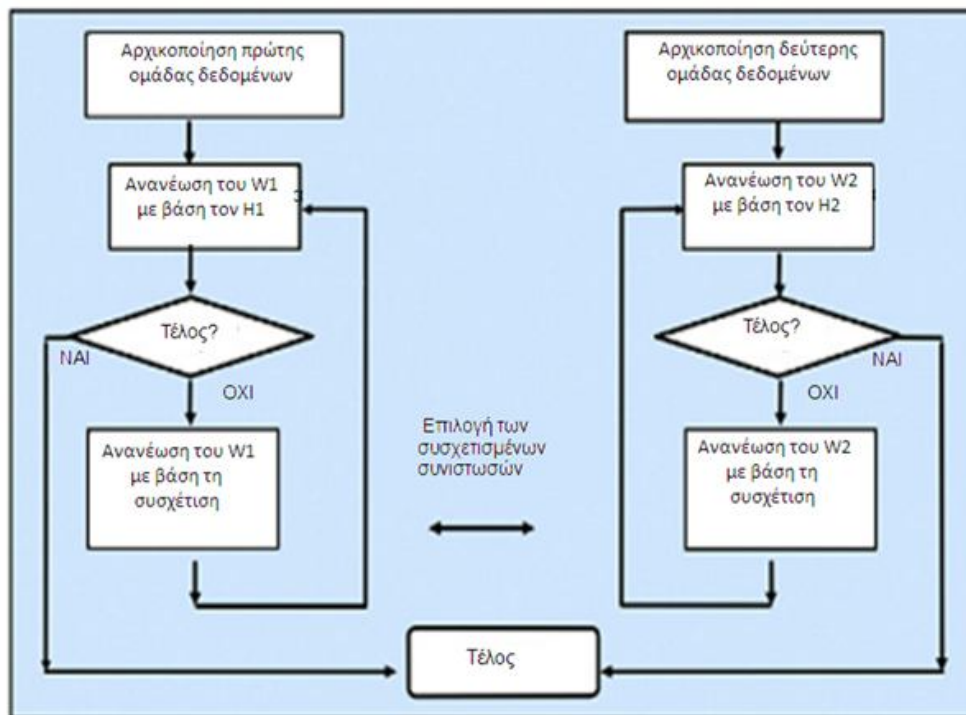
Στην πράξη ο paraICA τρέχει δύο ή και περισσότερους απλούς ICA ταυτοχρόνως, δηλαδή μπορούμε να περιγράψουμε τη διαδικασία ως εξής για δύο διαφορετικά σετ δεδομένων:

$$X_1 = A_1 \times S_1; S_1 = W_1 \times X_1;$$

$$X_2 = A_2 \times S_2; S_2 = W_2 \times X_2$$

Επίσης χρειάζονται δύο συναρτήσεις για την εξασφάλιση της ανεξαρτησίας των πηγών δεδομένων, για παράδειγμα με τη μεγιστοποίηση της εντροπίας σε κάθε σετ δεδομένων. Τέλος όμως χρειάζεται και μία επιπλέον συνάρτηση που προσπαθεί να ορίσει τη σχέση των δύο σετ δεδομένων, μέσω της σχέσης των A_1, A_2 . Ακολουθεί το εποπτικό διάγραμμα του αλγορίθμου.

Παράλληλη Ανάλυση σε Ανεξάρτητες Συνιστώσες



Εικόνα 7: Ομαδοποίηση δύο πηγών δεδομένων με παράλληλη ανάλυση σε ανεξάρτητες συνιστώσες [22]

Το σημαντικό πλεονέκτημα του paraICA είναι ότι κρατά τα χαλαρά κριτήρια του απλού ICA και δεν περιορίζει τον ερευνητή με επιπλέον παραδοχές, όπως κοινό χρονικό πεδίο. Επιπλέον, τα αποτελέσματα του paraICA είναι εύρωστα και οι έρευνες που τον χρησιμοποιούν συνεχώς αυξάνονται.

Μία από τις πρώτες έρευνες που εξέτασε τις δυνατότητες του paraICA σε εξετάσεις του εγκεφάλου είναι η εργασία των Lyu και Calhoun το 2007 [22]. Οι ερευνητές δημιούργησαν δύο σετ δεδομένων που μιμούνταν αποτελέσματα fMRI και EEG, με βάση άλλες δικές τους έρευνες και προσέθεσαν διάφορα

επίπεδα θορύβου ή επιπλέον συνδέσεις. Η προσέγγιση του paraICA συγκρίθηκε με την προσέγγιση του απλού ICA. Τα κριτήρια που χρησιμοποιήθηκαν ήταν η ακρίβεια στην προσέγγιση των πηγών και των μεταξύ τους συνδέσεων.

(A) Signal to noise ratio	paraICA		ICA		
	Connection Accuracy	Component accuracy	Connection Accuracy	Component accuracy	
10	0.8337	0.9672	0.6786	0.9659	
20	0.8383	0.9667	0.7747	0.9658	
30	0.9278	0.9676	0.8606	0.9659	
(B) Connection strength					
1.00	0.9278	0.9676	0.8606	0.9659	
0.80	0.7911	0.9661	0.7348	0.9663	
0.60	0.5850	0.9665	0.5205	0.9657	
0.45	0.5303	0.9637	0.4138	0.9656	
(C) Data1/Data2 Estimated component number					
8	4	0.4149	0.7860	0.4148	0.7860
	8	0.8275	0.9630	0.6635	0.9657
	16	0.8068	0.9608	0.6652	0.9633
	20	0.8240	0.9651	0.6639	0.9654
4	8	0.1739	0.7917	0.1738	0.7917
8		0.8492	0.9623	0.6636	0.9657
16		0.8953	0.9596	0.6636	0.9657
20		0.9587	0.9599	0.6635	0.9657

Πίνακας 1: Σύγκριση επιδόσεων paraICA με ICA [22]

Βλέπουμε ότι και οι δύο αλγόριθμοι προσεγγίζουν καλά τις διαφορετικές πηγές. Όμως ο paraICA αναγνωρίζει πολύ καλύτερα τις μεταξύ τους συνδέσεις για όλα τα διαφορετικά επίπεδα θορύβου ή δυναμική των συνδέσεων. Όμως και οι δύο αλγόριθμοι, αλλά ιδιαίτερα ο paraICA, εμφανίζουν κακές σχετικά επιδόσεις όταν υποτιμάται ή υπερτιμάται πολύ ο αριθμός των πηγών. Αυτή είναι μια παρατήρηση που θα πρέπει να ληφθεί υπόψη σε μελλοντικές εφαρμογές του αλγορίθμου.

Συμπερασματικά, ο paraICA παράγει καλά αποτελέσματα, ακόμα και σε ακραίες συνθήκες. Η διαφορετική του προσέγγιση από άλλες εφαρμογές του ICA επιτρέπει την καλύτερη προσέγγιση των σχέσεων διάφορων πηγών δεδομένων.

Ο paraICA όντως χρησιμοποιείται σε πολλά συστήματα ενοποίησης δεδομένων. Περιπτώσεις εφαρμογής του θα εξεταστούν στο αντίστοιχο κεφάλαιο.

2.3.4 Παλινδρόμηση

Η ανάλυση με παλινδρόμηση είναι μία στατιστική τεχνική για την εκτίμηση των σχέσεων μεταξύ μεταβλητών. Συνήθως η παλινδρόμηση δείχνει πώς μεταβάλλεται η τιμή μιας εξαρτημένης μεταβλητής όταν αλλάζει η τιμή μιας ανεξάρτητης μεταβλητής, ενώ η τιμή των υπόλοιπων ανεξάρτητων μεταβλητών παραμένει σταθερή. Σε όλες πάντως τις περιπτώσεις ο στόχος της ανάλυσης είναι η εύρεση μιας προσεγγιστικής συνάρτησης που περιλαμβάνει όλες τις ανεξάρτητες μεταβλητές. Οι πιο γνωστές εφαρμογές της είναι η μέθοδος ελαχίστων τετραγώνων και η γραμμική παλινδρόμηση. Επέκταση της γραμμικής παλινδρόμησης είναι το γενικό γραμμικό μοντέλο (general linear model), που χρησιμοποιείται για την αντιμετώπιση πολυπαραγοντικών μοντέλων, υποθέτοντας ανεξάρτητες μεταβλητές.

Η ανάλυση με παλινδρόμηση χρησιμοποιείται ευρύτατα στον τομέα της πρόβλεψης και στη μάθηση υπολογιστών. Μία ακόμα κλασική εφαρμογή της είναι η ανεύρεση των ανεξάρτητων μεταβλητών που καθορίζουν την εξαρτημένη, η ανάδειξη δηλαδή της συσχέτισης ανάμεσα σε μεταβλητές. Σε αυτήν τη μορφή αξιοποιείται εκτεταμένα και στην έρευνα της ΝΑ. Η παλινδρόμηση είναι ένα πολύ ισχυρό εργαλείο, αλλά η αποτελεσματικότητά της εξαρτάται από τον τρόπο που προέκυψαν τα δεδομένα και τη συμβατότητά τους με τη συγκεκριμένη μέθοδο παλινδρόμησης. Πολλές φορές θα χρειαστεί να ληφθούν παραδοχές για τα δεδομένα που δεν είναι πάντα ελέγξιμες. Η λανθασμένη σχέση αιτιατού ανάμεσα σε δύο μεταβλητές είναι τέλος ένας πολύ συνηθισμένος κίνδυνος στις τεχνικές παλινδρόμησης.

Οι μέθοδοι παλινδρόμησης εξακολουθούν να έχουν ερευνητικό ενδιαφέρον, κυρίως προς την αντιμετώπιση συσχετισμένων δεδομένων και τη διαχείριση ελλιπών δεδομένων. Αυτές οι νέες κατευθύνσεις έχουν αξιοποιηθεί και στην έρευνα για τη ΝΑ.

2.3.5 Παλινδρόμηση κορυφογραμμής

Η μέθοδος των ελαχίστων τετραγώνων που προαναφέρθηκε αποδεδειγμένα παράγει την καλύτερη γραμμική μη προκατειλημμένη (unbiased) προσέγγιση. Στην πολύ πιθανή περίπτωση όμως που οι μεταβλητές είναι συσχετισμένες, η ακρίβεια της παραπάνω προσέγγισης μειώνεται κατακόρυφα.

Η παλινδρόμηση κορυφογραμμής (ridge regression, συστηματοποίηση κατά Tikhonov) είναι μια κοινή μέθοδος για την κανονικοποίηση προβλημάτων που δεν μπορούν να προσεγγιστούν καλά από την απλή παλινδρόμηση, είτε γιατί δεν υπάρχουν κάποιες ανεξάρτητες μεταβλητές είτε γιατί δεν είναι ανεξάρτητες.

Αυτό επιτυγχάνεται με μία συνάρτηση που πρακτικά “τιμωρεί” το μεγάλο μέγεθος των παραμέτρων της παλινδρόμησης. Το κόστος είναι ότι οι νέες προσεγγίσεις αποκλείεται να ταυτίζονται με τις πραγματικές τιμές και συνήθως τείνουν να τις υποτιμούν. Πάντως η λύση που δίνει στο πρόβλημα της συσχέτισης μεταβλητών είναι εξαιρετικά εύρωστη και η χρήση της έχει εξαπλωθεί σε πολλούς επιστημονικούς κλάδους.

2.3.6 Μέθοδος least absolute shrinkage and selection operator

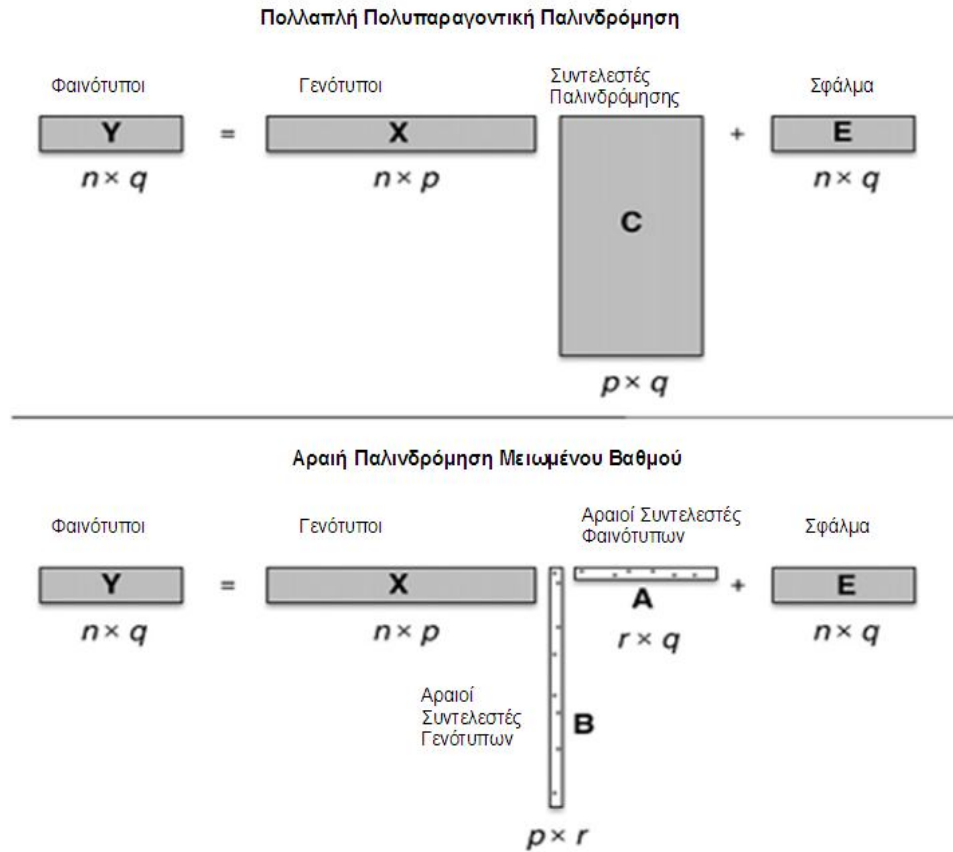
Η μέθοδος LASSO (Least Absolute Shrinkage and Selection Operator) είναι μια κανονικοποιημένη εκδοχή της μεθόδου ελαχίστων τετραγώνων, που περιορίζει το μέγεθος του διανύσματος παραμέτρων. Το πλεονέκτημά της σε σχέση με την παλινδρόμηση κορυφογραμμής είναι ότι τείνει να προτιμά λύσεις με λιγότερες μη μηδενικές παραμέτρους, δηλαδή μειώνει τον αριθμό των ανεξάρτητων μεταβλητών από τις οποίες εξαρτάται η λύση.

2.3.7 Ελαστικό δίκτυο

Η μέθοδος του ελαστικού δικτύου (elastic net) αποτελεί ένα συνδυασμό της παλινδρόμησης κορυφογραμμής και της LASSO μεθόδου. Ουσιαστικά το ελαστικό δίκτυο υπολογίζει το αποτέλεσμα σε δύο επίπεδα, πρώτα τις συνιστώσες σύμφωνα με την παλινδρόμηση κορυφογραμμής και μετά συρρικνώνει τις συνιστώσες με τη LASSO μέθοδο. Αυτό είναι απαραίτητο γιατί η LASSO μέθοδος μόνη της δυσκολεύεται να αντιμετωπίσει στενά συσχετισμένες μεταβλητές.

2.3.8 Παλινδρόμηση μειωμένου βαθμού

Όπως έχουμε δει, η πολλαπλή γραμμική παλινδρόμηση δυσκολεύεται να αντιμετωπίσει την αλληλεξάρτηση των πολυδιάστατων μεταβλητών. Για να αποκτήσει πραγματικό πολυμεταβλητό χαρακτήρα το μοντέλο της παλινδρόμησης, η παλινδρόμηση μειωμένου βαθμού (reduced rank regression) περιορίζει το βαθμό του πίνακα συντελεστών της παλινδρόμησης. Αυτό γίνεται με την εφαρμογή μιας ομαδικής επιβάρυνσης (penalty) όμοια με αυτή που γίνεται στη μέθοδο LASSO σε κάθε σειρά του πίνακα συντελεστών. Σε ειδικές περιπτώσεις, η παλινδρόμηση μειωμένου βαθμού καταλήγει σε απλή εφαρμογή του PCA αλγορίθμου. Η παλινδρόμηση μειωμένου βαθμού με μια απλή επέκταση των πινάκων (ώστε να μπορούν να αναστραφούν) προσφέρει σταθερά αποτελέσματα και σε περιπτώσεις που ο αριθμός των μεταβλητών είναι μεγαλύτερος από τον αριθμό των πειραματικών παρατηρήσεων. Αυτό είναι το πλαίσιο στο οποίο θα τη συναντήσουμε, στο κεφάλαιο για τις έρευνες στο σύνολο του γονιδιώματος, ως αραιή παλινδρόμηση μειωμένου βαθμού (sparse reduced rank regression, sRRR).



Και η πολλαπλή πολυπαραγοντική παλινδρόμηση και η αραιή παλινδρόμηση μειωμένου βαθμού είναι πολυπαραγοντικά μοντέλα, αλλά για να δώσει έγκυρα αποτελέσματα το πρώτο μοντέλο πρέπει το δείγμα n να είναι μεγαλύτερο της διάστασης p ή να επιβληθούν περιορισμοί στον πίνακα συντελεστών C .

Εικόνα 8: Ομαδοποίηση με sRRR και με πολλαπλή πολυπαραγοντική παλινδρόμηση [23]

2.3.9 Γραμμική διακριτική ανάλυση

Η γραμμική διακριτική ανάλυση (Linear Discriminant Analysis, LDA) είναι μία μέθοδος στατιστικής επεξεργασίας που, όμοια με την παλινδρόμηση, επιδιώκει να εκφράσει μια μεταβλητή ως έναν γραμμικό συνδυασμό χαρακτηριστικών. Η διαφορά με την παλινδρόμηση είναι ότι η εξαρτημένη μεταβλητή είναι κατηγορική (για παράδειγμα ανήκει σε κάποια κλάση) και όχι αριθμητική. Πρακτικά η τεχνική LDA προσπαθεί να διαφοροποιήσει την εκάστοτε κλάση μέσω του γραμμικού συνδυασμού των μεταβλητών. Ο LDA αλγόριθμος επίσης απαιτεί οι ανεξάρτητες μεταβλητές να έχουν κανονική κατανομή.

2.3.10 Επιβαρυσμένη γραμμική διακριτική ανάλυση

Αντίστοιχα με τις LASSO μεθόδους στην παλινδρόμηση, όταν στην LDA μέθοδο υπάρχει υψηλή συσχέτιση των μεταβλητών, απαιτείται η επιβολή κάποιου ορίου. Αυτό γίνεται με την εφαρμογή ενός πίνακα βαρών, όπως στην παλινδρόμηση κορυφογραμμής. Η παραλλαγή αυτή του κλασικού αλγορίθμου LDA ονομάζεται επιβαρυσμένη γραμμική διακριτική ανάλυση (Penalized LDA, επιβαρυσμένο LDA).

2.3.11 Λογιστική παλινδρόμηση

Η λογιστική παλινδρόμηση (logistic regression) είναι μία παραλλαγή της κλασικής παλινδρόμησης. Προβλέπει τις τιμές μίας ή περισσότερων μεταβλητών εμπειρικά μέσω ενός μοντέλου ποιοτικής απόκρισης. Πρακτικά, ο πυρήνας της παλινδρόμησης καθορίζεται από μία λογιστική συνάρτηση. Είναι μία από τις πιο δημοφιλείς μεθόδους στον τομέα πρόβλεψης τιμών.

2.3.12 Πολυδιακριτή ανάλυση

Η πολυδιακριτή ανάλυση (Multiresolution Analysis, MRA) είναι μία μέθοδος ανάλυσης σημάτων που χρησιμοποιείται εκτεταμένα στο μετασχηματισμό διακριτών κυματομορφών [24]. Πρακτικά επιδιώκει τη διάσπαση του σήματος σε ένα σύνολο φωλιασμένων υποχώρων, μέσω κάποιας συνάρτησης προσέγγισης. Στο πλαίσιο αυτής της μελέτης, θα συναντηθεί στην ανάλυση του γονιδιώματος, με την εφαρμογή μεθόδων παλινδρόμησης σε κάθε υποχώρο.

2.4 Ταξινομητές

Οι ταξινομητές (classifiers) αντιπροσωπεύουν συνήθως το τελικό στάδιο επεξεργασίας δεδομένων. Η λήψη της απόφασης για το αν τελικά ένα στοιχείο ανήκει σε κάποια κλάση (για παράδειγμα στην περίπτωση της NA, αν ο ασθενής έχει NA ή MCI) γίνεται συνήθως από κάποιον ταξινομητή. Το πρόβλημα της απόφασης για την κλάση που ανήκει ένα στοιχείο είναι από τα βασικότερα προβλήματα στη μάθηση υπολογιστών και στη στατιστική επεξεργασία δεδομένων. Οι αλγόριθμοι ταξινόμησης συνήθως λαμβάνουν ως προϋποθέσεις τις κατηγορίες των δεδομένων, μέσω ενός πίνακα χαρακτηριστικών (feature vector) και μία ομάδα εκπαίδευσης (training set), δηλαδή ένα μέρος των πειραματικών δεδομένων που η ταξινόμησή τους είναι γνωστή και που αξιοποιείται για να εκπαιδευτεί το σύστημα στην κατάταξη των υπόλοιπων δεδομένων (επιβλεπόμενη μάθηση, supervised learning). Το μέγεθος της ομάδας εκπαίδευσης είναι από τα κυριότερα σημεία που απαιτούν προσοχή σε μία έρευνα, καθώς μία πολύ μικρή ομάδα δεν θα μπορέσει να εκπαιδεύσει το σύστημα επαρκώς, ενώ μία πολύ

μεγάλη ομάδα ως προς το σύνολο των πειραματικών δεδομένων θα οδηγήσει σε υπερπαραμετροποίηση (overfitting), δηλαδή στην υπερβολική προσαρμογή του συστήματος στα συγκεκριμένα δεδομένα, που συνεπάγεται κακή απόδοση σε μία διαφορετική ομάδα πειραμάτων. Ταυτόχρονα υπάρχουν και αλγόριθμοι μάθησης χωρίς επίβλεψη (unsupervised learning), που επιδιώκουν τη δημιουργία ενός μοντέλου κατηγοριοποίησης χωρίς κάποιες γνωστές περιπτώσεις, μία διαδικασία που από αρκετούς δεν ορίζεται ως ταξινόμηση αλλά ως ομαδοποίηση (clustering). Η ομαδοποίηση αξιοποιεί εγγενή χαρακτηριστικά των δεδομένων, όπως η απόσταση σε ένα πίνακα τιμών. Αρκετοί από τους αλγόριθμους που είδαμε στο προηγούμενο κεφάλαιο μπορούν να θεωρηθούν αλγόριθμοι ομαδοποίησης [25].

Η ταξινόμηση μπορεί να γίνει με πολλούς διαφορετικούς τρόπους. Από τις πλέον κοινές μεθόδους στην ταξινόμηση δεδομένων και αναγνώριση χαρακτηριστικών είναι οι μέθοδοι που περιλαμβάνουν τη δημιουργία ενός πυρήνα (kernel). Αυτές οι μέθοδοι περιλαμβάνουν τη μεταμόρφωση του πίνακα δεδομένων σε έναν πολυδιάστατο ευκλείδειο χώρο, όπου κάθε συντεταγμένη αντιστοιχεί σε ένα χαρακτηριστικό (feature space). Με την αύξηση του αριθμού των διαστάσεων και την προβολή των δεδομένων σε αυτές πολλές μέθοδοι μπορούν να εφαρμοστούν για την εύρεση δομών στα δεδομένα. Συνήθως οι μέθοδοι δεν υπολογίζουν την προβολή κάθε δεδομένου σε αυτό τον πολυδιάστατο χώρο, γιατί κάτι τέτοιο θα ήταν υπολογιστικά απαγορευτικό. Αντίθετα δημιουργούν συναρτήσεις (συναρτήσεις πυρήνα, kernel functions) που εντοπίζουν τις συσχετίσεις των δεδομένων στον πολυδιάστατο χώρο και επιτρέπουν στον αλγόριθμο να λειτουργεί σε αυτόν.

Οι μέθοδοι αυτές είναι αρκετά γενικές και μπορούν να αντιμετωπίσουν πολλές διαφορετικές περιπτώσεις δεδομένων. Θα τις συναντήσουμε συχνά στην πορεία.

Στην περίπτωση της ΝΑ, η ταξινόμηση συνίσταται κατεξοχήν στην κατηγοριοποίηση των δειγμάτων σε πάσχοντες από ΝΑ, πάσχοντες από εξελισσόμενο ΜCI (δηλαδή με προοπτική εξέλιξης σε ΝΑ), πάσχοντες από ΜCI με σταθερή προοπτική και υγιείς (ομάδα ελέγχου). Έτσι στα κλασικότερα προβλήματα ταξινόμησης, χρησιμοποιούνται για παράδειγμα μορφομετρικά δεδομένα του ιπόκαμπου από MRI της βάσης της ADNI, προκειμένου να διαχωρίσουν το δείγμα σε πάσχοντες από ΝΑ και σε υγιείς. Δυσκολότερο και πιο ενδιαφέρον πρόβλημα είναι ο διαχωρισμός ασθενών με στάσιμο ΜCI και ασθενών με εξελισσόμενο ΜCI. Σε οποιαδήποτε περίπτωση τα προβλήματα διαχωρισμού είναι κυρίως δυαδικά (δηλαδή η κατάταξη απαιτεί το διαχωρισμό ή σε μία κατηγορία ή σε μία άλλη) και οι επιδόσεις της ταξινόμησης στις έρευνες που θα μελετήσουμε αξιολογούνται κυρίως με τους αντίστοιχους δείκτες: την ακρίβεια (accuracy), την ευαισθησία (sensitivity) και την ειδικότητα (specificity).

Η ακρίβεια είναι ίσως ο πιο χαρακτηριστικός από αυτούς τους δείκτες και εκφράζει τον βαθμό ομοιότητας της κατάταξης που έγινε στη σωστή κατάταξη. Εδώ να σημειώσουμε ότι «σωστή κατάταξη» στην περίπτωση των μελετών της ΝΑ θεωρείται αυτή που έχει γίνει από άνθρωπο και αφορά κυρίως τη διάγνωση που έχει προκύψει από τις νευροψυχολογικές δοκιμασίες. Στην περίπτωση ενός πειράματος στη ΝΑ, η ευαισθησία μίας ταξινόμησης εκφράζει το ποσοστό των ασθενών που όντως έχουν ΝΑ που ταξινομήθηκαν ως νοσούντες από το μοντέλο υπό εξέταση. Η ειδικότητα εκφράζει το ποσοστό των πραγματικά υγιών που κατατάσσεται στους υγιείς από τον ταξινομητή. Σε έναν ιδανικό ταξινομητή και οι τρεις παραπάνω δείκτες έχουν τιμή 100%. Φυσικά στην πραγματικότητα οι τιμές είναι χαμηλότερες. Εδώ αξίζει να σημειώσουμε ότι στην περίπτωση της ΝΑ, η ευαισθησία έχει μεγαλύτερη σημασία από την ειδικότητα, δεδομένου ότι το κόστος λανθασμένης ταξινόμησης ενός υγιούς στην ομάδα νοσούντων είναι πολύ χαμηλότερο από το κόστος ταξινόμησης ενός ασθενή με ΝΑ στους υγιείς, δεδομένου ότι σε αυτή την περίπτωση η ασθένεια θα εξελιχθεί και κάθε δυνατότητα αναστροφής της πορείας θα χαθεί.

Η σχέση της ταξινόμησης με την εξαγωγή και επιλογή χαρακτηριστικών που εξετάσαμε παραπάνω δεν είναι ιδιαίτερα αυστηρή, με την έννοια ότι η στατιστική παραγοντοποίηση εξ' ορισμού καταλήγει σε κάποιου είδους κατάταξη, ενώ αντίστοιχα πολλοί ταξινομητές απαιτούν τη μελέτη και αξιοποίηση των δομών στα δεδομένα. Πολλοί αλγόριθμοι μεταμόρφωσης στο χώρο χαρακτηριστικών, όπως η γραμμική διακριτική ανάλυση, μπορούν κάλλιστα να θεωρηθούν αλγόριθμοι ταξινόμησης. Σε μία τυπική έρευνα πάντως, υπάρχει συνήθως κάποιο τελικό στάδιο ταξινόμησης, δηλαδή χρησιμοποιείται το αποτέλεσμα για παράδειγμα του αλγορίθμου PCA ως είσοδος σε κάποιον ταξινομητή, όπως σε μια μηχανή διανυσμάτων υποστήριξης, που θα βελτιστοποιήσει την ταξινόμηση.

2.4.1 Μηχανές διανυσμάτων υποστήριξης

Οι μηχανές διανυσμάτων υποστήριξης (Support Vector Machines, SVMs) είναι μοντέλα επιβλεπόμενης μάθησης και χρησιμοποιούνται κατεξοχήν για ταξινόμηση δεδομένων. Είναι η σημαντικότερη από τις μεθόδους πυρήνα (kernel methods).

Τα SVMs είναι στην πιο απλή τους μορφή δυαδικοί γραμμικοί ταξινομητές. Πρακτικά λαμβάνουν ως είσοδο την ομάδα εκπαίδευσης, όπου κάθε στοιχείο ανήκει σε μία από δύο κλάσεις. Στο στάδιο της εκπαίδευσης, ένα SVM προβάλλει τα δεδομένα ως στοιχεία στο χώρο και δημιουργεί μία επιφάνεια που διαχωρίζει με τον βέλτιστο τρόπο τα δεδομένα σύμφωνα με την κατηγορία στην οποία ανήκουν (προσπαθώντας η επιφάνεια να διαχωρίζει τις δύο κατηγορίες με τη μέγιστη δυνατή ακρίβεια και με το μεγαλύτερο κενό ανάμεσα στα στοιχεία της ομάδας εκπαίδευσης). Στη συνέχεια αξιοποιεί αυτό το μοντέλο για την ταξινόμηση των υπόλοιπων δεδομένων.

Σε περίπτωση που τα δεδομένα δεν είναι γραμμικά διαχωρίσιμα, τα SVMs προβάλλουν τα δεδομένα σε χώρο περισσότερων διαστάσεων, όπου είναι ευκολότερο να βρεθεί κάποιο υπερεπίπεδο που να τα διαχωρίζει σε μία από τις δύο κλάσεις. Αν για παράδειγμα τα δεδομένα της ομάδας εκπαίδευσης είναι τρισδιάστατα, δεν είναι γραμμικά διαχωρίσιμα αν δεν μπορεί να βρεθεί ένα δισδιάστατο επίπεδο που να τα χωρίζει ικανοποιητικά. Γι' αυτό προβάλλουν τα δεδομένα σε ένα χώρο περισσότερων διαστάσεων, όπου είναι κατανοητό διαισθητικά ότι τα δεδομένα θα είναι πιο αραιά τοποθετημένα και άρα θα μπορεί να βρεθεί ένα υπερεπίπεδο που να τα διαχωρίζει ικανοποιητικά. Η προβολή αυτού του υπερεπιπέδου στις δύο διαστάσεις δεν είναι γραμμική. Με αυτό τον τρόπο τα SVMs πραγματοποιούν μη γραμμική ταξινόμηση.

Τα SVMs όπως τα είδαμε παραπάνω είναι δυαδικοί ταξινομητές. Όμως μπορούν αξιοποιηθούν και για προβλήματα με περισσότερες κατηγορίες με τη μετατροπή ενός προβλήματος ταξινόμησης με πολλές κατηγορίες σε ένα σύνολο δυαδικών προβλημάτων (multi-class SVM). Αυτό γίνεται συνήθως είτε με την αναγωγή του προβλήματος στο διαχωρισμό κάθε κατηγορίας από το σύνολο των υπολοίπων ή στο δυαδικό διαχωρισμό κάθε κατηγορίας ως προς κάθε άλλη.

Τα SVMs έχουν χρησιμοποιηθεί εκτεταμένα στην ιατρική απεικόνιση για την ανάλυση fMRI και γενικά για τον έλεγχο μορφολογικών χαρακτηριστικών, όπως επίσης και στα συστήματα υποστήριξης αποφάσεων.

2.4.2 Αλγόριθμος ADABoost

Ο ADABoost είναι ένας μετα-αλγόριθμος ταξινόμησης που επιλέγει διαδοχικά ταξινομητές (που μπορεί κάλλιστα να είναι αδύναμοι, δηλαδή λίγο καλύτεροι από την τυχαία επιλογή) από μια βιβλιοθήκη και τους κατανέμει βάρη ανάλογα με το σφάλμα τους. Κάθε επανάληψη του ADABoost μοιράζει “προτεραιότητες” σε κάθε παράδειγμα, ώστε αν κάποιο παράδειγμα με υψηλή προτεραιότητα ταξινομηθεί λάθος, στην επόμενη επανάληψη να ταξινομηθεί σωστά. Έτσι, οι ταξινομητές προσαρμόζονται σταδιακά στα δυσκολότερα παραδείγματα.

Το σημαντικότερο πλεονέκτημα του ADABoost είναι ότι μπορεί να αξιολογήσει και να χρησιμοποιήσει πλήθος ταξινομητών, δηλαδή ότι μπορεί να χρησιμοποιηθεί χωρίς την επιλογή ταξινομητή από ανθρώπινο παράγοντα. Ένα σύστημα που επιλέγει αυτόματα τα βέλτιστα χαρακτηριστικά για ταξινόμηση ελαχιστοποιεί τον ανθρώπινο κόπο, αλλά και την ανάγκη για εξειδικευμένο προσωπικό.

Όπως προαναφέραμε ο ADABOOST αλγόριθμος προσφέρεται για αναγνώριση σχημάτων. Χαρακτηριστικό παράδειγμα αποτελεί η ακόλουθη εργασία [26], όπου ο ADABOOST εκπαιδεύεται με επιτυχία να διαχωρίζει τον ιππόκαμπο του εγκεφάλου σε sMRI εικόνες. Δεδομένου ότι ο όγκος του ιππόκαμπου είναι σημαντικός βιοδείκτης για τη ΝΑ, αυτή είναι μια εργασία που απαιτείται συχνά, όμως ως επί το πλείστον γίνεται από εξειδικευμένο προσωπικό. Η αυτοματοποίηση της συγκεκριμένης διαδικασίας επιταχύνει πολύ την κλινική μελέτη και όντως έχουν προταθεί διάφορες άλλες μέθοδοι αυτοματοποίησης, όμως ο ADABOOST προσφέρει μια οικονομική λύση στο πρόβλημα.

Το δείγμα της έρευνας επιλέχθηκε από τη βάση δεδομένων της ADNI και αποτελείται από τρισδιάστατες sMRI εικόνες του εγκεφάλου. Την ομάδα εκπαίδευσης αποτελούσαν 21 εικόνες, 7 από υγιείς, 7 από ασθενείς με MCI και 7 από νοσούντες από ΝΑ. Σε αυτές τις εικόνες ειδικοί επέλεξαν τον ιππόκαμπο. Το παραπάνω δείγμα θεωρήθηκε από τους ερευνητές ότι ήταν κατάλληλο μέγεθος για την ομάδα εκπαίδευσης, καθώς οτιδήποτε μικρότερο θα έπληττε την απόδοση του αλγορίθμου. Όλες οι εικόνες είχαν ληφθεί με το ίδιο πρωτόκολλο και κανονικοποιήθηκαν με την ίδια διαδικασία προκειμένου να εξαλειφθούν οι όποιες διαφοροποιήσεις οφείλονταν σε εξωτερικούς παράγοντες. Επιλέχθηκαν μια σειρά από χαρακτηριστικά που θεωρήθηκε ότι θα επέτρεπαν την ταξινόμηση, ενώ ταυτόχρονα επιβλήθηκαν ορισμένα όρια, προκειμένου να κρατηθεί μόνο η ευρύτερη περιοχή του εγκεφάλου που περιλαμβάνει τον ιππόκαμπο και να μην επιβαρυνθεί το σύστημα με άχρηστα voxels.

Τέλος, οι ερευνητές εφήρμοσαν τον ADABOOST μέσω μιας επέκτασης (μοντέλο αυτόματου πλαισίου, auto context model) και εκπαιδύσαν το σύστημα. Τα αποτελέσματα ήταν όντως εντυπωσιακά: η εικόνα που επιλεγόταν ήταν πολύ κοντά σε αυτήν που επιλέχθηκε από το εξειδικευμένο προσωπικό, παρότι ο ειδικός ήταν άλλος από αυτόν που εκπαιδύσε το σύστημα.

Precision, recall, relative overlap (R.O.), similarity index (S.I.), Hausdorff distance, and mean distance are reported for the leave-one-out analysis ($N=21$) when trained on rater 1 and tested on rater 2

	Precision	Recall	R.O.	S.I.	Hausdorff	Mean
Left	0.860	0.863	0.754	0.859	3.08 mm	0.00416 mm
Right	0.833	0.848	0.719	0.835	3.82 mm	0.00633 mm

We note that the Hausdorff errors are only slightly higher than in Table 3, which is to be expected as these errors incorporate differences of opinion between the two human raters.

Πίνακας 2: Επιδόσεις του ADABOOST στην επιλογή εικόνας [26]

Χαρακτηριστικό είναι ότι οι εκτιμήσεις των δύο ειδικών για την κατάτμηση διαφοροποιούνταν περισσότερο μεταξύ τους παρά με το αποτέλεσμα του αλγορίθμου. Η ομάδα συνέκρινε τα αποτελέσματα και με τα αποτελέσματα άλλων αυτοματοποιημένων μεθόδων κατάτμησης με θετικά για αυτούς αποτελέσματα:

	Recall	R.O.	S.I.
Powell et al. (2008) (Left: N=5)	0.82	0.72	0.84
Powell et al. (2008) (Right: N=5)	0.83	0.74	0.85
Fischl et al. (2002) (Left: N=134)	N.A.	~0.78	N.A.
Fischl et al. (2002) (Right: N=134)	N.A.	~0.80	N.A.
Hogan et al. (2000) (Left: N=5)	N.A.	0.7378	N.A.
Hogan et al. (2000) (Right: N=5)	N.A.	0.7578	N.A.

Πίνακας 3: Αποτελέσματα άλλων αυτοματοποιημένων μεθόδων κατάτμησης μέχρι το 2008. Ο ADABOost έχει καλύτερες επιδόσεις [26]

Επίσης οι ερευνητές δοκίμασαν αν ισχύουν καθολικά αποδεκτές για τη ΝΑ παρατηρήσεις, για παράδειγμα ότι όσο εξαπλώνεται η ασθένεια, ο όγκος του εγκεφάλου μειώνεται (δοκιμάστηκε και με sMRI μετά την πάροδο έξι μηνών στους ίδιους ασθενείς). Και σε αυτές τις περιπτώσεις ο ADABOost συμφώνησε με τις γνωστές θέσεις, επιβεβαιώνοντας την καλή του λειτουργία.

Συνολικά ο παραπάνω αλγόριθμος παρουσιάζει αρκετά καλά αποτελέσματα και, δεδομένων των κριτηρίων που εφήρμοσαν οι ερευνητές, πείθει ότι προσφέρει μία λύση για το πρόβλημα.

Ωστόσο από τότε προέκυψαν νέες καλύτερες μέθοδοι, οπότε η ίδια ερευνητική ομάδα επανεξέτασε τον αλγόριθμο το 2010 [27], καθώς είχαν προταθεί πολλές αυτοματοποιημένες μέθοδοι που χρησιμοποιούσαν SVMs σε συνδυασμό με κάποιο σύστημα επιλογής χαρακτηριστικών (πχ PCA). Οι ερευνητές πρότειναν αντί του PCA, που δεν επιλέγει τα καλύτερα χαρακτηριστικά, να δοκιμασθεί ο συνδυασμός ADABOost με εφαρμογή ενός SVM σε επόμενο στάδιο. Δοκίμασαν όντως το συγκεκριμένο σύστημα και το συνέκριναν με την προηγούμενη εργασία τους, ένα SVM όπου επιλέχθηκαν με το χέρι τα ενδιαφέροντα χαρακτηριστικά καθώς και μια άλλη πλήρως αυτοματοποιημένη μέθοδο με την ονομασία FreeSurfer.

Διαδικασία για την εκπαίδευση ενός ADA-SVM δένδρου βάθους D:

- Επιλογή των σημαντικών χαρακτηριστικών από τον αλγόριθμο ADABOOST και ταξινόμηση της ομάδας εκπαίδευσης (παραδείγματα) με αυτά τα χαρακτηριστικά μέσω ενός SVM
- Δοκιμή όλων των παραδειγμάτων μέσω του SVM ώστε να παραχθούν δύο κλάσεις, η θετική και η αρνητική
- Αν κάποιο παράδειγμα είναι στο διάστημα ανάμεσα στις δύο κλάσεις, εισάγεται και στις δύο
- Αν το δένδρο έχει βάθος μεγαλύτερο από D, ο αλγόριθμος τελειώνει
- Αλλιώς εκπαιδεύεται το θετικό "παιδί" μόνο στα δεδομένα που ο προηγούμενος κόμβος θεώρησε θετικά
- Και εκπαιδεύεται το αρνητικό "παιδί" μόνο στα δεδομένα που ο προηγούμενος κόμβος θεώρησε αρνητικά

Η διαδικασία είναι όμοια με τον αλγόριθμο του ADABOOST μόνο που χρησιμοποιείται SVM για την ταξινόμηση και όχι ο ADABOOST αλγόριθμος

Εικόνα 9: Αλγόριθμος ταξινόμησης ADABOOST σε συνδυασμό με SVM και διαφορά από απλό ADABOOST [27]

Οι ερευνητές αξιολόγησαν τις επιδόσεις των μεθόδων κατάτμησης με μια σειρά κριτηρίων. Στα περισσότερα από αυτά, είναι σαφές από τον παρακάτω πίνακα ότι ο ADABOOST μαζί με SVM πέτυχαν τις καλύτερες επιδόσεις. Ο ADABOOST μόνος του έχει λίγο χειρότερες επιδόσεις.

	Ada-SVM				AdaBoost				FreeSurfer			
	Left		Right		Left		Right		Left		Right	
	Normal	AD	Normal	AD	Normal	AD	Normal	AD	Normal	AD	Normal	AD
Precision	5.13	-5.71	0.90	-0.92	6.13	-6.98	2.29	-2.40	1.21	-1.24	-1.56	1.51
Recall	1.10	-1.12	0.17	-0.17	1.66	-1.72	1.22	-1.25	5.86	-6.63	6.29	-7.19
R.O.	5.10	-5.68	0.94	-0.96	6.25	-7.15	2.90	-3.08	5.57	-6.27	4.23	-4.62
S.I.	3.31	-3.54	0.64	-0.65	4.16	-4.54	1.94	-2.02	3.77	-4.07	2.89	-3.06
Hausdorff	-12.42	9.95	4.38	-4.80	-15.50	11.83	2.28	-2.39	-8.67	7.39	-0.68	0.67
Mean	-79.09	15.04	-70.44	0.38	-92.32	15.07	-42.73	-8.52	-46.22	-32.45	-33.73	-43.54

Πίνακας 4: Επιδόσεις ADABOOST, ADABOOST με SVM και FreeSurfer. Τα μεγέθη εκφράζουν το ποσοστό της διαφοράς σε σχέση με την κατάτμηση από άνθρωπο [27]

Όλοι οι παραπάνω αλγόριθμοι έδωσαν ακριβείς προσεγγίσεις. Η μέθοδος ADABOOST με SVM εμφάνισε ελαφρώς ακριβέστερες προσεγγίσεις από τις υπόλοιπες. Όμως ο FreeSurfer διαχειρίστηκε την ποικιλία των δεδομένων με επιτυχία και, δεδομένου ότι είχε εκπαιδευθεί από άλλη ομάδα, και άρα άλλο εργαστήριο και σαρωτή, παρήγαγε πολύ σταθερά αποτελέσματα.

Ο λόγος που το SVM με τα επιλεγμένα από άνθρωπο χαρακτηριστικά απέτυχε, φαίνεται παρακάτω. Αυτά είναι τα βέλτιστα χαρακτηριστικά που επέλεξε ο αλγόριθμος ADABOOST για την ταξινόμηση. Παρατηρούμε ότι τα χαρακτηριστικά είναι ποικίλα και δεν προκύπτει κάποιο προφανές μοτίβο για την επιλογή τους. Όντως, ο ADABOOST επέλεξε από μια τεράστια βιβλιοθήκη χαρακτηριστικών (18000) που προφανώς δεν είναι διαχειρίσιμη από έναν άνθρωπο. Παράλληλα όμως αναδεικνύεται και το πλεονέκτημα του αυτοματοποιημένου αλγορίθμου.

Left			Right		
Channel	Name	Neighborhood	Channel	Name	Neighborhood
Prior Image	Mean Filter	3,6,3	Prior Image	Mean Filter	6,7,3
Prior Image	Haar Filter	6,3,3 (3D)	Prior Image	Haar Filter	7,7,6 (3D)
Prior Image	Mean Filter	6,5,1	T1-weighted Image	Haar Filter	7,7,3 (3D)
Tissue Classification Image	Mean Filter	3,1,1	Prior Image	Standard Deviation Filter	7,6,6
Prior Image	Standard Deviation Filter	7,6,6	Prior Image	Haar Filter	5,4,5 (3D)
T1-weighted Image	Haar Filter	4,5,7 (3D)	Tissue Classification Image	Intensity	n.a.
Prior Image	Haar Filter	1,3,6 (3D)	Prior Image	Mean Filter	6,5,1
Prior Image	Haar Filter	7,7,2 (2D)	Prior Image	Gradient Filter	5,7,2 (y)
T1-weighted Image	Intensity	n.a.	T1-weighted Image	Haar Filter	3,1,7 (3D)
Tissue Classification Image	Haar Filter	3,2,4 (3D)	Prior Image	Haar Filter	5,3,1 (2D)

Πίνακας 5: Χαρακτηριστικά που επέλεξε ο αλγόριθμος ADABOOST για την επιλογή εικόνας [27]

Τέλος ο αλγόριθμος ADA-SVM σε ορισμένα σημεία εμφάνισε καλύτερα αποτελέσματα ακόμα και από τον ειδικό που επέλεξε με το χέρι τα αποτελέσματα. Όντως, ορισμένες περιοχές του ιππόκαμπου είναι δύσκολο να διαχωριστούν με απόλυτη ακρίβεια, λόγω της περίπλοκης επιφάνειας. Η μεγαλύτερη ακρίβεια, όμως, που μπορεί να προσφέρει ο αυτοματοποιημένος αλγόριθμος, επιτρέπει τη δημιουργία καλύτερων βιοδεικτών για τη νόσο.

2.4.3 Διανύσματα υπολοίπων

Στην προσπάθεια δημιουργίας πρακτικών βιοδεικτών απεικόνισης για τη ΝΑ μια ενδιαφέρουσα πρόταση έρχεται από τον David Clark [28], σε μια έρευνα που χρηματοδοτήθηκε από την ADNI. Ο Clark προτείνει τη χρήση διανυσμάτων υπολοίπων (residual vectors) για τη δημιουργία ενός ταξινομητή που θα αξιολογεί FDG-PET εικόνες και θα τις κατατάσσει ανάλογα με το στάδιο της ασθένειας.

Τα διανύσματα υπολοίπων είναι μια σχετικά απλή εφαρμογή στην κατηγορία της ανάκτησης πληροφορίας (information retrieval). Η ανάκτηση πληροφορίας αποτελεί ένα ευρύ επιστημονικό πεδίο που ασχολείται με τη γρήγορη εύρεση σχετικών στοιχείων από μεγάλες βάσεις δεδομένων. Κατεξοχήν χρησιμοποιείται σε μηχανές αναζήτησης, όπου κύριο αντικείμενο των βάσεων είναι κείμενο. Μία κλασική προσέγγιση σε αυτό το πρόβλημα είναι η ταξινόμηση του αριθμού των λέξεων πολλών κειμένων σε πίνακα και η αντιμετώπιση των στηλών και των γραμμών σαν διανύσματα. Έτσι επιτρέπεται η σύγκριση των κειμένων και η απάντηση ερωτήσεων με απλές γεωμετρικές μετρήσεις, όπως η ευκλείδεια απόσταση και η συνημιτονική ομοιότητα. Το παραπάνω μοντέλο έχει αποδειχθεί πολύ χρήσιμο και έχει επεκταθεί στην επεξεργασία εικόνων ή μουσικής.

Στην προκειμένη περίπτωση, ο ταξινομητής λειτουργεί ως εξής: δημιουργούμε το διάνυσμα-πρότυπο που προκύπτει από FDG-PET εικόνες νοσούντων από ΝΑ (ή ως δείκτη διάγνωσης εικόνες των ασθενών που ανέπτυξαν ΝΑ) και αναζητούμε στο διανυσματικό χώρο του συνόλου των εικόνων παρεμφερή διανύσματα.

Ο στόχος της έρευνας ήταν να εξετάσει αν ο παραπάνω ταξινομητής μπορεί να χρησιμοποιηθεί για διαχωρισμό ασθενών με ΝΑ ή ΜCI και αν το σύστημα μπορεί να προβλέψει την εξέλιξη ενός ασθενούς. Το δείγμα εξάχθηκε από τη βάση δεδομένων της ADNI και το αποτελούσαν 254 συμμετέχοντες.

Για να εξαχθεί το διάνυμα-πρότυπο οι ερευνητές προσπάθησαν να βρουν κάποιο διάνυμα που θα είχε υψηλή συνημιτονική ομοιότητα με τις PET εικόνες των νοσούντων από ΝΑ και χαμηλή συνημιτονική ομοιότητα με τις PET εικόνες των υπολοίπων (ουσιαστικά αποτελεί εφαρμογή της μεθόδου των ελαχίστων τετραγώνων).

Η μέθοδος πέτυχε υψηλή ευαισθησία και ειδικότητα στην ταξινόμηση της νόσου κατά τη σύγκριση ασθενών με ΝΑ και υγιών (94.8% και 86% αντίστοιχα) και δείχθηκε ότι μπορεί να χρησιμοποιηθεί για την πρόβλεψη της εξέλιξης της νόσου σε συνδυασμό με άλλους δείκτες. Εκτεταμένη εφαρμογή της στην ταξινόμηση της ΝΑ θα δούμε στην πορεία. Τα σημαντικά πλεονεκτήματα της μεθόδου πάντως είναι ότι είναι απλή σε σχέση με αντίστοιχους απεικονιστικούς δείκτες, ελάχιστα απαιτητική υπολογιστικά και μπορεί να εφαρμοστεί σε κάθε είδους εικόνες.

2.4.4 Ταξινομητής τυχαίου δάσους

Το τυχαίο δάσος (random forest) είναι μία μέθοδος εκμάθησης για ταξινόμηση δεδομένων που δημιουργεί ένα πλήθος δένδρων απόφασης στο στάδιο της εκπαίδευσης. Κάθε περίπτωση του δείγματος εφαρμόζεται σε κάθε δένδρο και υπολογίζεται πόσο κοντά είναι κάθε περίπτωση στις υπόλοιπες. Η εγγύτητα των περιπτώσεων χρησιμοποιείται για να ταξινομήσει τα δεδομένα ή για να προβλέψει τιμές (έπειτα από κάποια κανονικοποίηση).

ΚΕΦΑΛΑΙΟ 3: ΕΦΑΡΜΟΓΕΣ ΓΕΝΕΤΙΚΩΝ ΔΕΙΚΤΩΝ

3.1 Νόσος του Αλτσχάιμερ και γενετική

Η πεποίθηση ότι η ΝΑ είναι κυρίως γενετική ασθένεια είναι καλά τεκμηριωμένη. Η ασθένεια είναι κληρονομική σε ποσοστό 60-75%. Στο EOAD η ασθένεια έχει ταυτιστεί ήδη από το 1995 με τρία γονίδια (APP, PSEN1, PSEN2). Όσοι φέρουν αυτά τα γονίδια με συγκεκριμένες μεταλλάξεις θα αναπτύξουν σε μεγάλο ποσοστό ΝΑ. Όσον αφορά το LOAD, επίσης από νωρίς έγινε σαφές ότι η αλλέλη e4 στο γονίδιο APOE είναι επιβαρυντικός παράγοντας [29]. Μέχρι σήμερα αυτή θεωρείται ότι είναι η κατεξοχήν γενετική αιτία για ανάπτυξη ΝΑ και σε όλες τις σχετικές μελέτες είναι ο πλέον επιβαρυντικός δείκτης. Ωστόσο το γονίδιο APOE ερμηνεύει λιγότερο από το 30% του συνολικού πληθυσμού που εμφανίζει LOAD [30] και μέχρι πρόσφατα δεν υπήρχαν οι υποδομές για ανίχνευση γονιδίων που απλά προσθέτουν στην πιθανότητα να έχει κάποιος ΝΑ.

3.2 Συνδυασμός με απεικονιστικές τεχνικές

Η επιρροή της διακύμανσης του γενότυπου στον εγκέφαλο πρέπει με κάποιο τρόπο να μετρηθεί. Η πλέον διαδεδομένη μέθοδος χρησιμοποιεί απεικονιστικούς δείκτες για να αξιολογήσει την έκφραση των γενετικών δεικτών στον εγκέφαλο (αν και όπως θα δούμε χρησιμοποιούνται και δείκτες CSF ή ακόμη και απευθείας οι γενετικοί δείκτες). Σε αυτή την κατεύθυνση ιδιαίτερη ώθηση έχει δώσει η διαπίστωση ότι η επιρροή κάποιου γενετικού πολυμορφισμού είναι πιο προφανής σε νευροαπεικονιστικά δεδομένα παρά σε άλλα ποσοτικά δεδομένα, όπως οι νευροψυχολογικές δοκιμασίες.

Τα πρωτόκολλα της νευροαπεικόνισης που μελετήσαμε στο προηγούμενο κεφάλαιο, ιδιαίτερα αυτά που αφορούν τον όγκο του κροταφικού λοβού και του ιππόκαμπου, εδώ βρίσκουν εφαρμογή. Πράγματι, τέτοιες έρευνες απαιτούν ένα αυτοματοποιημένο πρωτόκολλο για να έχουν όλες μία κοινή ερευνητική βάση, ενώ παράλληλα, σε έρευνες τέτοιου μεγέθους ο διαχωρισμός των απεικονιστικών χαρακτηριστικών από άνθρωπο είναι εξαιρετικά επίπονη διαδικασία, λόγω του μεγάλου αριθμού των δεδομένων.

3.3 Μελέτες υποψήφιων γονιδίων

3.3.1 Βιοδείκτες για την πρόωμη ανίχνευση της νόσου του Αλτσχάιμερ

Όπως έχει ήδη επισημανθεί, η επιτυχής πρόβλεψη της εμφάνισης ΝΑ είναι ένας από τους βασικότερους στόχους στον αγώνα κατά της νόσου. Οι ασθενείς που εμφανίζουν κλινικά συμπτώματα της νόσου έχουν ήδη ανεπανόρθωτη βλάβη στον εγκέφαλό τους, ενώ παράλληλα η ανίχνευση των σταδίων που προηγούνται της νόσου εκτιμάται ότι θα βοηθήσει στην κατανόηση του μηχανισμού της. Επίσης, η ταύτιση ορισμένων γονιδίων με επιβαρυντικές για τον εγκέφαλο αλλαγές μπορεί να δώσει νέους μοριακούς στόχους για την ανάπτυξη φαρμάκων.

3.3.2 Εφαρμογή στη νόσο του Αλτσχάιμερ πρόωμης εκδήλωσης: Πρόωμες διαφοροποιήσεις

Σε αυτό το πλαίσιο έχουν γίνει πρόσφατα αρκετές έρευνες που προσπαθούν να εντοπίσουν τα προκλινικά στάδια της ΝΑ σε φορείς γονιδίων που έχουν συνδεθεί με τη ΝΑ. Μία από τις πιο αξιόλογες είναι η σχετική έρευνα του Reiman [31], όπου οι ερευνητές εντόπισαν διαφοροποιήσεις του εγκεφάλου δύο δεκαετίες πριν την προβλεπόμενη εμφάνιση της νόσου. Το μειονέκτημα της έρευνας είναι ότι χρησιμοποίησε στο δείγμα φορείς της presenilin-1 μετάλλαξης, μίας από τις μεταλλάξεις που ευθύνονται για το EOAD, την πιο σπάνια μορφή της ΝΑ. Το πλεονέκτημα είναι ότι όσοι από το δείγμα φέρουν τη μετάλλαξη θα εμφανίσουν σε μεγάλο ποσοστό τη νόσο, άρα οι νευρολογικές διαφορές που προέκυψαν μπορούν να αποδοθούν με μεγαλύτερη βεβαιότητα στη ΝΑ.

Το δείγμα περιείχε 44 φορείς και μη της μετάλλαξης, ηλικίας 22-29 ετών. Η εκτιμώμενη ηλικία εμφάνισης συμπτωμάτων EOAD είναι τα 45 έτη. Ενώ οι νευροψυχολογικές δοκιμασίες δεν έδειξαν διαφορά ανάμεσα στους φορείς της μετάλλαξης και στους μη φορείς, οι sMRI εικόνες έδειξαν μικρότερο όγκο φαιάς ουσίας στον βρεγματικό λοβό των φορέων, ενώ οι φορείς είχαν και μεγαλύτερη συγκέντρωση Αβ1-42 πρωτεΐνης. Τα παραπάνω είναι συνεπώς υποψήφιοι βιοδείκτες για την πρόωμη πρόβλεψη του EOAD. Οι παραπάνω βιοδείκτες είναι εδραιωμένοι βιοδείκτες στην έρευνα της ΝΑ, ωστόσο η μελέτη τους σε τόσο πρώιμο στάδιο υποδεικνύει ότι έχουν ρόλο και στην ανάπτυξη της νόσου.

Η επιτυχία της έρευνας οφείλεται στο ότι εντόπισε υποψήφιους βιοδείκτες σε πολύ πρώιμο στάδιο. Ωστόσο το δείγμα που χρησιμοποιήθηκε ήταν σχετικά μικρό, ενώ δεν ξέρουμε και αν το συμπέρασμα μπορεί να επεκταθεί και στο LOAD πέρα από το EOAD. Τα αποτελέσματα δεν μπορούν να συγκριθούν άμεσα με άλλες έρευνες, δεδομένου ότι η συγκεκριμένη είναι αρκετά εξειδικευμένη (χρήση δείγματος EOAD και επικέντρωση σε μικρές ηλικίες). Αντίστοιχες έρευνες στο LOAD επικεντρώθηκαν σε άλλους βιοδείκτες, όπως θα δούμε στην πορεία. Παρά τα παραπάνω η έρευνα θέτει ενδιαφέροντα ερωτήματα

για την ανάπτυξη της νόσου και θέτει ένα όριο στο πόσο νωρίς μπορούν να εντοπιστούν ίχνη ανάπτυξης ΝΑ (δηλαδή στην περίπτωση EOAD σε ηλικίες 22 με 29). Ενδιαφέρον επίσης έχει αν οι διαφορές που εντοπίστηκαν είναι νευροεκφυλιστικής φύσεως ή δημιουργούνται κατά την ανάπτυξη, δηλαδή σε ακόμη μικρότερη ηλικία.

	Hemisphere	Atlas coordinates*			p value†
		x	y	z	
Locations with greater activation during associative memory encoding and viewing of novel face-name pairs					
Hippocampus‡	Right	+33	-8	-19	<0.0001 (0.001)
Hippocampus‡	Left	-30	-16	-13	0.001
Parahippocampal gyrus‡	Right	+33	-33	+3	0.001 (0.014)
Locations with less deactivation during associative memory encoding and viewing of novel face-name pairs					
Precuneus‡	Right	+18	-57	+42	0.001 (0.008)
Precuneus‡	Left	-12	-66	+54	0.002 (0.010)
Posterior cingulate‡	Right	+21	-50	+17	0.002 (0.009)
Locations with less grey matter volume					
Parietal lobe‡	Right	+55	-18	+38	0.0004 (0.009)
Parietal lobe‡	Left	-38	-75	+48	0.0005
Temporal lobe‡	Right	+51	-27	+5	0.001
Fusiform gyrus‡	Right	+38	-61	-7	0.0008
Fusiform gyrus‡	Left	-36	-45	-8	0.0008
Parahippocampal gyrus‡	Right	+28	-24	-22	0.002
Frontal lobe	Right	+44	+14	+18	0.0006
Occipital lobe	Left	-26	-61	+21	0.002
Cerebellum	Left	-6	-55	-11	0.002

*Coordinates from Talairach's brain atlas,¹⁹ such that x is the distance in mm to the right (+) or left (-) of midline, y is the distance anterior (+) or posterior (-) to the anterior commissure, and z is the distance above (+) or below (-) a horizontal plane through the anterior and posterior commissures. †Uncorrected for multiple comparisons. Findings that remained significant after correcting for multiple comparisons in the postulated search regions are indicated by the corrected p values shown in parentheses. ‡Located in a search region affected by Alzheimer's disease.

Table 2: Location and magnitude of the most significant functional and structural brain measures in young adult PSEN1 E280A mutation carriers, when compared with non-carriers

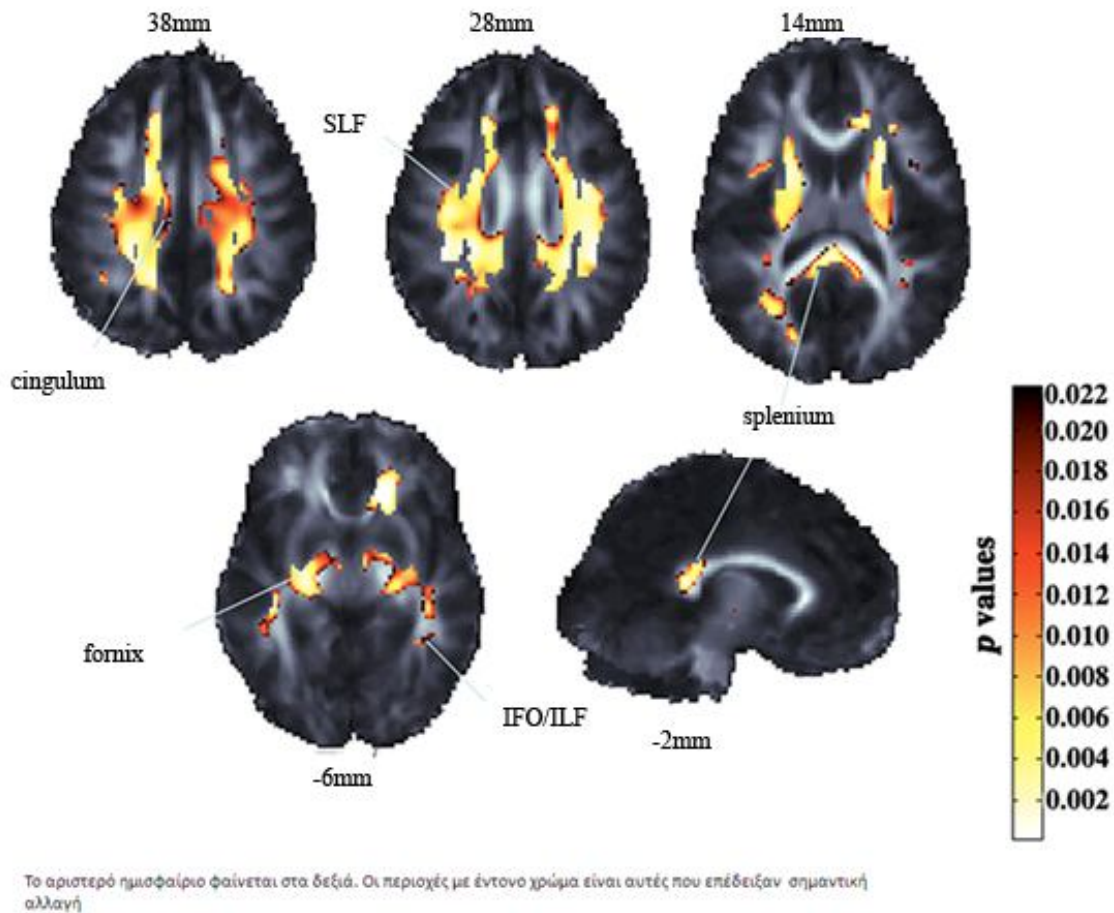
Πίνακας 6: Αλλαγές στον εγκέφαλο από το presinilin-1[31]

3.3.3 Πρώιμες διαφοροποιήσεις στη νόσο του Αλτσχάιμερ όψιμης εκδήλωσης: Αλλαγές στη μικροδομή της λευκής ουσίας

Αντίστοιχη έρευνα αλλά με στόχο το γονίδιο CLU έγινε πρόσφατα [32], με αξιόλογα αποτελέσματα. Η C-αλλέλη του CLU έχει πρόσφατα ταυτοποιηθεί ως επιβαρυντικός παράγοντας για την ανάπτυξη ΝΑ. Η διαφορά με την προηγούμενη έρευνα είναι ότι το συγκεκριμένο γονίδιο αφορά την εμφάνιση LOAD, όχι EOAD. Καθώς το LOAD αφορά τη συντριπτική πλειοψηφία των ασθενών έχει

σαφέστατα μεγαλύτερο ερευνητικό ενδιαφέρον, ωστόσο, καθώς οι φορείς του CLU έχουν πολύ μικρότερη πιθανότητα να εμφανίσουν ΝΑ από κάποιον με τις μεταλλάξεις του EOAD, τα αποτελέσματα είναι λιγότερο διαφωτιστικά για την πορεία ανάπτυξης της νόσου, ενώ μεγαλύτερο δείγμα απαιτείται για την εξαγωγή συμπερασμάτων.

Όντως, στην προκειμένη έρευνα χρησιμοποιήθηκε δείγμα 398 υγιών φορέων και μη της C-αλλέλης του CLU, με μέσο όρο ηλικίας τα 23.6 χρόνια. Από τις MRI στις οποίες υποβλήθηκαν προέκυψε ότι κάθε αντίγραφο της C-αλλέλης μείωνε τη συνοχή της λευκής ουσίας σε πολλές περιοχές του εγκεφάλου που έχουν συνδεθεί με την εμφάνιση ΝΑ. Η εξήγηση που δόθηκε είναι ότι έτσι δημιουργείται κάποια τοπική αδυναμία που επιτρέπει την επέκταση της νόσου. Δεδομένου ότι αλλαγές στη μικροδομή της λευκής ουσίας είναι μέρος της παθολογίας της ΝΑ, μπορεί να εξαχθεί η υπόθεση ότι τα δύο γεγονότα συνδέονται μεταξύ τους. Τα αποτελέσματα είναι ομολογουμένως εντυπωσιακά, δεδομένου ότι καταγράφονται διαφορές στον εγκέφαλο ατόμων που ίσως να εμφανίσουν τη νόσο 5 δεκαετίες αργότερα. Και πάλι, η συγκεκριμένη βιβλιογραφία είναι αρκετά μικρή, ώστε τα αποτελέσματα να μην μπορούν να συγκριθούν αποτελεσματικά. Η προσφορά αυτών των ερευνών εντοπίζεται κυρίως στη διερεύνηση πιθανών αιτιών ανάπτυξης της νόσου.

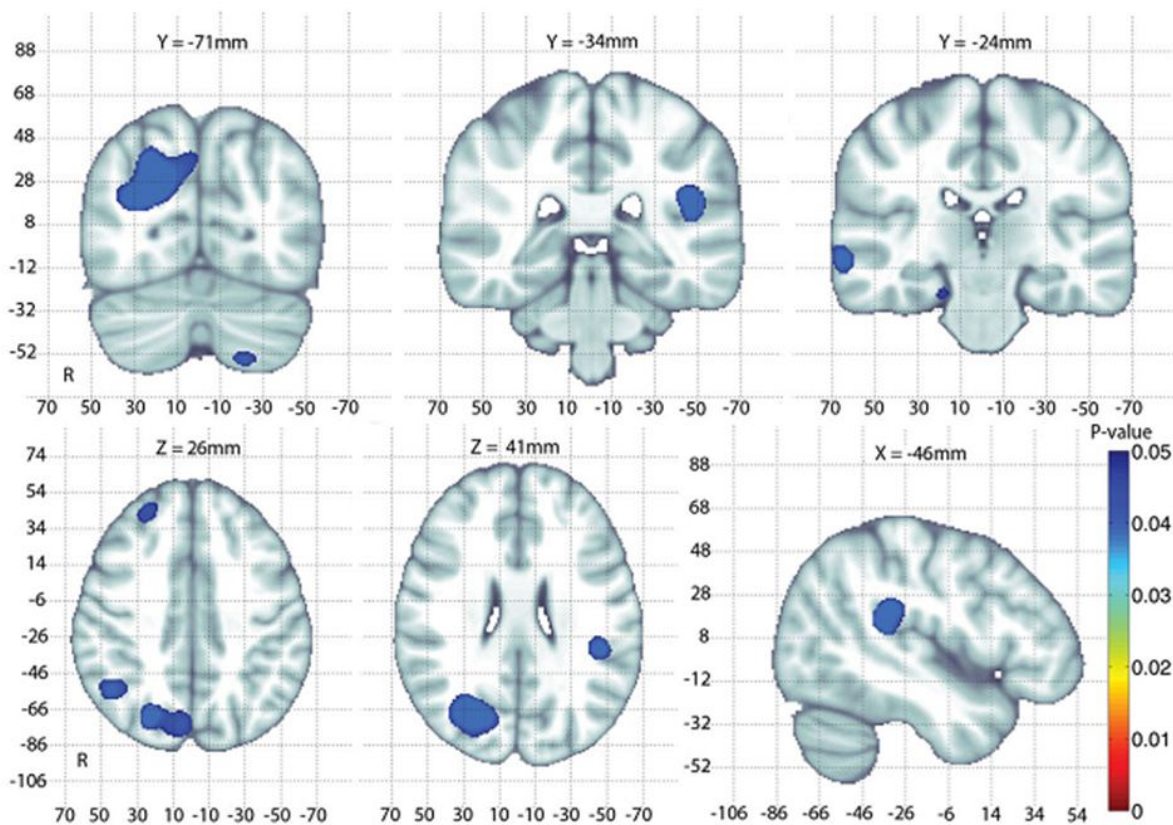


Εικόνα 10: Απεικόνιση των περιοχών του εγκεφάλου που σε μικρή ηλικία εμφανίζουν μεγαλύτερη μεταβολή στη συνοχή της λευκής ουσίας με κάθε επιπλέον *CLU-C* αλλέλη [32]

3.3.4 Πρώιμες διαφοροποιήσεις στη νόσο του Αλτσχάιμερ όψιμης εκδήλωσης: Αλλαγές στον όγκο

Μία αντίστοιχη ενδιαφέρουσα έρευνα έγινε από τους Hibar et al. [33] με θέμα την επιρροή του *GAB2* γονιδίου στον όγκο του εγκεφάλου. Ο ρόλος του *GAB2* στην παθολογία της ΝΑ είναι ήδη καλά τεκμηριωμένος, η ανίχνευση λοιπόν του ρόλου του σε έναν νεότερο ηλικιακά πληθυσμό κρίθηκε αναγκαία. Όπως το δείγμα το αποτελούσαν 755 δίδυμοι για τους οποίους ήταν διαθέσιμες sMRI εικόνες και πληροφορίες για το γενότυπό τους. Από τις T1 εικόνες MRI δημιουργήθηκαν τρισδιάστατοι χάρτες του εγκεφάλου. Από τις μετρήσεις προέκυψε ότι, πράγματι, συγκεκριμένα SNPs του *GAB2* επηρεάζουν τον όγκο περιοχών του εγκεφάλου και κυρίως του βρεγματικού λοβού, όπως επίσης και τον όγκο του εγκεφαλονωτιαίου υγρού συνολικά.

Το αξιόλογο της συγκεκριμένης έρευνας είναι ότι εκμεταλλεύθηκε την περιοριστική υπόθεση ότι το GAB2 θα διαφοροποιεί τον όγκο του εγκεφάλου σε νέους ανθρώπους, ώστε να μελετήσει την επιρροή του GAB2 σε όλο τον εγκέφαλο και όχι σε συγκεκριμένα voxels που είχαν μέχρι τώρα συνδεθεί με αυτό. Η ανάγκη για γενετικές μελέτες στο σύνολο του εγκεφάλου γίνεται όλο και πιο σαφής, καθώς οι περιοχές που επηρεάζει ένα γονίδιο μπορεί να επεκτείνονται ή να μειώνονται με την πάροδο των ετών, ή, ακόμη πιθανότερο, ένα γονίδιο μπορεί να έχει μικρή επιρροή σε κάθε μεμονωμένη περιοχή αλλά ικανή στο σύνολο της μορφολογίας του εγκεφάλου. Προς το παρόν είναι δύσκολο να διεξαχθούν γενετικές μελέτες που θα συμπεριλάβουν όλο τον εγκέφαλο και όλο το γονιδίωμα, αφενός γιατί δεν έχουν συσσωρευθεί τα απαιτούμενα δείγματα, αφετέρου γιατί το υπολογιστικό κόστος είναι τεράστιο. Η διεξαγωγή λοιπόν γενετικών ερευνών βασισμένων σε “υποψήφια” γονίδια είναι μια λύση προς αυτή την κατεύθυνση.



Στον παραπάνω χάρτη φαίνονται οι περιοχές όπου το γονίδιο GAB2 έχει τη μεγαλύτερη έκφραση στη μορφολογία του εγκεφάλου. Κυρίως επηρεάζεται ο δεξιός πλευρικός λωβός.

Εικόνα 11: Απεικόνιση των περιοχών του εγκεφάλου που ο όγκος μεταβάλλεται σε πρώιμη ηλικία από το γονίδιο GAB2 [33]

3.4 Μελέτες συσχέτισης γονιδιώματος

Όπως δείχθηκε παραπάνω, οι μελέτες υποψήφιων γονιδίων είναι ένας αποτελεσματικός τρόπος έρευνας στον τομέα της γενετικής, αλλά απαιτούν κάποια ισχυρή υπόθεση, για παράδειγμα ότι το γονίδιο APOE συνδέεται στενά με την εμφάνιση NA. Σε πιο απλές ασθένειες αυτό θα ήταν αρκετό, καθώς μπορούν να ερμηνευθούν σε κάποιο ικανοποιητικό ποσοστό, όπως το EOAD ερμηνεύθηκε από την ανακάλυψη των υπεύθυνων γονιδίων. Στην περίπτωση όμως του LOAD, η ανακάλυψη του APOE ως επιβαρυντικού παράγοντα δεν οδήγησε στην αιτιολόγηση της νόσου. Ταυτόχρονα, είναι βέβαιο ότι δεν θα βρεθούν νέα γονίδια τόσο επιβαρυντικά για την εμφάνιση LOAD όσο το APOE.

Συνεπώς χρειάζεται μια διαφορετική προσέγγιση. Το κενό αυτό έρχονται να καλύψουν οι μελέτες συσχέτισης γονιδιώματος (Genome Wide Association Studies, GWAS), μία πρόσφατη εξέλιξη της γενετικής επιδημιολογίας (η πρώτη μεγάλη σχετική μελέτη δημοσιεύτηκε μόλις το 2005). Οι GWAS είναι έρευνες που γίνονται σε όλο το γονιδίωμα και αυτό είναι και το μεγαλύτερό τους πλεονέκτημα [34].

Σε μία τυπική GWAS η διαδικασία που ακολουθείται είναι η εξής: αναλύεται το DNA των νοσούντων από την ασθένεια και συγκρίνεται με το DNA άλλων που δεν φέρουν τη νόσο (ομάδα ελέγχου, controls), μέσω SNP πινάκων. Αν, για κάποια από τις εκατομμύρια γενετικές παραλλαγές που εντοπιστούν, η συχνότητα των αλλέλων είναι σημαντικά μεγαλύτερη στους νοσούντες από ό,τι στους υγιείς, το συγκεκριμένο SNP θεωρείται ότι συνδέεται με την εμφάνιση της ασθένειας. Έτσι, περιοχές του ανθρώπινου γονιδιώματος σημειώνονται για περαιτέρω μελέτη. Αν και οι GWAS έρευνες καθεαυτές δεν εξάγουν αιτιότητα, προσφέρουν νέα ερευνητικά πεδία για ασθένειες που ήταν μέχρι τώρα δυσεξήγητες. Ακριβώς γι' αυτό, πάνω από 1200 σχετικές μελέτες έχουν γίνει μέσα σε λιγότερο από 7 χρόνια. Καθώς οι υπολογιστικές δυνατότητες των ερευνητικών κέντρων βελτιώνονται, η τάση των GWAS είναι να μεγαλώνουν, με πρόσφατες έρευνες να χρησιμοποιούν δείγματα των 200.000 ατόμων. Η στατιστική συσχέτιση πιστοποιείται καλύτερα με μεγάλης τάξης μεγέθη, καθώς ένα από τα μεγάλα προβλήματα των GWAS είναι η επαναληψιμότητα των αποτελεσμάτων. Μία άλλη τάση είναι η ανάλυση ποσοτικών φαινοτυπικών δεδομένων μέσω των GWAS, με την έμφαση να δίνεται σε ερευνητικούς βιοδείκτες (στην περίπτωση της NA σε απεικονιστικούς δείκτες).

Τα αποτελέσματα των GWAS μπορούν να αξιοποιηθούν με πολλούς τρόπους: η πρόγνωση της ασθένειας μέσω γενετικών δεικτών είναι ίσως ο πλέον προφανής. Η ταύτιση της εμφάνισης της ασθένειας με την ύπαρξη ορισμένων γενετικών παραλλαγών δημιουργεί και τους αντίστοιχους

φαρμακευτικούς στόχους. Επίσης, παράγεται και μία μηχανιστική αιτιολόγηση της ασθένειας που μπορεί να οδηγήσει στην ερμηνεία εμφάνισής της.

Οι GWAS έχουν βέβαια και αρκετά μειονεκτήματα. Το σημαντικότερο από αυτά είναι η στατιστική περιπλοκότητα του εγχειρήματος [35]. Ο τεράστιος αριθμός στατιστικών δοκιμών που χρειάζεται μπορεί πολύ εύκολα να οδηγήσει σε λανθασμένες πρακτικές και πολλές έρευνες έχουν καταλήξει σε μη επαναλήψιμα συμπεράσματα. Αυτό φυσικά θα αντιμετωπιστεί με την εξέλιξη των στατιστικών προτύπων, που θα μειώσουν το βάρος εργασίας των ερευνητών. Τέλος, το εξ ορισμού πολύ μεγάλο πεδίο έρευνας μίας GWAS δεν αποδίδει πάντα καλά αποτελέσματα, αλλά, μέχρι τώρα, έχει δώσει ενδιαφέροντα αποτελέσματα σε αρκετές δυσεπίλυτες ασθένειες.

Αυτή είναι και η περίπτωση της NA, όπου οι GWAS εντόπισαν νέους γενετικούς στόχους που έχουν ταυτοποιηθεί αρκετές φορές [36]. Η συμβολή τους δείχθηκε αναμφισβήτητα σημαντική, αν και η επέκταση αυτών των ερευνών προς καινούριες κατευθύνσεις έχει ήδη αρχίσει να απαιτείται από την επιστημονική κοινότητα. Οι περισσότερες GWAS στη NA αξιοποιούν χάρτες γενετικής έκφρασης (πχ ο Allen εγκεφαλικός χάρτης γονιδιακής, Allen Brain Atlas for gene expression) για να συνδέσουν περιοχές του γονιδιώματος με φαινότυπους από MRI (ή κάποια άλλη απεικονιστική μέθοδο). Ωστόσο οι έρευνες δεν περιορίζονται μόνο σε MRI ή απεικονιστικά δεδομένα, αλλά περιλαμβάνουν και δείκτες CSF ή και την αξιολόγηση καθαρά γενετικών δεικτών. Η αξιολόγηση της επιρροής των γονιδίων γίνεται με βάση την επιρροή τους σε αυτούς τους βιοδείκτες και συνήθως είναι μικρή. Χαρακτηριστικά, σε μερικούς από τους πλέον έγκυρους βιοδείκτες, όπως η μορφομετρία του βρεγματικού λοβού, η συνολική μείωση του όγκου από την επιρροή όλων των επιβαρυντικών SNPs δεν ξεπερνά το 15%, ενώ το 10% περίπου από αυτό αποδίδεται συνήθως στο γονίδιο APOE. Προφανώς το ποσοστό που μένει (5%) είναι πολύ μικρό μέγεθος και αυξάνει τη δυσκολία του εγχειρήματος, ωστόσο είναι και ο κύριος τρόπος αξιολόγησης της σημασίας των αποτελεσμάτων. Αυτός είναι και ο λόγος που διαφορετικοί βιοδείκτες μελετώνται και άλλες μέθοδοι επεξεργασίας εξετάζονται.

Το πιο κοινό πρόβλημα στην εφαρμογή των GWAS στη NA είναι ότι η επιτυχημένη εφαρμογή τους απαιτεί πολύ μεγάλο δείγμα. Εκ των πραγμάτων, από τη στιγμή που αξιοποιούνται απεικονιστικές μέθοδοι, το δείγμα για τη NA δεν μπορεί να είναι πολύ μεγάλο. Ακόμα και η βάση της ADNI, που είναι η μεγαλύτερη στο χώρο, θεωρείται μικρή για τις απαιτήσεις των GWAS. Απαιτείται συνεπώς ένα πολύ καλά δομημένο δείγμα και εξαιρετικά λεπτομερείς στατιστικοί έλεγχοι. Ακόμα και έτσι όμως, τα δεδομένα που προκύπτουν από μία έρευνα συχνά ανατρέπονται και χρειάζονται πολλές μελέτες για να γίνει αποδεκτό ότι όντως κάποιο SNP σχετίζεται με την εμφάνιση της NA.

Καθίσταται λοιπόν σαφές, ότι απαιτούνται μεγαλύτερα μεγέθη δείγματος για να μπορούν να διασταυρωθούν τα αποτελέσματα. Προς το παρόν, οι ερευνητές διαχωρίζουν συνήθως σε τρεις ή τέσσερις ομάδες το δείγμα (υγιείς, εξελισσόμενο και μη MCI, NA) αλλά δύο σημαντικοί παράγοντες δεν μπορούν να αποκλειστούν χωρίς μεγαλύτερα δείγματα: πρώτον, η γενετική διαστρωμάτωση του πληθυσμού (ουσιαστικά η πιθανή διαφορετική συχνότητα κάποιων SNPs, εξαιτίας της διαφορετικής καταγωγής) που επηρεάζει τη συχνότητα και άρα τη επιρροή κάποιων SNPs (αν και η βάση της ADNI περιέχει κατεξοχήν καυκάσιους, οπότε αυτός ο παράγοντας εν μέρει αντιμετωπίζεται με τον αποκλεισμό των υπολοίπων από το δείγμα) και, δεύτερον, το γεγονός ότι πολλοί ηλικιωμένοι εμφανίζουν νευροαπεικονιστικά συμπτώματα NA, χωρίς αντίστοιχα νοητικά συμπτώματα. Αυτό δυσκολεύει το διαχωρισμό των ασθενών ηλικιωμένων από τους υγιείς με καθαρά νευροαπεικονιστικά κριτήρια. Ωστόσο πολλά ενδιαφέροντα αποτελέσματα έχουν προκύψει και, με την εξέλιξη των στατιστικών αλγορίθμων και τη μεγέθυνση του διαθέσιμου δείγματος, η κατάσταση θα βελτιωθεί.

Προς το παρόν πάντως το μόνο γονίδιο που τα SNPs του εμφανίζονται σε όλες τις έρευνες είναι το APOE. Το συγκεκριμένο γονίδιο χρησιμοποιείται λοιπόν σαν έλεγχος για τον καλό ορισμό της έρευνας. Δεδομένου ότι οι GWAS είναι σχετικά λίγες στη NA, η αξιολόγησή τους είναι αρκετά δύσκολη. Αυτή τη στιγμή πολλές διαφορετικές μέθοδοι αξιοποιούνται για τη διεξαγωγή τους που παράγουν αντίστοιχα διαφορετικά αποτελέσματα. Ο αποτελεσματικότερος έλεγχος για την ευστάθεια των αποτελεσμάτων είναι αφενός η διασταύρωση των συμπερασμάτων με άλλες GWAS (δηλαδή αν κάποια SNPs έχουν εντοπιστεί ως υποψήφια και από άλλους) και γενικότερα με τη γενετική βιβλιογραφία της NA (δηλαδή αν άλλες έρευνες εκτός των GWAS έχουν συνδέσει κάποιο γονίδιο με τη NA ή ακόμη και αν ο γνωστός ρόλος του γονιδίου δικαιολογεί σύνδεση με την ανάπτυξη της NA).

3.4.1 Πρώτη εφαρμογή μελέτης συσχέτισης γονιδιώματος στη λειτουργία του εγκεφάλου: Διάγνωση σχιζοφρένειας

Η ομάδα που είχε μελετήσει τον αλγόριθμο του parallel ICA, επέκτεινε τη μέθοδο και την εφήρμοσε σε δείγμα ασθενών με σχιζοφρένεια [37]. Η εφαρμογή του paraICA στα GWAS είναι όντως κατάλληλη: ουσιαστικά δύο ομάδες δεδομένων χρησιμοποιούνται, οι fMRI εικόνες και τα γενετικά δεδομένα. Επιδιώκεται τα παραπάνω να διασπαστούν στα βασικά τους μέρη και να βρεθούν οι πιθανές αλληλεπιδράσεις τους. Για την περίπτωση του GWAS κάτι τέτοιο είναι η σχέση ανάμεσα σε ένα SNP και κάποιο voxel.

Τα δεδομένα παρήχθησαν από 20 ασθενείς με σχιζοφρένεια και 43 υγιείς και αφορούσαν εξετάσεις fMRI που λήφθηκαν κάτω από συγκεκριμένα ακουστικά ερεθίσματα και στοιχεία για το γενότυπο. Τα

δεδομένα υπέστησαν επεξεργασία με πολλαπλή γραμμική παλινδρόμηση. Όντως βρέθηκε συσχέτιση ανάμεσα σε περιοχές του εγκεφάλου και συγκεκριμένα γονίδια που έχουν και στο παρελθόν συνδεθεί με τη σχιζοφρένεια. Ο αλγόριθμος του paraICA έδωσε σταθερά αποτελέσματα σε διαφορετικά επίπεδα συνδεσιμότητας και ενθαρρύνει την εφαρμογή του στο μέλλον.

Η παραπάνω εργασία είναι από τις πρώτες GWAS που έγιναν. Δεδομένου ότι η σχιζοφρένεια έχει υψηλή κληρονομικότητα, η εφαρμογή του GWAS σε αυτόν τον τομέα είναι σίγουρα κατάλληλη. Τα αποτελέσματα που παρήχθησαν ενθάρρυναν την διενέργεια πολλών αντίστοιχων μελετών, ιδιαίτερα στη ΝΑ, όπως θα δούμε στη συνέχεια.

3.4.2 Πρώτη μεγάλη μελέτη συσχέτισης γονιδιώματος στη νόσο του Αλτσχάιμερ: Σημασία στατιστικής επεξεργασίας

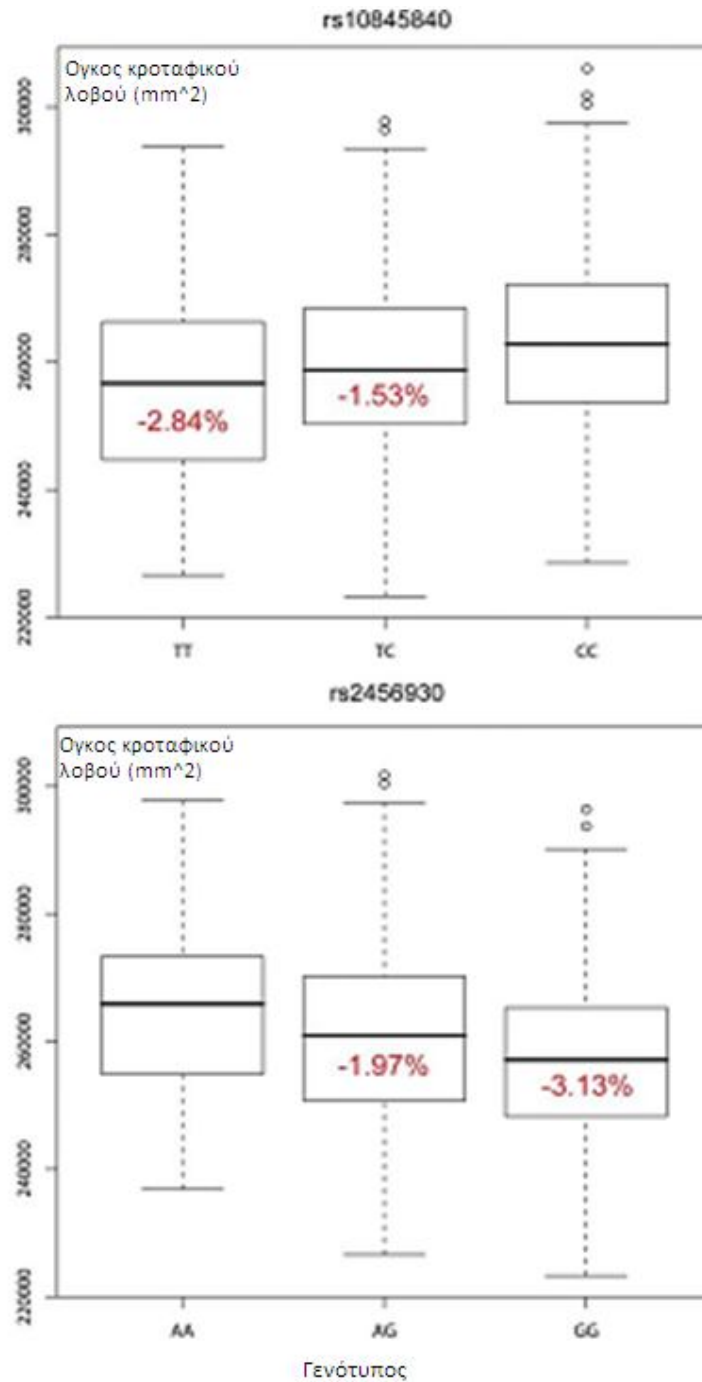
Μετά την επιτυχία των GWAS σε άλλες ασθένειες, άρχισαν να εφαρμόζονται και στη ΝΑ. Μία τυπική προσπάθεια είναι η μελέτη του Stein το 2010 [38]. Οι ερευνητές αξιοποίησαν τη βάση της ADNI για την εξαγωγή τρισδιάστατων ογκομετρικών στοιχείων του κροταφικού λοβού και του ιππόκαμπου (TBM και μοντέλο αυτόματου πλαισίου από τις sMRI εξετάσεις). Το δείγμα διαχωρίστηκε σε ασθενείς με ΝΑ, ασθενείς με MCI και υγιείς, ανάλογα με την επίδοσή τους στο MMSE. Οι ερευνητές χρησιμοποίησαν περισσότερους από 500.000 γονιδιακούς δείκτες (SNPs) και αποτίμησαν τη στατιστική επιρροή που είχε ο καθένας στα voxels της απεικόνισης του λοβού. Η μελέτη αυτή είναι από τις πρώτες μεγάλης κλίμακας GWAS που έγιναν για τη ΝΑ.

Καθώς ο αριθμός των SNPs είναι τεράστιος, οι περισσότερες μελέτες προσπαθούν να εντοπίσουν περιοχές του γονιδιώματος που συνδέονται στατιστικά με τα χαρακτηριστικά που τις ενδιαφέρουν. Στην προκειμένη περίπτωση, οι ερευνητές εντόπισαν δύο περιοχές του γονιδιώματος που συνδέονταν με τον όγκο του κροταφικού λοβού. Επικεντρώθηκαν συνεπώς σε αυτές τις περιοχές και μελέτησαν διεξοδικότερα τις συχότητες των αλλέλων στους ασθενείς με ΝΑ και στους υγιείς. Παρατηρούμε ότι ουσιαστικά η GWAS έδωσε κάποιο στόχο, ο οποίος ερευνήθηκε σε δεύτερο στάδιο σαν μια μελέτη υποψήφιων γονιδίων.

Η στατιστική ανάλυση χρησιμοποίησε τον κλασικό αλγόριθμο παλινδρόμησης. Αφού έβγαλαν τα αποτελέσματα, και για να βεβαιωθούν ότι η ομαδοποίηση του δείγματος δεν επηρεάζει την ανάλυση, οι ερευνητές χρησιμοποίησαν έναν αντιμεταθετικό αλγόριθμο, που ουσιαστικά αντάλλαξε τυχαία φαινότυπους ανάμεσα σε αντικείμενα κάθε ομάδας του δείγματος. Στη συνέχεια εφήρμοσαν πάλι τη στατιστική ανάλυση και, αφού ακολούθησαν αυτή τη διαδικασία αρκετές φορές, συγκράτησαν τα SNPs που εμφάνισαν στατιστική επιρροή συνολικά.

Παρότι το δείγμα προέρχεται από τη βάση δεδομένων της ADNI που είχε μαζέψει στοιχεία από διαφορετικές περιοχές της Βορείου Αμερικής, οι ερευνητές θέλησαν να ελέγξουν και για την πιθανή επιρροή της γενετικής συγγένειας των αντικειμένων του δείγματος. Συνέταξαν λοιπόν ένα γραμμικό μοντέλο που αποτιμούσε την επιρροή των παραπάνω SNPs, λαμβάνοντας υπόψη του έναν πίνακα συγγένειας του δείγματος. Εδώ να σημειώσουμε ότι οι ερευνητές είχαν ήδη αφαιρέσει από το δείγμα της ADNI τους πληθυσμούς που δεν είχαν περιγραφεί ως καυκάσιοι (που είναι ο μεγαλύτερος πληθυσμός της ADNI), για να ελέγξουν την επιρροή της διαστρωμάτωσης στο δείγμα. Ελέγχοντας το δείγμα, οι ερευνητές βεβαίωσαν ότι τα αντικείμενα όντως είχαν τον ίδιο ευρύτερο γενότυπο.

Τελικά, μετά τους παραπάνω ελέγχους οι ερευνητές συγκράτησαν δύο SNPs που εμφάνισαν επαρκή στατιστική επιρροή. Και τα δύο SNPs αφορούσαν τον όγκο του κροταφικού λοβού και βρίσκονταν εντός του GRIN2B γονιδίου (το συγκεκριμένο γονίδιο έχει συνδεθεί με τη λειτουργία της μνήμης). Η επιρροή που είχαν τα παραπάνω SNPs στον όγκο του λοβού φαίνεται στις ακόλουθες γραφικές παραστάσεις. Μερικά άλλα υποψήφια γονίδια συγκρατήθηκαν επίσης για μελλοντική μελέτη, παρότι είχαν μικρότερη επιρροή (RNF220, UTP20, KIAA0743).



Τα παραπάνω σχήματα δείχνουν τον όγκο του κροταφικού λοβού οργανωμένο κατά τα SNPs με τη μεγαλύτερη επιρροή. Τα κουτιά αντιπροσωπεύουν τις ετεροζυγείς ή ομοζυγείς αλλέλες για τα δύο SNPs. Οι κόκκινοι αριθμοί δείχνουν τη μείωση του όγκου του λοβού από τα συγκεκριμένα SNPs σε σχέση με το μέσο όρο των υπολοίπων.

Διάγραμμα 1: Μεταβολή στον όγκο του κροταφικού λοβού με την παρουσία δύο SNPs με τη μεγαλύτερη επιρροή [38]

Είναι κοινά αποδεκτό ότι η e4 αλλέλη του APOE έχει μετρήσιμη επιρροή στην εμφάνιση της ΝΑ. Σε αυτό το πλαίσιο, οι ερευνητές αποφάσισαν να αποτιμήσουν την επίδραση των παραπάνω SNPs στην ομάδα του δείγματος που έφερε την e4 αλλέλη. Στα άτομα που έφεραν τη συγκεκριμένη αλλέλη, έπειτα από την εφαρμογή ενός μοντέλου πολλαπλής παλινδρόμησης, οι ερευνητές κατέληξαν ότι η αλληλεπίδραση των δύο SNPs με την e4 αλλέλη δεν ήταν ισχυρή και δεν ήταν, άρα, υπεύθυνη για την επιρροή που εμφάνισαν τα δύο SNPs.

Τα παραπάνω SNPs εμφανίζονται να έχουν επιρροή από 1.5% έως 3% στον όγκο του κροταφικού λοβού ανάλογα με τα αντίτυπα των αλλέλων. Μία τέτοια επιρροή είναι όντως σημαντική, δεδομένου ότι στον κύριο επιβαρυντικό παράγοντα, APOE-ε4, αποδίδεται επιρροή μικρότερη του 10% στον ίδιο βιοδείκτη (όγκος του βρεγματικού λοβού), όπως θα δούμε στην πορεία. Επίσης, αυτά τα SNPs έχουν δοκιμαστεί εκτεταμένα στο πλαίσιο αυτής της έρευνας και σίγουρα αποτελούν υπονήφιους στόχους για μελλοντική μελέτη. Ωστόσο, αξιοποιώντας τους πίνακες αλληλεπίδρασης των SNPs, και άλλοι πιθανοί στόχοι μπορεί να ανιχνευθούν, που ενώ δεν πέρασαν το προβλεπόμενο όριο, είχαν αρκετά μεγάλη στατιστική επιρροή.

Συνολικά, είδαμε ότι οι ερευνητές έδωσαν μεγάλο βάρος στην καλή στατιστική επεξεργασία των δεδομένων και εφήρμοσαν πολλούς ελέγχους προκειμένου να ελέγξουν τα στοιχεία τους. Όντως, δεδομένου του τεράστιου αριθμού των στόχων (πάνω από 500.000 SNPs και ένας αντίστοιχος αριθμός από voxels), το δείγμα της ADNI φαντάζει σχετικά μικρό, συνεπώς απαιτούνται εκτεταμένοι στατιστικοί έλεγχοι. Η συγκεκριμένη εργασία αποτελεί μια πρώτη προσπάθεια προς αυτή την κατεύθυνση. Το γονίδιο GRIN2B εξακολουθεί να συνδέεται με τη ΝΑ, παρότι τα SNPs που εντόπισαν δεν επιβεβαιώθηκαν στις περισσότερες ακόλουθες μελέτες, και οι ίδιοι οι ερευνητές εξήγαγαν άλλα αποτελέσματα, όταν ακολούθησαν διαφορετική στατιστική προσέγγιση. Τέλος, οι ερευνητές κατέληξαν ότι ο όγκος του ιππόκαμπου δεν προσφέρεται σαν δείκτης στο πλαίσιο της γενετικής απεικόνισης, ίσως γιατί η κληρονομικότητα των χαρακτηριστικών του είναι μικρότερη. Παρόλα αυτά, ο συγκεκριμένος δείκτης έχει αξιοποιηθεί με επιτυχία σε άλλες GWAS για τη ΝΑ και αυτό είναι ένα μάλλον λανθασμένο συμπέρασμα.

3.4.3 Νέοι δείκτες σε μελέτη συσχέτισης γονιδιώματος: Florbetapir-PET

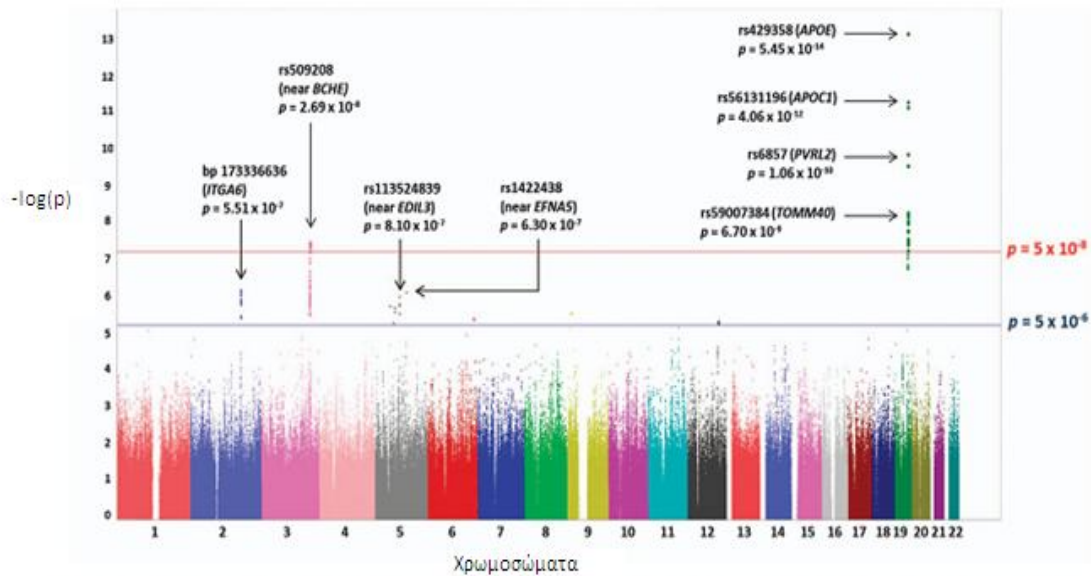
Έχουμε ήδη αναφέρει ότι οι GWAS δεν περιορίζονται μόνο στα MRI δεδομένα. Συγκεκριμένα για τη ΝΑ, άλλοι δείκτες, όπως η συγκέντρωση tau στο CSF ή η συσσώρευση αμυλοειδικών-βήτα πεπτιδίων στον εγκέφαλο έχουν απασχολήσει την ερευνητική κοινότητα. Αυτός ο τελευταίος δείκτης χρησιμοποιήθηκε εκτεταμένα στην ακόλουθη έρευνα [5].

Ένα από τα κεντρικά χαρακτηριστικά της παθολογίας της ΝΑ είναι η δημιουργία Αβ νευροϊνιδιακών πλακών. Αυτό δεν συναντάται μεν κατά αποκλειστικότητα στη ΝΑ (πολλοί υγιείς ηλικιωμένοι ή πάσχοντες από άλλες μορφές άνοιας έχουν όμοιες δομές στον εγκέφαλο), αλλά θεωρείται ότι είναι μία από τις αιτίες ανάπτυξης της ασθένειας. Συνεπώς η πυκνότητα των νευροϊνιδιακών πλακών έχει ερευνητικό ενδιαφέρον για τη ΝΑ.

Σε αυτόν τον τομέα, πρόσφατα υπήρξαν σημαντικές εξελίξεις: ο ραδιοϊχνηλάτης florbetapir μεγεθύνει πολύ την ικανότητα της PET να διακρίνει τις νευριτικές πλάκες, λόγω της τάσης του να συνδέεται με αυτές και της μεγάλης ημιζωής του. Δεδομένου ότι ο δείκτης έχει εγκριθεί για διαγνωστικούς σκοπούς με αρκετά υποσχόμενα αποτελέσματα (εγκρίθηκε από την FDA το 2011), ήταν θέμα χρόνου να αξιοποιηθεί σε κάποιο GWAS. Η μεγαλύτερη GWAS που αξιοποιεί το συγκεκριμένο δείκτη δημοσιεύτηκε πρόσφατα με ενδιαφέροντα συμπεράσματα.

Οι επιστήμονες χρησιμοποίησαν την πυκνότητα των Αβ-πλακών ως δείκτη για την εξέλιξη της ΝΑ και εξήγαγαν τις αντίστοιχες γενετικές συσχετίσεις. Ουσιαστικά, η ομάδα υπολόγισε τον μέσο όρο της αναλογίας συγκέντρωσης της ακτινοβολίας στον εγκέφαλο ως ολικό ποσοτικό φαινότυπο. Χρησιμοποιήθηκε η βάση δεδομένων της ADNI με δείγμα 555 PET εξετάσεων και στα δεδομένα εφαρμόστηκε γραμμική παλινδρόμηση.

Καθώς είναι γνωστό ότι το πιο επιβαρυντικό γονίδιο για τη ΝΑ, το APOE, καθορίζει άμεσα τη συσσώρευση των νευροϊνιδιακών πλακών, οι επιστήμονες περίμεναν το APOE να έχει τη μέγιστη επιρροή. Όντως, η αλλέλη APOE-e4 είχε τη μεγαλύτερη συσχέτιση. Πολλά SNPs με σημαντική επιρροή είχε και το γονίδιο BCHE, και η συσχέτιση παρέμεινε ισχυρή ανεξάρτητα από το APOE (η APOE-e4 εντάχθηκε στο σύστημα ως συντελεστής σε δεύτερο στάδιο, καθώς αλληλεπιδρά με πολλά γονίδια και η παρουσία του επηρεάζει τα αποτελέσματα). Εδώ να σημειωθεί ότι το BCHE έχει προταθεί στο παρελθόν σαν επιβαρυντικός δείκτης για τη ΝΑ και έχει συνδεθεί με τη δημιουργία νευροϊνιδιακών πλακών. Αρκετά ακόμη SNPs βρέθηκαν κοντά στο να έχουν στατιστική σημασία, αλλά δεν πέρασαν το όριο που είχε τεθεί.



Πάνω από 6.000.000 SNPs δοκιμάστηκαν για πιθανή σχέση με τη συγκέντρωση Αβ. Τα SNPs που είχαν στατιστική σημασία σε σχέση με όλο το γονίδιο είναι αυτά που ξεπερνούν τη κόκκινη γραμμή. Τα περισσότερα είναι στο χρωμόσωμα 19, δηλαδή κοντά ή μέσα στο APOE. Τα SNPs που υπήρξε απλά ένδειξη ότι σχετίζονται είναι όσα ξεπερνούν τη γαλάζια γραμμή.

Διάγραμμα 2: τα SNPs με τη μεγαλύτερη επιρροή από το florbetapir GWAS [5]

Ακόλουθες αναλύσεις δεν βρήκαν κάποια συσχέτιση στην επιρροή του BCHE και του APOE και σε παράγοντες όπως η ηλικία, το φύλο ή η μόρφωση, οπότε οι παραπάνω παράγοντες θεωρήθηκαν ότι έχουν ανεξάρτητη και προσθετική λειτουργία στη συσσώρευση των Αβ πρωτεϊνών. Στη μετέπειτα γραμμική παλινδρόμηση που εφαρμόστηκε, παράγοντες όπως μόρφωση, φύλο και ηλικία ερμήνευσαν συνολικά μόλις το 2.7% της διακύμανσης των τιμών της πυκνότητας Αβ, ενώ το APOE-ε4 ήταν υπεύθυνο για ένα επιπλέον 10.7% της διακύμανσης και το BCHE για ένα 4.3%. Συνολικά, δηλαδή, οι δύο αυτοί γενετικοί δείκτες εξηγούσαν περίπου το 15% της διακύμανσης της συγκέντρωσης Αβ, που είναι μεγάλο ποσοστό για τα δεδομένα των GWAS στη ΝΑ, που συνήθως κινείται κάτω από το 5% για όλα τα γονίδια εκτός του APOE. Αυτό το αποτέλεσμα, αν επιβεβαιωθεί και από άλλους ερευνητές, πείθει για την αξία της florbetapir-PET σαν δείκτη. Οι μορφομετρικοί δείκτες μέχρι τώρα είναι ακριβέστεροι από δείκτες όπως το FDG-PET για τη συγκέντρωση Αβ, αλλά το florbetapir-PET μπορεί να αλλάξει άρδην την κατάσταση, υπό τον όρο ότι η ταύτισή του με την πορεία της νόσου κριθεί επαρκώς σε βάθος χρόνου. Επίσης, το florbetapir-PET μπορεί να δειχθεί ενδιαφέρων δείκτης και για την παρακολούθηση της εξέλιξης της ΝΑ.

Οι επιστήμονες καταλήγουν ότι θα ήθελαν να επαναλάβουν την έρευνα με περισσότερα δεδομένα, καθώς πιστεύουν ότι αρκετά SNPs θα αποκτούσαν στατιστική σημασία με μεγαλύτερο δείγμα. Αυτό είναι ένα πάγιο ζήτημα των ερευνών GWAS στη ΝΑ. Τέλος να επισημανθεί ότι ένα πιθανό μειονέκτημα

της έρευνας είναι ότι εφάρμοσε μαζική γραμμική παλινδρόμηση, δηλαδή εξέτασε την επιρροή κάθε SNP χωριστά. Όπως θα δούμε στην πορεία, η τάση για τα μορφομετρικά GWAS είναι η χρήση πολυπαραγοντικών τεχνικών, που αξιοποιούν την αλληλεπίδραση των SNPs και αυξάνουν τη στατιστική ισχύ των ερευνών. Δεν γνωρίζουμε σε αυτό το στάδιο, αν κάτι τέτοιο θα είχε νόημα στην περίπτωση της florbetapir-PET (η αυξημένη στατιστική σημασία SNPs σε περιοχές κοντά στο APOE και στο BCHE δίνει μια τέτοια ένδειξη, δηλαδή η γειννίαση φαίνεται να υποδεικνύει έκφραση με αντίστοιχο τρόπο), ωστόσο είναι ένα θέμα που θα άξιζε να διερευνηθεί σε επόμενη έρευνα.

3.4.4 Πέρα από τους απεικονιστικούς δείκτες: Εκτεταμένη μελέτη συσχέτισης γονιδιώματος για δείκτες του εγκεφαλονωτιαίου υγρού

Τα GWAS όμως δεν περιορίζονται στη συσχέτιση απεικονιστικών δεικτών με γονίδια. Ενδιαφέρον έχει και η ταύτιση της επιρροής των γονιδίων σε άλλους δείκτες, όπως οι δείκτες του CSF. Έχουμε ήδη προαναφέρει ότι οι βασικοί δείκτες του CSF για τη NA είναι τρεις: η συνολική συγκέντρωση tau, η συγκέντρωση p-tau και η συγκέντρωση Αβ42 πρωτεΐνης. Συνολικά αυτοί οι δείκτες είναι αρκετά έγκυροι και προσφέρονται για GWAS, δεδομένου ότι είναι πιο απλοί από τους μορφομετρικούς δείκτες (πχ μετράται ο μέσος όρος της Αβ συγκέντρωσης αντί για τα 10^5 voxels που υπάρχουν σε ένα τυπικό απεικονιστικό GWAS). Σε οποιαδήποτε περίπτωση, οι απεικονιστικές τεχνικές εξακολουθούν να παράγουν τους πιο ακριβείς δείκτες, αλλά σίγουρα οι CSF δείκτες προσφέρουν διαφορετική πληροφορία από τους απεικονιστικούς δείκτες, όπως διαπιστώνεται σε έρευνα συστημάτων απόφασης που θα δούμε στην πορεία, και άρα έχουν κάποια αξία.

Σε αυτό το πλαίσιο, η ακόλουθη έρευνα είναι αρκετά αξιόλογη: οι ερευνητές πραγματοποίησαν το μεγαλύτερο GWAS για CSF που έχει γίνει ποτέ, με δείγμα 1269 άτομα [39]. Για να το πετύχουν αυτό χρησιμοποίησαν τέσσερις βάσεις δεδομένων για τη NA (ADNI, Knight-ADRC, Consortium of Alzheimer's Disease Centers (University of Washington και University of Pennsylvania)). Οι ερευνητές ομαλοποίησαν τα δεδομένα των διαφορετικών βάσεων και κανονικοποίησαν για παράγοντες όπως γένος ή πληθυσμιακή διαστρωμάτωση (και δεδομένου ότι έχουμε έναν απλό δείκτη είναι σαφέστατα ευκολότερο να ομολοποιηθούν τα δεδομένα σε σχέση με τους διαφορετικούς σαρωτές και πρότυπα προεπεξεργασίας που υπάρχουν στις απεικονιστικές βάσεις δεδομένων). Οι ομάδες της έρευνας ήταν δύο, ασθενείς με NA και υγιείς. Χρησιμοποιήθηκε γραμμική παλινδρόμηση για να αποτιμηθεί η επιρροή κάθε ξεχωριστού SNP στους δείκτες. Συνολικά ερευνήθηκαν πάνω από 5.500.000 SNPs.

Για να αποτιμήσουν τη διακύμανση (variance) στις τιμές των δεικτών που ερμηνεύεται από τα διαθέσιμα SNPs, η ομάδα εφάρμοσε μία τεχνική διαμερισματοποίησης του γονιδιώματος (πρόκειται για

τεχνικές που εμπλουτίζουν περιοχές ενδιαφέροντος στο γονιδίωμα, αντίστοιχες με την αναδιάταξη του γονιδιώματος αλλά πιο περιορισμένες και με μικρότερο κόστος υπολογιστικά). Κατέληξαν ότι περίπου 7% της απόκλισης για το επίπεδο της total-tau και 15% για την p-tau ερμηνεύεται από το διαθέσιμο γενότυπο. Δεδομένου ότι, σύμφωνα με τους ερευνητές, μόλις το 0.25-0.29% της απόκλισης οφειλόταν στον πιο γνωστό επιβαρυντικό παράγοντα (APOE), γίνεται σαφές ότι υπάρχουν πολλοί άλλοι γενετικοί παράγοντες που επηρεάζουν τους δείκτες tau του CSF. Η σημασία του APOE ήταν μεγαλύτερη στο δείκτη Αβ42, κάτι που είναι γνωστό από τη βιβλιογραφία (θεωρείται ότι το APOE εκφράζεται μέσω ενός μηχανισμού που περιλαμβάνει την Αβ πρωτεΐνη). Πέρα από το APOE, βρέθηκαν και άλλες περιοχές του γονιδιώματος που επηρέαζαν τους tau δείκτες στο CSF.

Table 2. Genome-wide Significant SNPs for CSF Tau and Ptau

CHR	SNP	MAF	Closest Gene	SNP Type/Location	Tau	Ptau	Aβ ₄₂	SNP Type
19	rs769449	0.190	APOE	Intron	1.95×10^{-16}	2.56×10^{-16}	9.02×10^{-47}	Imputed
19	rs12972970	0.193	PVRL2	Intron	8.24×10^{-16}	1.28×10^{-16}	3.74×10^{-40}	Imputed
19	rs34342646	0.193	PVRL2	Intron	5.54×10^{-16}	1.17×10^{-16}	4.17×10^{-40}	Imputed
19	rs34404554	0.208	TOMM40	Intron	3.58×10^{-15}	1.33×10^{-16}	1.01×10^{-39}	Imputed
19	rs11556505	0.210	TOMM40	Synonymous	3.48×10^{-15}	1.68×10^{-16}	1.06×10^{-39}	Imputed
19	rs2075650	0.213	TOMM40	Intron	4.28×10^{-15}	5.91×10^{-16}	2.21×10^{-39}	Typed
19	rs71352238	0.201	TOMM40	Intergenic	1.78×10^{-15}	2.46×10^{-16}	5.68×10^{-39}	Imputed
19	rs12972156	0.187	PVRL2	Intron	3.03×10^{-15}	2.31×10^{-15}	6.35×10^{-37}	Imputed
3	rs9877502	0.386	SNAR-I	Intergenic	4.98×10^{-09}	1.68×10^{-07}	0.022	Imputed
3	rs1316356	0.352	SNAR-I	Intergenic	2.80×10^{-07}	1.96×10^{-07}	0.102	Typed
9	rs514716	0.136	GLIS3	Intron	1.07×10^{-08}	3.22×10^{-09}	0.026	Imputed
9	rs622536	0.114	GLIS3	Intron	6.68×10^{-07}	4.50×10^{-08}	0.029	Imputed
9	rs59860681	0.114	GLIS3	Intron	6.68×10^{-07}	4.50×10^{-08}	4.66×10^{-3}	Imputed
9	rs623295	0.114	GLIS3	Intron	6.68×10^{-07}	4.50×10^{-08}	4.66×10^{-3}	Imputed
9	rs624290	0.114	GLIS3	Intron	6.68×10^{-07}	4.50×10^{-08}	4.66×10^{-3}	Typed
6	rs6922617	0.084	NCR2	Intergenic	2.55×10^{-05}	3.58×10^{-08}	3.69×10^{-3}	Typed
6	rs11966476	0.057	NCR2	Intergenic	1.59×10^{-05}	4.97×10^{-08}	3.99×10^{-3}	Imputed

Πίνακας 7: Τα SNPs με τη μεγαλύτερη επιρροή για CSF,t-tau και p-tau βιοδείκτες [39]

Πολλά από αυτά τα γονίδια είναι γνωστό ότι επηρεάζουν τη δομή του εγκεφάλου, συνεπώς δεν είναι απίθανο να συνδέονται με τη ΝΑ. Δεδομένου ότι όλες οι ομάδες του γενότυπου παρήγαγαν κάποιο σήμα εντός λογικών ορίων, οι ερευνητές θεώρησαν ότι τα παραπάνω αποτελέσματα δεν οφείλονται σε στατιστικό λάθος. Η ομάδα, λοιπόν, συνέδεσε αρκετά γονίδια με δείκτες για τη ΝΑ. Μέχρι αυτή την εργασία οι GWAS στο CSF είχαν εξάγει στατιστική σημασία μόνο για το APOE. Οι επιστήμονες αποδίδουν το αποτέλεσμα στο πολύ μεγαλύτερο δείγμα που χρησιμοποίησαν (οι προηγούμενες έρευνες είχαν αξιοποιήσει μόνο το δείγμα της ADNI, ενώ εδώ αξιοποιήθηκε δείγμα 1209 ατόμων, όπου μόλις 394 ήταν από την ADNI – τριπλασιασμός του δείγματος). Εδώ να σημειωθεί ότι τα αποτελέσματα από διαφορετικές βάσεις δεδομένων δεν είχαν σημαντική διακύμανση, δηλαδή οι διαφορετικές κλινικές

διαδικασίες σε κάθε βάση δεν επηρέασαν το αποτέλεσμα. Συνολικά η συγκεκριμένη έρευνα έχει ενδιαφέρον γιατί δείχνει τι είναι δυνατόν να επιτευχθεί με ένα σχετικά μεγάλο δείγμα και ορισμένους απλούς δείκτες.

Προφανώς ένα αντίστοιχο δείγμα είναι αρκετά δύσκολο να επιτευχθεί στο πλαίσιο των απεικονιστικών μεθόδων, αλλά γίνονται προσπάθειες προς αυτή την κατεύθυνση. Η διακύμανση των δεικτών (7% για t-tau και 15% για p-tau) είναι ικανό μέγεθος, ιδιαίτερα γιατί το APOE έχει μικρή σχετικά επιρροή, δηλαδή παρότι δεν βρέθηκε κάποιο SNP με πολύ μεγάλη επιρροή, το άθροισμά τους έχει μεγαλύτερη ή τουλάχιστον αντίστοιχη επιρροή στο CSF από τους περισσότερους μορφομετρικούς δείκτες. Συμπεραίνει κανείς λοιπόν ότι οι CSF δείκτες προσφέρονται ως δείκτες για GWAS. Ωστόσο απαιτείται περαιτέρω σύνδεσή τους με τη NA, δεδομένου ότι δεν είναι τόσο αποτελεσματικοί βιοδείκτες όσο τα μορφομετρικά χαρακτηριστικά (σύγκριση των επιδόσεων της ταξινόμησης για διαφορετικούς βιοδείκτες γίνεται στο κεφάλαιο που ακολουθεί). Επίσης δεδομένου ότι αντίστοιχες έρευνες στο παρελθόν είχαν αποτύχει, η επανάληψη των αποτελεσμάτων είναι απαραίτητη.

3.5 Πολυπαραγοντικές μελέτες συσχέτισης γονιδιώματος

3.5.1 Γιατί απαιτούνται πολυπαραγοντικές μελέτες συσχέτισης γονιδιώματος: Δομή του γενότυπου

Οι μέχρι τώρα προσεγγίσεις των GWAS αντιμετώπιζαν κάθε SNP ως ανεξάρτητο παράγοντα. Αυτή η τακτική έχει όμως κάποιους περιορισμούς: πρώτα από όλα λίγα SNPs μπορεί να αποκτήσουν στατιστική σημασία, ιδιαίτερα εξαιτίας του μικρού δείγματος και του τεράστιου αριθμού των SNPs και των voxels. Παράλληλα, ξέρουμε ότι υπάρχει υψηλή γενετική αιτιότητα στη NA, αλλά σίγουρα πλέον δεν υπάρχουν αρκετά σημαντικοί ανεξάρτητοι γενετικοί παράγοντες που να την εξηγούν. Επίσης, έχει πλέον εδραιωθεί από πλήθος μελετών η άποψη ότι υπάρχει δομή στα γενετικά δεδομένα, δηλαδή ότι ομάδες SNPs ή και γονιδίων αλληλεπιδρούν στην έκφρασή τους στον εγκέφαλο (και αυτή η επίδραση επιπλέον περιορίζεται κυρίως σε γειτονικές περιοχές του γονιδιώματος. Μία προφανής δομή για τα γενετικά δεδομένα για παράδειγμα είναι τα loci, δηλαδή οι περιοχές ενός χρωμοσώματος που υπάρχει η ακολουθία ενός γονιδίου ή ενός τμήματος του DNA) [40]. Αυτό συνεπάγεται ότι η στατιστική σημασία κάποιου SNP δεν μπορεί να αποτιμηθεί ανεξάρτητα από τα υπόλοιπα, καθώς θα υπάρχουν πολλά SNPs που θα εμφανίζονται να έχουν κάποια στατιστική σημασία που όμως προκύπτει από την εξάρτησή τους από άλλα SNPs (για παράδειγμα η έκφραση κάποιων SNPs εξαρτάται από το APOE-e4). Παράλληλα, καθώς τα πολυδιάστατα δεδομένα έχουν κάποια μη τυχαία συσχέτιση, τα διανύσματα των παραγόντων είναι εξαρτημένα (μη ορθογώνια), άρα η διακύμανση των παραγόντων της παλινδρόμησης είναι μεγάλη

και δεν εξάγονται στατιστικά εύρωστα αποτελέσματα. Τέλος, η πολυπαραγοντική προσέγγιση είναι λιγότερο επίπονη υπολογιστικά από τη μαζική μονοπαραγοντική προσέγγιση, γιατί εξετάζονται οι πολύ λιγότερες συσχετίσεις των γενετικών και απεικονιστικών ομάδων, αντί για κάθε γενετική και απεικονιστική μονάδα. Όλοι αυτοί οι παράγοντες, οδήγησαν στην εφαρμογή πολυπαραγοντικών μεθόδων στα GWAS στη ΝΑ, που προσπαθούν να ομαδοποιήσουν το γενότυπο και να εντοπίσουν τη στατιστική σημασία κάθε σχηματισμού.

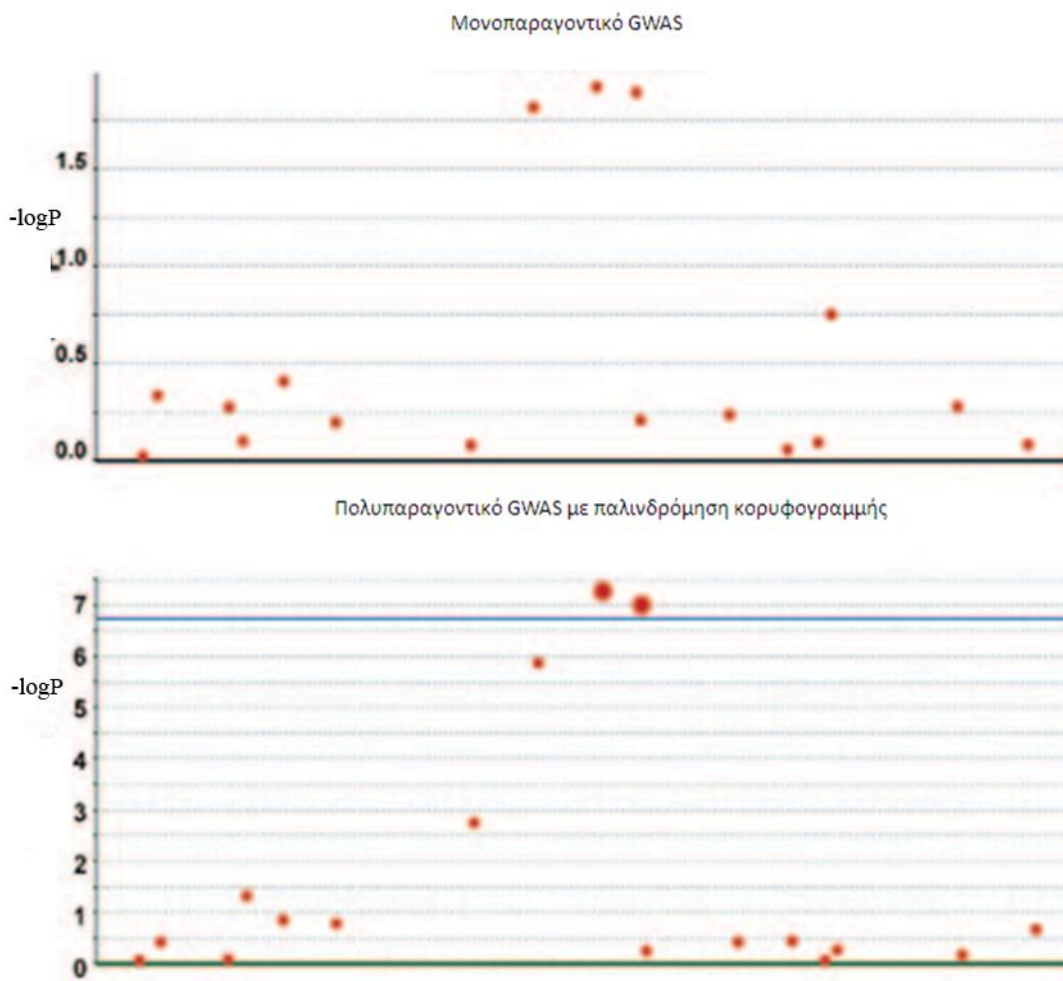
3.5.2 Παλινδρόμηση κορυφογραμμής για την ομαδοποίηση πολυμορφισμών: Αύξηση στατιστικής ισχύος

Όπως έχουμε δει τα GWAS παράγουν νέα ερευνητικά στοιχεία, αλλά απαιτούν και τη διαχείριση τεράστιων βάσεων δεδομένων. Συνεπώς οι ερευνητές αναζητούν συνεχώς την εφαρμογή νέων στατιστικών μεθόδων. Σε αυτό το πλαίσιο οι ερευνητές της παραπάνω εργασίας εφαρμόζουν παλινδρόμηση κορυφογραμμής για να διαχειριστούν αλληλοεξαρτόμενα πολυδιάστατα δεδομένα. Εδώ έγκειται και η πρωτοτυπία της συγκεκριμένης έρευνας [41]. Οι έρευνες που είδαμε μέχρι τώρα ασχολούνταν με την ανεξάρτητη επιρροή κάθε SNP στον φαινότυπο. Αντίθετα εδώ εικάζεται ότι, ενώ ένα SNP μόνο του μπορεί να μην έχει ανιχνεύσιμη στατιστική επιρροή, η αλληλεπίδραση με άλλα SNPs θα σπρώξει κάποια από αυτά στην περιοχή στατιστικής συσχέτισης. Αυτή η τακτική έχει εφαρμοστεί επιτυχώς σε άλλες ασθένειες.

Οι ερευνητές προσπάθησαν λοιπόν να ανιχνεύσουν SNPs που συνδέονται με χαρακτηριστικά του όγκου του κροταφικού λοβού και του ιππόκαμπου και που δεν είχαν εντοπιστεί με τα κλασικά μονοπαραγοντικά GWAS. Τα δεδομένα παράχθηκαν με την εφαρμογή της TBM μεθόδου στο δείγμα MRI της ADNI.

Οι ερευνητές προσπάθησαν να εντοπίσουν SNPs που ήταν κοντά στο όριο στατιστικής συσχέτισης με την κλασική μονοπαραγοντική παλινδρόμηση που είχαν εφαρμόσει παλιότερα, ελπίζοντας ότι η στατιστική συσχέτιση θα δυνάμωνε με την εφαρμογή της πολυπαραγοντικής παλινδρόμησης κορυφογραμμής και θα αναδείκνυε τα πλεονεκτήματα των πολυπαραγοντικών μεθόδων. Αυτό συνέβη για ορισμένες περιπτώσεις, αλλά όχι για όλες (για παράδειγμα ενώ ένα SNP έχει ικανή επιρροή σε ένα επιβαρυντικό γονίδιο στην απλή παλινδρόμηση, άλλο SNP φαίνεται να έχει μεγαλύτερη επιρροή όταν λαμβάνουμε υπόψη τα υπόλοιπα SNPs). Βέβαια, υποστηρίζεται ότι με τα πολυπαραγοντικά συστήματα θα υπάρχουν και λιγότερες λανθασμένες κατατάξεις των SNPs (false positives). Συνολικά τα αποτελέσματα διαφοροποιούνται αρκετά και, χωρίς περισσότερες πολυπαραγοντικές και μη έρευνες, είναι δύσκολο να διασταυρωθούν. Ωστόσο σε άλλες αντίστοιχες περιπτώσεις, όπως η έκφραση του

γονιδιώματος σε άλλες ασθένειες, προέκυψαν πιο έγκυροι ερευνητικοί στόχοι με τις μεθόδους παραγοντοποίησης, ενώ και εδώ, το άθροισμα της στατιστικής σημασίας που αποκτούν τα SNPs μεγαλώνει. Όπως θα δούμε και σε μετέπειτα έρευνες [42], η αθροιστική επιρροή των SNPs που εντοπίστηκαν στον όγκο του κροταφικού λοβού με πολυπαραγοντικές μεθόδους έφθασε να ξεπερνά το 20%, από το 15% στις μονοπαραγοντικές.



Διάγραμμα της σχέσης του όγκου του ιππόκαμπου με το rs2912975 SNP.
 Στο μονοπαραγοντικό GWAS η p -τιμή του στο σύνολο του γονιδίου είναι $1.19 \cdot 10^{-2}$
 Λαμβάνοντας υπόψη και τα κοντινά του SNPs, μέσω της παλινδρόμησης κορυφογραμμής, το ίδιο SNP έχει p -τιμή $4.98 \cdot 10^{-8}$.

Διάγραμμα 3: Μεγάλη διαφοροποίηση επιρροής ενός SNP-υποψήφιου στόχου με εφαρμογή παλινδρόμησης κορυφογραμμής στο GWAS [41]

Πάντως οι ερευνητές εντόπισαν τρία νέα SNPs (που στη σχετική βιβλιογραφία σχετίζονται με ογκομετρικά χαρακτηριστικά) και τα δύο από αυτά επιβεβαίωσαν την επιρροή τους σε ακόλουθους στατιστικούς ελέγχους. Συνολικά η έρευνα όντως παράγει διαφορετικά συμπεράσματα από τις

μονοπαραγοντικές μεθόδους. Τα αποτελέσματα θα πρέπει να στηριχθούν από άλλες αντίστοιχες έρευνες.

3.5.3 Προσομοίωση με αραιή μέθοδο μειωμένου βαθμού για μελέτη συσχέτισης γονιδιώματος στη νόσο του Αλτσχάιμερ

Στην επόμενη μελέτη οι ερευνητές προσπάθησαν και αυτοί να αντιμετωπίσουν την αλληλεξάρτηση των SNPs [23]. Για να το πετύχουν αυτό εφήρμοσαν μία παραλλαγή του reduced-rank regression, που εξετάστηκε σε προηγούμενο κεφάλαιο. Οι ερευνητές επέλεξαν σε πρώτη φάση να μην χρησιμοποιήσουν πραγματικά δεδομένα, αλλά ένα λογισμικό που προσομοιώνει τον ανθρώπινο γενότυπο, ενώ δημιούργησαν και διάφορους φαινότυπους του εγκεφάλου με πρότυπο τις εικόνες MRI της ADNI.

Εδώ οι ερευνητές προτείνουν τη επιβολή ενός ορίου στο βαθμό του πίνακα συντελεστών της παλινδρόμησης, μειώνοντας έτσι τον αριθμό των συντελεστών που πρέπει να υπολογιστούν και επιτρέποντας την αξιοποίηση της πολυπαραγοντικής φύσης των δεδομένων.

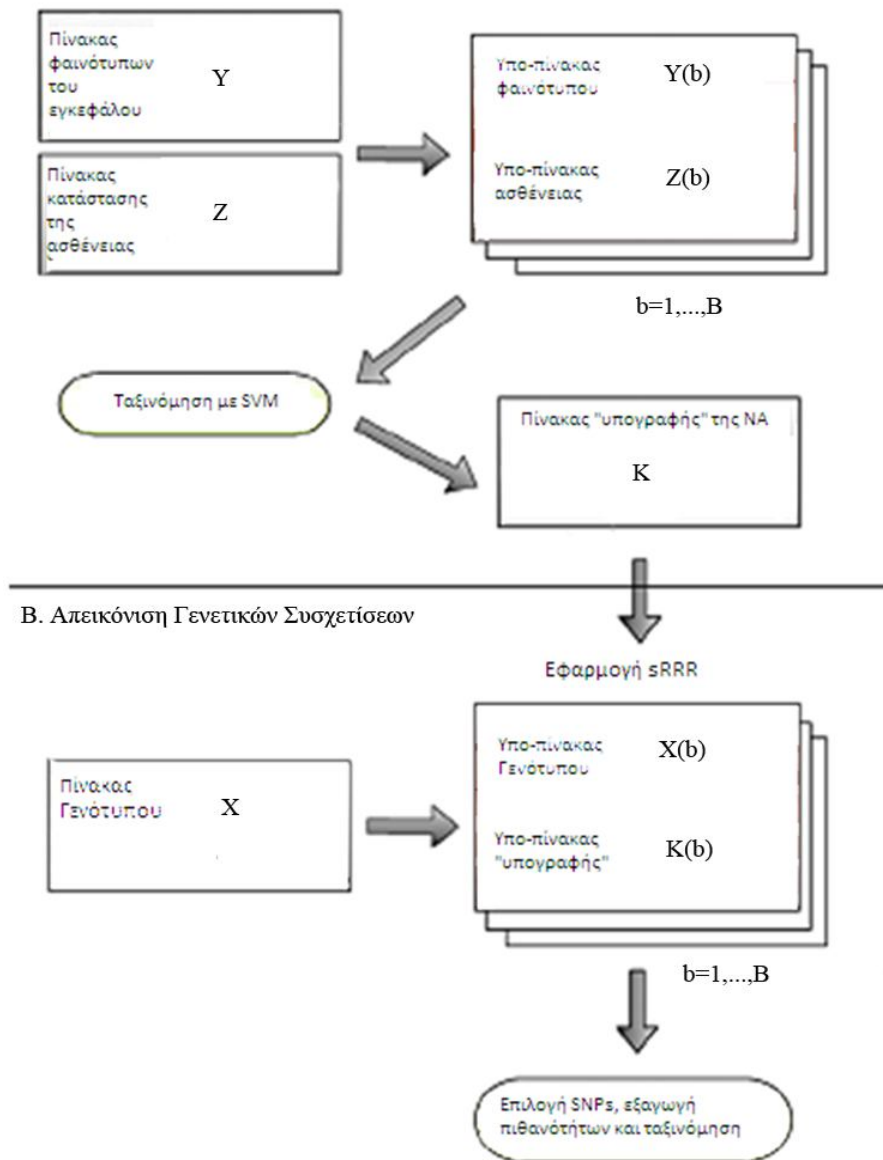
Η ομάδα εισήγαγε κάποια γενετική συσχέτιση στους φαινότυπους και δοκίμασε τις επιδόσεις του συστήματος σε σύγκριση με την απλή πολυπαραγοντική προσέγγιση. Οι ερευνητές κατέληξαν ότι η υπολογιστική ισχύς του μοντέλου τους είναι ανώτερη από την απλή πολυπαραγοντική προσέγγιση, για αυτό και επέκτεινε στη συνέχεια την εφαρμογή του σε πραγματικά δεδομένα.

3.5.4 Εφαρμογή της αραιής μεθόδου μειωμένου βαθμού σε πραγματικά δεδομένα: Πλήρης πολυπαραγοντική εφαρμογή

Αφού επιβεβαίωσαν ότι η παραπάνω μέθοδος προσέφερε καλά αποτελέσματα στην προσομοίωση, η ομάδα αποφάσισε να δοκιμάσει τον αλγόριθμο σε πραγματικά δεδομένα [42]. Αποφάσισαν όμως να μην χρησιμοποιήσουν το σύνολο του εγκεφάλου για τους φαινότυπους, αλλά να περιοριστούν στις αλλαγές στα πλέον σημαντικά voxels για την πορεία της ασθένειας. Αυτό έγινε εφικτό με την επεξεργασία των MRI με τον αλγόριθμο του επιβαρυσμένου LDA, που πρακτικά ενθάρρυνε τα voxels μιας περιοχής ενδιαφέροντος να παραμείνουν μαζί. Αυτή η επιλογή είναι ίσως μειονέκτημα για την έρευνα, δεδομένου ότι η ΝΑ έχει παθολογία που εκτείνεται στο σύνολο του εγκεφάλου, ωστόσο τέτοιοι περιορισμοί έχουν τεθεί από σχεδόν όλες τις έρευνες GWAS γιατί έτσι ενισχύεται η υπολογιστική ισχύς (SNPs που δεν θα ήταν στατιστικά σημαντικά στο σύνολο του εγκεφάλου, εδώ αποκτούν σημασία). Είναι ενδιαφέρον πάντως ότι η LDA μέθοδος επέλεξε περιοχές που είναι γνωστές για τη σχέση τους με τη ΝΑ (ιπόκαμπος, πλευρικές κοιλίες) και άρα αυτοματοποιεί μία διαδικασία που γινόταν ούτως ή άλλως.

Μετά την εφαρμογή της LDA μεθόδου, οι ερευνητές εφήρμοσαν τον αλγόριθμο του sRRR. Η εφαρμογή της επιβαρυσμένης LDA μεθόδου συμπληρώνεται απόλυτα με τον sRRR αλγόριθμο, αφού στόχος είναι και στις δύο περιπτώσεις η εισαγωγή αραιότητας (sparsity) για να αποκοπούν οι περιττές μεταβλητές. Σε πιο θεωρητικό επίπεδο, αυτό μπορεί να στηριχθεί από το γεγονός ότι οι αυθεντικές γενετικές συσχετίσεις δημιουργούν όμοιες δομές σε γειτονικά pixels (άλλωστε από εκεί προέκυψε και η ανάγκη για πολυπαραγοντική παλινδρόμηση) και οτιδήποτε άλλο μπορεί να θεωρηθεί θόρυβος και άρα να αφαιρεθεί. Για την κατάταξη της σημασίας των SNPs δεν χρησιμοποιήθηκε κάποιο κλασικό μοντέλο cross-validation (όπως leave one out), αλλά δημιουργήθηκαν τυχαίες υποομάδες του πληθυσμού στις οποίες εφαρμόστηκε ο sRRR. Τα SNPs κατατάχθηκαν με βάση τη συχνότητα με την οποία επιλέχθηκαν στις διάφορες υποομάδες.

A. Πολυπαραγοντική Επιλογή Φαινότυπων



Εικόνα 12: Επιλογή φαινότυπων με LDA και επιβολή sRRR στα δεδομένα για εξαγωγή γενετικών συσχετίσεων. Ο πίνακας υπογραφής εκφράζει τον πίνακα απεικονιστικών βιοδεικτών που επιλέχθηκαν [42]

Το δείγμα που αξιοποιήθηκε ήταν της ADNI. Μεγάλη σημασία έχει το γεγονός ότι οι ερευνητές δεν πραγματοποίησαν απλώς τη συνηθισμένη σύγκριση (ασθενείς με NA έναντι υγιών) για να εξάγουν γενετικές συσχετίσεις. Αντίθετα, εφήρμοσαν την ίδια διαδικασία και για τη σύγκριση εξελισσόμενου MCI έναντι υγιών και εξελισσόμενου MCI έναντι στάσιμου MCI. Είναι λίγες οι GWAS που χρησιμοποίησαν αυτές τις συγκρίσεις, οι οποίες όμως έχουν τεράστιο ενδιαφέρον γιατί έτσι μπορεί να παρακολουθηθεί η γενετική συσχέτιση στα διαφορετικά στάδια της ασθένειας. Παράλληλα, και

ιδιαίτερα με την τελευταία σύγκριση, αποτιμάται η αξία των γενετικών δεικτών για τη διάγνωση (με την πρόγνωση αυτών που θα εμφανίσουν ΝΑ από MCI).

Τα συμπεράσματα που εξήγαγαν οι επιστήμονες ήταν σημαντικά. SNPs με υπολογίσιμη επιρροή, δηλαδή μεγαλύτερη του 0.5%, εμφανίστηκαν και στα τρία διαφορετικά πειράματα. Στο κλασικό πείραμα της εύρεσης γενετικών διαφορών μεταξύ ασθενών με ΝΑ και υγιών το άθροισμα της στατιστικής επιρροής των σημαντικών SNPs ξεπερνούσε το 20%, όπως φαίνεται στον παρακάτω πίνακα, δηλαδή υπάρχει σημαντική βελτίωση σε σχέση με τις κλασικότερες μεθόδους που συνήθως εντόπιζαν επιρροή μικρότερη του 15%. Ακολουθεί ο σχετικός πίνακας.

AD vs CN						
	SNP	Gene	Chr	MAF	HWE	\hat{p}_{xy}
Rank 1	APOE-ε4	APOE	19	0.276	0.083	0.996
	rs2075650	TOMM40	19	0.252	0.868	0.944
	rs3815501	BZW1	2	0.154	0.470	0.758
	rs11132507		4	0.284	0.643	0.705
	rs11132508		4	0.284	0.643	0.705
	rs1681052	LOC647946	18	0.077	0.647	0.681
	rs7761213		6	0.303	0.458	0.675
	rs17345545		1	0.266	0.422	0.647
	rs13340334	PDZD2	5	0.112	0.336	0.611
	rs17103124		14	0.152	0.623	0.605
Rank 2	rs9263844		6	0.161	1.000	0.772
	rs9263846		6	0.161	1.000	0.772
	rs7999394	MTRF1	13	0.408	0.027	0.746
	rs3794328	MTRF1	13	0.408	0.027	0.746
	rs11590365		1	0.114	0.547	0.621
	rs11204949		1	0.114	0.547	0.621
	rs11204971		1	0.114	0.547	0.621
	rs12405278	FLG	1	0.114	0.547	0.621
	rs215340		12	0.248	1.000	0.593
	rs7603289		2	0.345	0.330	0.585
Rank 3	rs727432	ADCY2	5	0.282	0.877	0.709
	rs11783329		8	0.398	0.435	0.589
	rs7114756	MAML2	11	0.161	0.059	0.581
	rs17309585		8	0.406	1.000	0.526
	rs10491327		5	0.183	0.143	0.520
	rs12534148	PDE1C	7	0.270	0.153	0.506

Πίνακας 8 : Τα σημαντικά SNPs στο πείραμα ασθενών με NA και υγιών οργανωμένα σύμφωνα με το γονίδιο. Στη δεξιά στήλη η εκτιμώμενη στατιστική σημασία [42]

Αντίστοιχα μεγάλο ενδιαφέρον έχουν και τα τέσσερα σημαντικά SNPs στη σύγκριση εξελισσόμενου MCI με NA. Αν και το άθροισμα της επιρροής τους είναι μικρότερο του 2%, το συγκεκριμένο πείραμα δεν έχει βρει αποτελέσματα σε άλλη εργασία που θα μελετήσουμε.

P-MCI vs CN						
SNP	Gene	Chr	MAF	HWE	\hat{p}_{xj}	
Rank 1	APOE-ε4	APOE	19	0s.271	0.083	0.982
	rs2883782	MYO3B	2	0.483	0.387	0.746
	rs2798062		9	0.256	0.105	0.712
	rs10934170		3	0.146	0.619	0.681
	rs17826780		4	0.102	0.318	0.665
	rs7843577		8	0.448	1.000	0.617
	rs2075650	TOMM40	19	0.246	0.180	0.589
	rs1405443		7	0.135	1.000	0.583
	rs758491	RBFOX1	16	0.352	1.000	0.566
	rs914166		21	0.150	0.624	0.548
Rank 2	rs13132552	SORBS2	4	0.348	1.000	0.792
	rs12633719		3	0.233	0.490	0.671
	rs11069874		13	0.158	0.242	0.661
	rs885339		13	0.158	0.242	0.661
	rs2381958		5	0.204	0.343	0.655
	rs10041184		5	0.217	0.146	0.639
	rs4265409		1	0.440	0.706	0.627
	rs7584948	ANTXR1	2	0.187	0.105	0.623
	rs501435	ODZ4	11	0.150	0.811	0.599
	rs1001684		5	0.265	0.874	0.595
Rank 3	rs705837	PRSS12	4	0.375	0.895	0.639
	rs11856999	MAP2K5	15	0.148	0.227	0.605
	rs7653663	MGLL	3	0.165	0.111	0.601
	rs12597064		16	0.240	1.000	0.570
	rs633398	NDST3	4	0.373	0.895	0.546
	rs631271	NDST3	4	0.371	1.000	0.544

Πίνακας 9: Τα σημαντικά SNPs στο πείραμα ασθενών με NA και ασθενών με εξελισσόμενο MCI. Στη δεξιά στήλη η εκτιμώμενη στατιστική επιρροή [42]

Και στις τρεις συγκρίσεις η APOE-ε4 ήταν με διαφορά ο σημαντικότερος παράγοντας (και στις τρεις συγκρίσεις η πιθανότητα να επιλεγθεί ήταν μεγαλύτερη του 0.9, στις δύο πρώτες συγκρίσεις ήταν κοντά στο 1). Αυτό επιβεβαιώνει την καλή λειτουργία του συστήματος, αφού ο ρόλος της συγκεκριμένης αλλέλης είναι διασταυρωμένος από πολλές εργασίες. Έχει ενδιαφέρον ότι η σημασία της παρέμεινε υψηλή και στις τρεις συγκρίσεις, δηλαδή η επιρροή που ασκεί είναι μεγάλη σε όλη τη διάρκεια της ασθένειας, και άρα έχει ρόλο στην ανάπτυξη, άλλα και στη μετάβαση από MCI σε NA.

Μεγάλη σημασία φάνηκε να έχει και ο δείκτης rs2075650 του TOMM40 γονιδίου, που επίσης είναι γνωστός για τη συμβολή του στη ΝΑ. Ωστόσο, δεν πρόβαλε ως σημαντικός παράγοντας στο διαχωρισμό των σταδίων MCI. Άλλα δύο γονίδια, τα COX7A2L και NRXN1 είχαν μεγάλη επίδοση στη σύγκριση εξελισσόμενου MCI με υγιείς, και καθώς και τα δύο αυτά γονίδια έχουν θεωρητικά συνδεθεί με τη ΝΑ, είναι υποψήφιοι στόχοι. Άλλα πέντε γονίδια είχαν ισχυρή επιρροή και δεν έχουν συνδεθεί με τη ΝΑ στη βιβλιογραφία. Η επιρροή τους θα διερευνηθεί σε περαιτέρω μελέτες.

Συνολικά, η συγκεκριμένη εργασία είναι πολύ αξιόλογη. Χρησιμοποίησε έναν πολλά υποσχόμενο αλγόριθμο, με τον οποίο η ομάδα είχε πειραματιστεί εκτεταμένα σε προηγούμενη έρευνα, προκειμένου να ενισχύσει τη στατιστική ισχύ του σήματος. Εφήρμοσε μια νέα μέθοδο κατάταξης και παρήγαγε νέους ερευνητικούς στόχους. Η συνολική στατιστική που εντόπισε ξεπερνά το 20%, μέγεθος μεγαλύτερο κατά τουλάχιστον 5% από τις περισσότερες μελέτες. Ο περιορισμός των voxels με το επιβαρυνόμενο LDA είναι επίσης σημαντικός, δεδομένου ότι προκύπτει από αυτοματοποιημένη διαδικασία και δεν ορίζει έναν αυθαίρετο στόχο. Σε πολλές έρευνες GWAS, πολλά SNPs είναι κοντά στο όριο στατιστικής σημασίας, χωρίς να το υπερβαίνουν όμως, καθώς το δείγμα είναι μικρό. Τέλος, οι ερευνητές αναφέρουν ότι συνέχεια στην έρευνα για τη ΝΑ θα πρέπει να δοθεί στο πλαίσιο των πολυτροπικών ερευνών. Είναι όντως πιθανό οι ανεξάρτητες έρευνες των GWAS να αρχίσουν να περιορίζονται, καθώς τα γενετικά στοιχεία εξετάζονται σε συνάρτηση με άλλους παράγοντες, όπως θα δούμε στη συνέχεια.

3.5.5 Πολυπαραγοντική ταξινόμηση του γονιδιώματος με πολυδιακριτή ανάλυση: Ανάπτυξη γενετικού δείκτη χωρίς απεικονιστικά δεδομένα

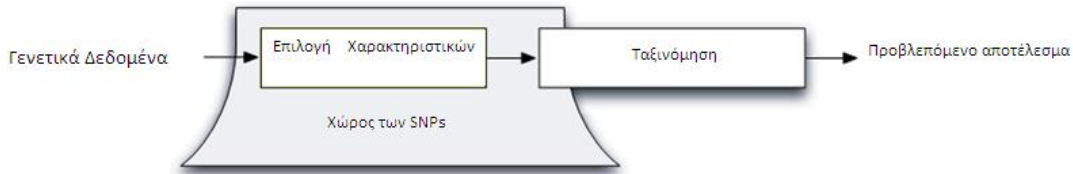
Έχουμε μέχρι τώρα δει ότι η τάση στα GWAS είναι η χρήση πολυπαραγοντικών μεθόδων για να αξιοποιούνται οι αλληλεπιδράσεις των SNPs και να αυξάνεται η υπολογιστική ισχύς. Οι περισσότεροι αλγόριθμοι βασίζονταν στην επιβολή κάποιου πίνακα-ορίου στη μαζική παλινδρόμηση. Μία αρκετά πρωτότυπη έρευνα εισάγει την έννοια της πολυδιακριτής ανάλυσης (multiresolution analysis, MRA) [24].

Οι συγγραφείς αναγνώρισαν την ανάγκη μιας πολυπαραγοντικής προσέγγισης στην εξαγωγή δομών του ανθρώπινου γονιδιώματος. Σε αυτό το πλαίσιο, θεώρησαν το γονιδίωμα ως ένα μονοδιάστατο σήμα, από το οποίο ήθελαν να εξάγουν ακολουθίες SNPs. Με την εφαρμογή της MRA επεδίωξαν να διασπάσουν το γονιδίωμα σε ένα σύνολο χώρων, όπου ίσως βρεθούν καλύτερα χαρακτηριστικά. Αν και η MRA έχει εφαρμοστεί και στο παρελθόν στην ανάλυση του ανθρώπινου DNA, αυτή είναι η πρώτη έρευνα για την πρόβλεψη κλινικού αποτελέσματος στη ΝΑ μέσω MRA. Οι ερευνητές δεν πρόβαλαν τη γενετική πληροφορία σε κάποιο απεικονιστικό βιοδείκτη της ΝΑ, αλλά

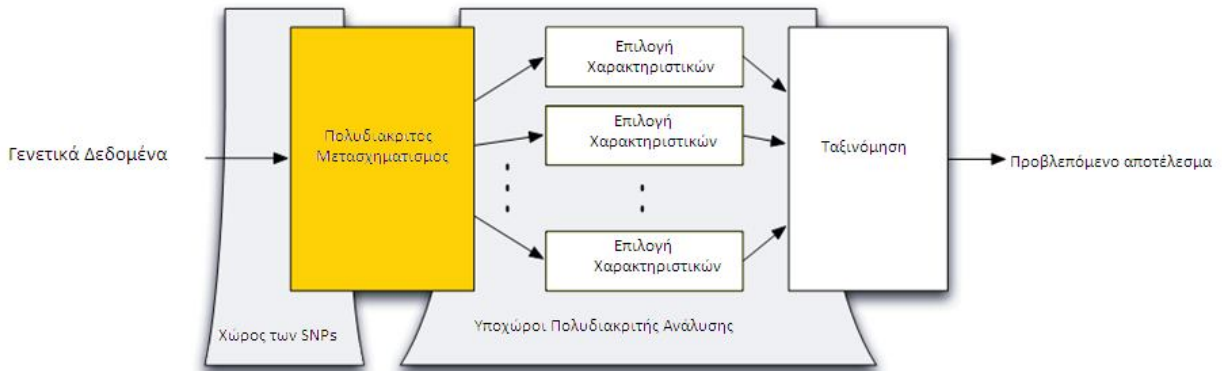
προσπάθησαν να φτιάξουν έναν γενετικό δείκτη που θα λύνει το ακόλουθο δυαδικό πρόβλημα: έχει ο ασθενής NA; Προφανώς, αν το συγκεκριμένο ερώτημα μπορούσε να απαντηθεί άμεσα με γενετικά δεδομένα, δεν θα χρησιμοποιούνταν απεικονιστικές μέθοδοι. Η αλήθεια είναι ότι καθαρά γενετικοί δείκτες έχουν παραχθεί με επιτυχία σε άλλες ασθένειες αλλά η περίπτωση της NA είναι πιο περίπλοκη. Συνεπώς η συγκεκριμένη εργασία έχει ενδιαφέρον κυρίως στο πλαίσιο της πρωτότυπης ομαδοποίησης που εφαρμόζει και στη δυνατότητα εφαρμογής της και σε άλλες πηγές δεδομένων. Γενικά η αναδιάταξη των γενετικών δεδομένων (deep gene resequencing) απαιτείται στο μέλλον της έρευνας της NA και ο συγκεκριμένος αλγόριθμος μπορεί να βρει εφαρμογή εκεί.

Η βασική ιδέα της μεθόδου είναι η ακόλουθη: Στο πλαίσιο της εκπαίδευσης το γονιδίωμα κάθε ανθρώπου αποδομείται σε έναν N -διάστατο χώρο, όπου N ο αριθμός των βασικών SNPs (που περιλαμβάνουν το σύνολο των αλλέλων). Κάθε τμήμα του γονιδιώματος πρέπει να ταξινομηθεί στην κλάση που συμμετέχει στην εμφάνιση NA ή που δεν συμμετέχει στη NA. Από κάθε υποχώρο εξάγονται χαρακτηριστικά και εφαρμόζεται ταξινόμηση. Εδώ εφαρμόζεται στατιστική επεξεργασία για την εύρεση των T -καλύτερων χαρακτηριστικών, τα οποία χρησιμοποιούνται για την ταξινόμηση (με λογιστική παλινδρόμηση και τυχαία δάση στην προκειμένη περίπτωση). Το σύνολο των ταξινομητών ενοποιείται με έναν απλό κανόνα.

Όταν φτάσουμε στην πρόβλεψη, κάθε δείγμα διασπάται με MRA, τα T -χαρακτηριστικά που βρέθηκαν παραπάνω χρησιμοποιούνται για τη δημιουργία ενός πίνακα χαρακτηριστικών, που είναι η είσοδος του ταξινομητή του υποχώρου. Η πρόβλεψη αποτελεί το συνδυασμό των προβλέψεων των ταξινομητών.



A. Η κλασική διαδικασία για την επεξεργασία γενετικών δεδομένων



B. Η επεξεργασία με τη πολυδιακριτή ανάλυση. Προσθέτει ένα επίπεδο προεπεξεργασίας

Εικόνα 13: Διαφορά διαδικασίας GWAS και πολυδιακριτής ανάλυσης σε συνδυασμό με GWAS [24]

Οι ερευνητές χρησιμοποίησαν τη βάση δεδομένων του Translational Genomics Research Institute με δείγμα 1411 ατόμων (861 με LOAD, 550 υγιείς). Η ομάδα επικεντρώθηκε στην ανάλυση της περιοχής κοντά στο APOE.

Η ομάδα εφήρμοσε το πείραμα για διαφορετικά επίπεδα αποδόμησης (3 και 5) και με τους δύο ταξινομητές (λογιστική παλινδρόμηση και τυχαία δάση) και συνέκρινε τα αποτελέσματα με την απευθείας εφαρμογή λογιστικής παλινδρόμησης ή τυχαίων δασών (χωρίς MRA). Σε όλες τις περιπτώσεις η εφαρμογή MRA βελτίωνε την επίδοση του συνολικού ταξινομητή (0.75 έναντι 0.70), ενώ η εφαρμογή λογιστικής παλινδρόμησης έδωσε καλύτερα αποτελέσματα από το τυχαίο δάσος. Τα αποτελέσματα που βρήκε η ομάδα ήταν τα αναμενόμενα: τα SNPs που συνθέτουν την APOE-ε4 ήταν στην κορυφή της λίστας, όπως και SNPs από γονίδια όπως το ApoC1, που συνδέονται με τη NA. Μερικά επιπλέον SNPs που είχαν στατιστική σημασία, δεν είχαν συνδεθεί με τη NA μέχρι τώρα.

Συνολικά, η ομάδα επιβεβαίωσε τα πλεονεκτήματα μιας πολυπαραγοντικής προσέγγισης στην ανάλυση του γονιδιώματος. Πρότεινε μια προσέγγιση με MRA, που έχει δώσει εύρωστα αποτελέσματα σε άλλες περιπτώσεις. Δεδομένου ότι η αποδόμηση με MRA τείνει να ομαδοποιεί τα γειτονικά SNPs, κάτι που στην περίπτωση της NA είναι θετικό (όμοια έκφραση των γειτονικών SNPs), η μέθοδος έχει

περιθώρια εφαρμογής σε ερευνητικό στάδιο. Ο γενετικός δείκτης που δημιούργησε για τη NA με τη συμμετοχή της MRA, βελτίωσε αισθητά τα αποτελέσματα. Τέλος, θα είχε ενδιαφέρον η εφαρμογή της μεθόδου και στη συσχέτιση με απεικονιστικά δεδομένα.

Ωστόσο, υπολογιστικά η μέθοδος θα είναι πολύ επίπονη υπολογιστικά για την εφαρμογή σε έρευνες που δουλεύουν όντως σε όλο το γονιδίωμα. Εδώ οι μελετητές χρησιμοποίησαν σαν βάσεις ένα περιορισμένο αριθμό από SNPs, ενώ συνολικά αξιοποιήθηκαν περίπου 300.000 SNPs, ένα σχετικά μικρό μέγεθος για τα δεδομένα των GWAS στη NA.

3.5.6 Εφαρμογή παράλληλης ανάλυσης σε ανεξάρτητες συνιστώσες σε πολυπαραγοντική μελέτη συσχέτισης γονιδιώματος: Αντιστοίχιση με βιολογικά δίκτυα του εγκεφάλου

Μία από τις πιο ενδιαφέρουσες GWAS είναι η ακόλουθη. Η ομάδα που εισήγαγε τη χρήση του paraICA σε GWAS [37] μέσω εκτεταμένων προσομοιώσεων, ολοκλήρωσε και ένα εκτεταμένο πολυπαραγοντικό GWAS με πραγματικά δεδομένα (βάση της ADNI) [19].

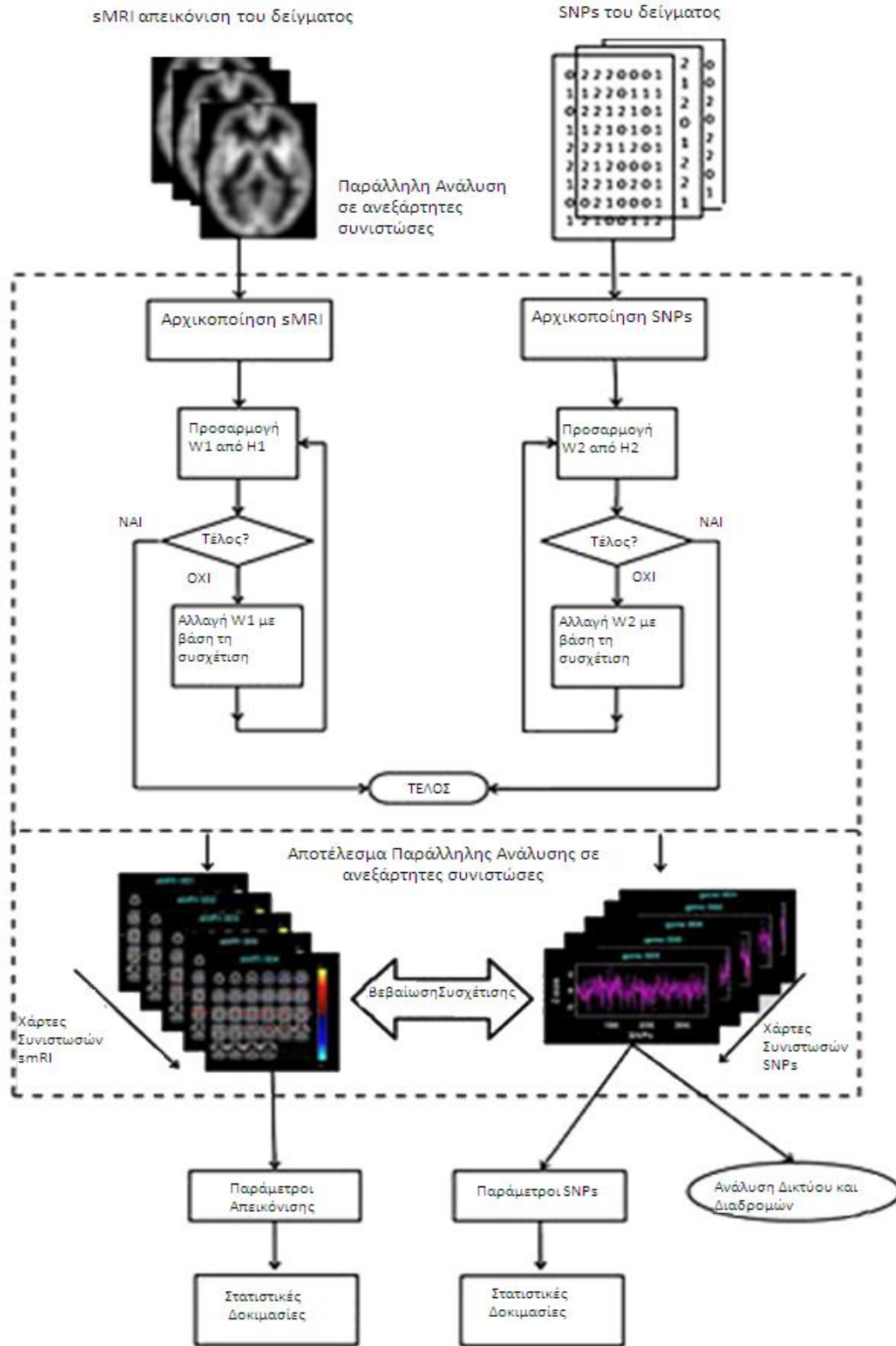
Οι μέχρι τώρα πολυπαραγοντικές μελέτες που είδαμε αξιοποιούσαν κατεξοχήν τη γειννίαση για την ομαδοποίηση των SNPs. Αυτή η μελέτη, όμως, εισήγαγε και τη γνωστή βιολογική λειτουργία των γονιδίων για να τα ομαδοποιήσει, κάτι που διαφωτίζει την αιτιολογία και την εξέλιξη της NA, καθώς και τη σχέση της με άλλες ασθένειες.

Οι ερευνητές εφήρμοσαν τον αλγόριθμο του paraICA για την ταυτόχρονη πολυπαραγοντική επεξεργασία των δύο διαφορετικών πηγών δεδομένων, των sMRI και του γενότυπου. Όμοια δηλαδή με τον αλγόριθμο της ομάδας της Vounou [42], που είδαμε νωρίτερα, εφαρμόζεται παραγοντοποίηση και στις δύο ομάδες δεδομένων. Όμως, ο paraICA έχει ένα εγγενές πλεονέκτημα: τα δύο διαφορετικά σετ δεδομένων επεξεργάζονται ταυτόχρονα. Έτσι, κατά την ομαδοποίηση δεν επιδιώκεται απλώς η μέγιστη ανεξαρτησία κάθε πηγής στο πλαίσιο κάθε ενός από τα δύο σετ δεδομένων (που είναι αντίστοιχη επεξεργασία με τη μέθοδο επιβαρυσμένου LDA της Vounou), αλλά και να μεγιστοποιηθεί η συσχέτιση των πηγών ανάμεσα στα δύο σετ δεδομένων. Θεωρητικά τουλάχιστον ο paraICA αλγόριθμος αναμένουμε να δώσει καλύτερα αποτελέσματα.

Η ερευνητική ομάδα ουσιαστικά ολοκλήρωσε δύο στάδια GWAS: μετά την αρχική επεξεργασία του γενότυπου και των MRI, εφήρμοσε μαζική μονοπαραγοντική παλινδρόμηση στα δεδομένα του γενότυπου, χωρίς κάποια διόρθωση για αλληλεπίδραση των SNPs. Επέβαλαν ένα σχετικά χαμηλό στατιστικό όριο και συγκράτησαν τα SNPs που είχαν επιρροή στην ασθένεια (ουσιαστικά διαφορετική συσχέτιση των SNPs στις τρεις πειραματικές ομάδες, τους υγιείς, τους ασθενείς με MCI και τους

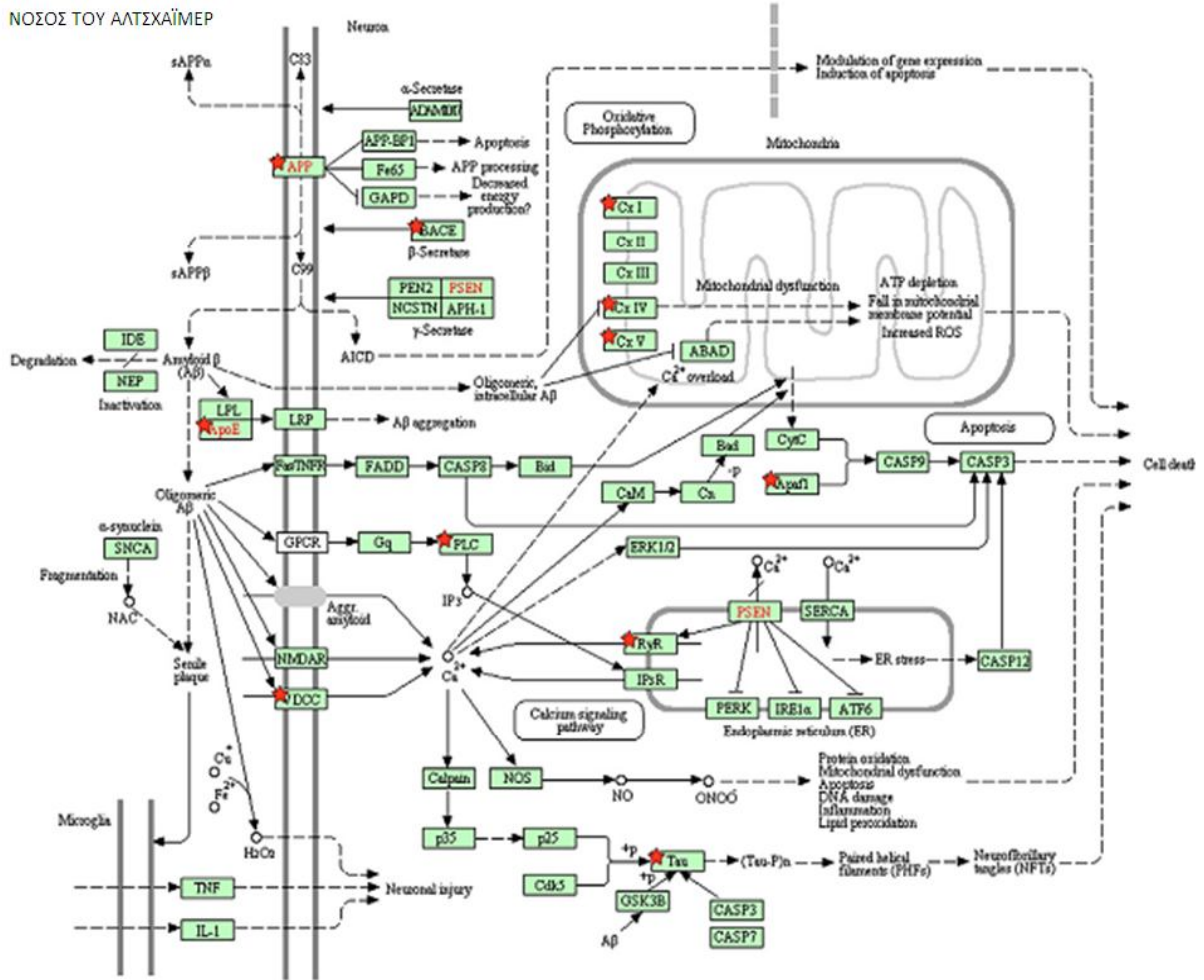
ασθενείς με ΝΑ). Αυτό το πρώτο στάδιο εφαρμόστηκε ώστε να απομονωθούν τα SNPs που έχουν κάποια επιρροή στην ασθένεια και να γίνει πιο διαχειρίσιμο το μέγεθος των δεδομένων. Είναι σίγουρα μια πιθανή αδυναμία της έρευνας, αλλά, δεδομένης της εκτεταμένης επεξεργασίας και χαρτογράφησης που ακολούθησε, κρίθηκε απαραίτητο. Τα αποτελέσματα που εξήγαγαν οι ερευνητές σε αυτό το σημείο ήταν σε γενικές γραμμές τα αναμενόμενα, με το APOE-ε4 να έχει τη μεγαλύτερη στατιστική επιρροή.

Στο δεύτερο στάδιο επέβαλαν τον paraICA αλγόριθμο στα δεδομένα που σε μεγάλο βαθμό θα σχετίζονταν με τη ΝΑ. Οι ομάδες των γονιδίων οργανώθηκαν από τον ICA αλγόριθμο σε σχέση με τη λειτουργία τους, με βάση τη συμμετοχή τους είτε σε κάποιο κοινό λειτουργικό δίκτυο του εγκεφάλου, είτε σε γειτονικές δομές του εγκεφάλου, ή γενικά σε κάποια κοινή βιολογική διαδικασία. Η ομαδοποίηση της MRI αφορούσε τα μορφομετρικά χαρακτηριστικά (όγκος, πάχος) τοπικών τμημάτων του εγκεφάλου, που οργανώθηκαν από τον ICA αλγόριθμο σε δίκτυα. Σε κάθε επανάληψη του paraICA αλγορίθμου αποτιμήθηκε η αλληλεπίδραση της έκφρασης (μορφομετρικής και λειτουργικής) κάθε ομάδας του γενότυπου με τις ομάδες των MRI. Συνολικά σχηματίστηκαν μόλις 12 και 8 ομάδες αντίστοιχα, δηλαδή κάθε ομάδα ήταν αρκετά εκτεταμένη.



Εικόνα 14: Σχηματική αναπαράσταση της εφαρμογής του paraICA στην επεξεργασία γενετικών και απεικονιστικών δεδομένων σε GWAS [19]

Τέσσερις ομάδες του γενότυπου βρέθηκαν κατεξοχήν να αλληλεπιδρούν με την ίδια ομάδα του εγκεφάλου με βάση τα δίκτυα που συμμετείχαν οι ομάδες και τις αντίστοιχες βιολογικές λειτουργίες.



Εικόνα 15: Κύρια δίκτυα του εγκεφάλου που επηρεάζονται από τη ΝΑ, η δυσλειτουργία τους και τα γονίδια που εκφράζονται μέσω αυτών [19]

Τα γονίδια με ανοιχτό χρώμα είναι όσα έχουν ήδη επισημανθεί. Όσα έχουν κόκκινο αστεράκι είναι τα επιβαρυντικά γονίδια, σύμφωνα με την έρευνα.

Όπως φαίνεται και από το παραπάνω σχήμα, πολλά γονίδια φέρεται να επηρεάζουν εγκεφαλικά δίκτυα και να συμμετέχουν στο LOAD. Πολλά από αυτά τα γονίδια ή τα SNPs δεν είχαν αποκτήσει

στατιστική σημασία σε απλά GWAS, αλλά η προσέγγισή τους στο πλαίσιο μεγάλων ομάδων, τους αποδίδει σημαντικό ρόλο.

Η έρευνα κινήθηκε σε διαφορετική κατεύθυνση από τις άλλες πολυπαραγοντικές μεθόδους, αφενός γιατί εφήρμοσε έναν αλγόριθμο που λαμβάνει υπόψη του την αλληλεξάρτηση κατά την ομαδοποίηση και όχι μετά, αφετέρου γιατί δεν αρκέστηκε στην ομαδοποίηση των SNPs σε ομάδες (loci) προκειμένου να αυξηθεί η στατιστική ισχύς. Εδώ η ομάδα αποφάσισε να αξιοποιήσει ουσιαστικά την πολυπαραγοντικότητα και σχημάτισε ομάδες με μεγάλο αριθμό γονιδίων (για παράδειγμα μία ομάδα είχε 169 σημαντικά γονίδια και 332 σημαντικά SNPs) και κριτήριο τη βιολογική λειτουργία. Ίσως έτσι δεν προβάλλεται ο ρόλος του εκάστοτε γονιδίου ή SNP, αλλά μπορεί να χαρτογραφηθεί ακριβέστερα η σχέση της NA με ασθένειες ή βιολογικές λειτουργίες, καθώς είναι γνωστό ότι πολλά γονίδια συμμετέχουν σε μία διαδικασία. Ένα ακόμη μειονέκτημα της έρευνας είναι η εφαρμογή της μαζικής μονοπαραγοντικής παλινδρόμησης σε πρώτο στάδιο. Έτσι αφαιρέθηκαν όσα SNPs δεν είχαν άμεσα σχέση με την ασθένεια, όμως ούτως ή άλλως ο στόχος ήταν να εντοπιστούν νέα γονίδια που είχαν εμμέσως σχέση με τη NA μέσω της συμμετοχής τους σε κάποια διαδικασία (βέβαια το γεγονός ότι αφαιρέθηκαν μόνο όσα δεν είχαν σημαντική διακύμανση στις τρεις πειραματικές ομάδες μειώνει αυτόν τον κίνδυνο). Μία ακόμη πιθανή αδυναμία της έρευνας είναι ο μικρός αριθμός των πηγών. Όπως είδαμε στο κεφάλαιο που αξιολογείται ο paraICA, η μεγάλη υποτίμηση ή υπερτίμηση του αριθμού των πηγών μπορεί να αλλοιώσει τα αποτελέσματα. Οι προηγούμενες έρευνες που εξετάσαμε κατέληξαν σε έναν πολύ μεγαλύτερο αριθμό πηγών, αν και βέβαια χρησιμοποίησαν άλλα κριτήρια για το σχηματισμό τους.

Γενικά η συγκεκριμένη έρευνα είναι δύσκολο να αξιολογηθεί δεδομένου ότι ο στόχος ήταν κυρίως να ξεδιαλύνει τη γενετική αιτιολογία της νόσου και τα διαφορετικά δίκτυα του εγκεφάλου που συμμετέχουν. Παρόλαυτα ακολουθεί ο πίνακας των σημαντικών SNPs, σύμφωνα με τις ομάδες που οργανώθηκαν, που είναι αρκετά εκτεταμένος. Παρατηρούμε ότι η πρώτη ομάδα έχει τη σημαντικότερη συσχέτιση.

Gene	Location	Function	rs #	AD (# alleles)		MCI (# alleles)		HC (# alleles)		Chi-Square	P-value	Weighting score
				A	B	A	B	A	B			
G1												
APOE3	19q13.2	N/A	E3	162	200	269	465	39	279	88	1.00E-16	1.000
APOE4	19q13.2	N/A	E4	209	153	521	273	358	60	80	1.00E-15	0.909
SHROOM2	Xp22.3	Intron	rs2405940	209	153	503	231	294	124	16.58	0.0003**	0.188
ZNF673	Xp11.3	Intron	rs7876304	103	259	250	484	172	246	14.04	0.0009**	0.160
SLC9A7	Xp11.23	Intron	rs1883255	99	263	241	493	165	253	13	0.0015**	0.148
SLC16A2	Xq13.2	Intron	rs479640	89	273	252	482	138	280	11.15	0.0038	0.127
MAG3	1p13.2	Intron	rs7552954	154	208	373	361	226	192	10.97	0.0041	0.125
COBL	7p12.2	Intron	rs1437490	195	167	374	360	255	163	10.9	0.0043	0.124
PLXDC2	10p12.31	Intron	rs2358839	102	260	265	469	150	268	7.54	0.0229	0.086
CTNND2	5p15.2	Intron	rs757459	215	147	388	346	210	208	6.93	0.0311	0.079
G2												
ZNF673	Xp11.3	Intron	rs7876304	103	259	250	484	172	246	14.04	0.0009**	1.000
SLC9A7	Xp11.3	Intron	rs1883255	99	263	241	493	165	253	13	0.0015**	0.926
TNXB	6p21.3	Intron	rs2239689	133	229	213	521	114	304	9.47	0.0088	0.675
ZSCAN12	6p22.1	Mis-sense	rs1361385	260	102	497	237	258	160	9.24	0.0098	0.658
EHMT2	6p21.3	Intron	rs2844457	146	216	236	498	131	287	8.9	0.0116	0.634
ZKSCAN3	6p22.1	Intron	rs213228	258	104	491	243	258	160	8.03	0.018	0.572
SLC44A4	6p21.3	Intron	rs2736428	139	223	227	507	127	291	7.41	0.0246	0.528
CFB	6p21.3	Intron	rs4151657	223	139	509	225	286	132	6.96	0.0307	0.496
C6orf15	6p21.3	Intron	rs3130977	259	103	504	230	264	154	6.7	0.0349	0.477
CCHCR1	6p21.3	Intron	rs746647	256	106	490	244	261	157	6	0.0496	0.427
G3												
ATP5G2	12q13.13	Intron	rs1800634	111	251	166	568	134	284	14.97	0.0006**	1.000
VPS13C	15q22.2	Intron	rs1981916	190	172	449	285	216	202	12.84	0.0016**	0.858
ATF7	12q13.13	Intron	rs784568	189	173	337	397	230	188	9.82	0.0073	0.656
CEP57	11q21	Intron	rs3017756	154	208	316	418	215	203	8.95	0.0113	0.598
EHMT2	6p21.3	Intron	rs2844457	146	216	236	498	131	287	8.9	0.0116	0.595
MTMR2	11q21	Intron	rs193364	225	137	437	297	219	199	8.66	0.0132	0.578
PTCHD2	1p36.22	Intron	rs2379135	160	202	259	475	165	253	8.32	0.0156	0.556
CNTN5	11q22.1	Intron	rs10892901	134	228	314	420	146	272	7.86	0.0196	0.525
SLC44A4	6p21.32	Intron	rs2736428	139	223	227	507	127	291	7.41	0.0246	0.495
C2	6p21.3	Intron	rs2734335	204	158	355	379	211	207	6.22	0.0445	0.415
G4												
SKIV2L	6p21.3	Intron	rs592229	227	135	388	346	239	179	9.69	0.0078	0.999
TNXB	6p21.3	Intron	rs2239689	133	229	213	521	114	304	9.47	0.0088	0.976
EHMT2	6p21.3	Intron	rs159445	268	94	486	248	277	141	7.71	0.0211	0.795
AKAP9	7q21.2	Intron	rs756647	234	128	461	273	234	184	7.39	0.0248	0.762
SLC44A4	6p21.3	Intron	rs2242665	227	135	397	337	238	180	7.34	0.0254	0.757
CFB	6p21.3	Intron	rs4151657	223	139	509	225	286	132	6.96	0.0307	0.718
ATF6B	6p12.3	Intron	rs2228628	242	120	526	208	315	103	6.9	0.0316	0.711
C6orf10	6p21.3	Intron	rs485774	203	159	455	279	272	146	6.81	0.0331	0.702
CYP51A1	7q21.2	Intron	rs2301559	233	129	454	280	234	184	6.34	0.0419	0.654
C2	6p21.3	Intron	rs2734335	204	158	355	379	211	207	6.22	0.0445	0.641

** Genes that significantly discriminate groups after correction for multiple comparisons (Bonferroni-corrected across 40 chi-square tests).

Πίνακας 10: πίνακας των σημαντικών SNPs σύμφωνα με την έρευνα [19]. Τα αποτελέσματα είναι οργανωμένα στις γενετικές ομάδες που σχηματίστηκαν και με κατάταξη αντίστοιχη της σημασίας τους

Σίγουρα πάντως η συγκεκριμένη έρευνα διαφωτίζει τη γενετική αιτιολογία της ΝΑ και μεταβάλλει την προοπτική μας για τους γενετικούς δείκτες στη ΝΑ. Πέρα από τους συνηθισμένους ελέγχους, όπως για παράδειγμα για την APOE-ε4, είναι πιθανό να προκύψει κάποιος ακριβέστερος δείκτης που περιλαμβάνει ολόκληρη ομάδα αλληλεπιδρώντων SNPs. Τέλος, οι ερευνητές εντόπισαν, εκτός από τα γνωστά επιβαρυντικά γονίδια, και πολλά νέα που θα πρέπει να επιβεβαιωθούν σε νέες μελέτες και θα μας δώσουν κάποια ένδειξη για την αξία της έρευνας.

3.6 Συνολικά για τις μελέτες συσχέτισης γονιδιώματος στη νόσο του Αλτσχάιμερ

Κλείνοντας το κεφάλαιο των GWAS στη ΝΑ, βλέπουμε ότι επικεντρωθήκαμε στη συσχέτιση απεικονιστικών δεικτών (κατεξοχήν sMRI) και γενετικών παραγόντων. Αυτό οφείλεται στην παραδοχή ότι αφενός η γενετική επιρροή εκφράζεται κυρίως σε φυσιολογικές λειτουργίες (σε σύγκριση για παράδειγμα με το MMSE) και αφετέρου ότι οι μορφομετρικές αλλαγές του εγκεφάλου είναι οι πιο πρώιμοι και ευαίσθητοι δείκτες για τη ΝΑ.

Γενικά τα GWAS εκπλήρωσαν την υπόσχεση εύρεσης γενετικών στόχων για τη ΝΑ. Δεδομένου ότι η επιρροή μεμονωμένων γενετικών παραγόντων στη ΝΑ είναι πολύ μικρή, δεν υπήρχε άλλη ερευνητική δυνατότητα. Σε μελέτες υποψήφιων γονιδίων στη συνέχεια ή και σε άλλα GWAS, πολλοί από αυτούς τους παράγοντες επιβεβαιώθηκαν.

Καθώς υπάρχει κληρονομικότητα στο LOAD, αλλά δεν υφίσταται κάποιος αποφασιστικός γενετικός παράγοντας, πιστεύεται ότι ομάδες από πολλά SNPs αλληλεπιδρούν και συμμετέχουν στην ανάπτυξή του. Κάνοντας αυτή την παραδοχή, οι τελευταίες έρευνες για το GWAS αξιοποιούν μεθόδους που ομαδοποιούν το γενότυπο. Πέρα από αυτή τη θεωρητική παραδοχή, ούτως ή άλλως μία ομάδα έχει μεγαλύτερη στατιστική επιρροή από ένα μεμονωμένο SNP, άρα έτσι περιοχές του γενότυπου που δεν έχουν στατιστική σημασία, μπορεί να αποκτήσουν (νέοι γενετικοί στόχοι).

Οι πολυπαραγοντικές μέθοδοι που είδαμε αξιοποίησαν διάφορους αλγόριθμους (όρια στη μαζική παλινδρόμηση, paraICA, κλπ). Οι περισσότερες αρκέστηκαν στο να ομαδοποιήσουν κάποια loci με βάση γενετικούς χάρτες και κριτήριο τη γειννιάσή τους (με τη δεύτερη παραδοχή ότι μας ενδιαφέρουν μορφομετρικά χαρακτηριστικά που σε γειτονικές περιοχές του εγκεφάλου συνήθως καθορίζονται από γειτονικά γονίδια). Αυτό από μόνο του έδωσε νέα αποτελέσματα, ίσως στατιστικά πιο εύρωστα. Η χρήση του paraICA, όπως την είδαμε στη συγκεκριμένη έρευνα, απομακρύνθηκε ακόμη περισσότερο από την “τοπικιστική” προσέγγιση των SNPs και χρησιμοποίησε ακόμη πιο εκτεταμένη ομαδοποίηση των γονιδίων με κριτήριο τη συμμετοχή τους σε ολόκληρα δίκτυα του εγκεφάλου. Ίσως έτσι προκύψει κάποιος αποτελεσματικός γενετικός δείκτης για τη ΝΑ, που θα αφορά ένα σύνολο επιβαρυντικών γονιδίων.

Συνολικά πάντως το ενδιαφέρον έχει στραφεί σε άλλες προσεγγίσεις, όπως η αλληλεπίδραση των γενετικών παραγόντων με το περιβάλλον και ο συνδυασμός όχι μόνο γενετικών και απεικονιστικών δεικτών, αλλά όλων των βιοδεικτών σε έναν παγκόσμιο δείκτη για τη ΝΑ.

3.7 Γενετικοί και περιβαλλοντικοί παράγοντες

Αν και οι γενετικοί παράγοντες είναι επιβαρυντικοί για τη ΝΑ, η τυχαιότητα της εμφάνισής του και η ποικιλία της παθολογίας του, έχει πείσει πολλούς ερευνητές ότι περιβαλλοντικές αλλαγές συμμετέχουν στην εμφάνιση της νόσου. Ειδικά για τη ΝΑ, έρευνες έχουν δείξει ότι το κάπνισμα, ο διαβήτης, οι διατροφικές ελλείψεις και η υπέρταση αυξάνουν την πιθανότητα να εμφανίσει κάποιος ΝΑ. Η έκθεση του ανθρώπου σε ορισμένα βακτήρια (πχ σπειροχάιτη) επίσης θεωρείται ότι μπορεί να ενεργοποιήσουν τους μηχανισμούς ανάπτυξης της ΝΑ [43]. Αντίστοιχα, παράγοντες όπως το υψηλό επίπεδο μόρφωσης, η συγκατοίκηση με άλλους ανθρώπους και η φυσική δραστηριότητα φαίνεται να μειώνουν την πιθανότητα να εμφανιστεί η νόσος.

Καθώς μέρος της παθολογίας της ΝΑ έχει ήδη εξηγηθεί με γενετικούς όρους, το περιβάλλον σε καμία περίπτωση δεν ευθύνεται αποκλειστικά για την εμφάνιση της νόσου. Η αλληλεπίδρασή του με επιβαρυντικά για τη ΝΑ γονίδια όμως είναι ένα ανοιχτό ερευνητικό ερώτημα. Το μοντέλο για την επίλυση της ΝΑ δηλαδή είναι

Περιβάλλον x Γενότυπος = ΝΑ [44]

Το μοντέλο αυτό δεν είναι γραμμικό, δηλαδή δεν υποστηρίζεται ότι υπάρχει κάποιος περιβαλλοντικός παράγοντας που οδηγεί σε ορισμένο φαινότυπο, αλλά πολλοί παράγοντες που δρουν συνεργατικά. Η αλληλεπίδραση των επιβαρυντικών παραγόντων χρεώνεται με τη μη ικανοποιητική ερμηνεία της ΝΑ μέχρι τώρα, καθώς οι έρευνες είχαν πιο γραμμικό χαρακτήρα και, όπως θα δούμε παρακάτω, πρέπει να ληφθεί υπόψη στο σχεδιασμό των σχετικών σύγχρονων πολυτροπικών συστημάτων.

Το παραπάνω μοντέλο δεν προκύπτει μόνο από την περιπλοκότητα της ΝΑ. Η επιρροή του περιβάλλοντος στην έκφραση του γενότυπου έχει μελετηθεί εκτεταμένα και έχουν διαπιστωθεί αλλαγές έπειτα από έκθεση σε βακτήρια, τοξίνες ή σε στρεσογόνους ψυχολογικούς παράγοντες. Ειδικά για τη ΝΑ, οι φορείς της αλλέλης ΑΡΟΕ-ε4, του κατεξοχήν επιβαρυντικού παράγοντα για τη νόσο, εμφανίζουν σε υψηλότερα ποσοστά τη ΝΑ, εφόσον έχουν και υψηλά επίπεδα χοληστερίνης. Τα παραπάνω είναι σημαντικά, καθώς μπορούν να οδηγήσουν στην πρόληψη. Παράλληλα όμως αποτελούν και εφιαλτήριο για την ερμηνεία του μηχανισμού της ΝΑ. Για παράδειγμα οι φορείς της ΑΡΟΕ-ε4 αλλέλης που έχουν ιστορικό τραύματος στο κρανίο εμφανίζουν 10 φορές μεγαλύτερη πιθανότητα να εμφανίσουν ΝΑ, από τους φορείς χωρίς αντίστοιχο ιστορικό [45], δηλαδή ένα ατύχημα ίσως ενεργοποιεί την αμυλοειδική συσσώρευση στον εγκέφαλο, εφόσον υπάρχει κάποια αντίστοιχη

προδιάθεση. Για το συγκεκριμένο θέμα πρότερου τραυματισμού η ADNI μαζεύει αυτή τη στιγμή στατιστικά δεδομένα.

Η αποτίμηση των περιβαλλοντικών παραγόντων στη ΝΑ ανοίγει ένα μεγάλο ερευνητικό πεδίο. Ο τρόπος με τον οποίο αυτοί επιδρούν στο ανθρώπινο γονιδίωμα είναι επίσης μεγάλη πρόκληση. Οι προσπάθειες σε αυτή την κατεύθυνση επικεντρώνονται κυρίως στο πεδίο της επιγενετικής.

3.7.1 Επιγενετική

Όπως έχει προαναφερθεί, η ΝΑ παρουσιάζει ποικίλη παθολογία και είναι εξαιρετικά δύσκολο να προβλεφθεί. Μελέτες σε μονοζυγωτικούς διδύμους δείχνουν ότι ενώ μοιράζονται το ίδιο γενετικό υλικό, δεν εμφανίζουν τη ΝΑ με τον ίδιο τρόπο ή και καθόλου [46]. Το πρόβλημα αυτό για αρκετούς ερευνητές μπορεί να εξηγηθεί μέσω της επιγενετικής.

Η επιγενετική αφορά τις (εν δυνάμει αναστρέψιμες) διαφοροποιήσεις στην έκφραση των γονιδίων ή στους κυτταρικούς φαινότυπους που δεν συνδέονται με αλλαγές στην ακολουθία του DNA. Ενώ το γενετικό υλικό είναι καθορισμένο, η λειτουργική έκφραση του γονιδίου μεταβάλλεται από τους επιγενετικούς μηχανισμούς, που λειτουργούν σαν “κουμπί έντασης” για την έκφραση των γονιδίων [47]. Αυτοί οι μηχανισμοί έχουν χρησιμοποιηθεί για να εξηγήσουν, έστω και μερικώς, πολλές χρόνιες νευρολογικές ασθένειες. Τα επιγενετικά χαρακτηριστικά είναι αναστρέψιμα, αλλά μπορούν κάλλιστα να είναι και κληρονομικά (μέσω της μίτωσης και της μειωτικής), κάτι που ταιριάζει με την υψηλή κληρονομικότητα της ΝΑ.

Η επιγενετική είναι συνεπώς ένα ενδιαφέρον μοντέλο για την ερμηνεία της πολυποίκιλης παθολογίας της ΝΑ. Το γεγονός ότι για το LOAD δεν έχει βρεθεί κάποιος καθοριστικός γενετικός παράγοντας (τα γονίδια που μελετήσαμε είναι κατεξοχήν επιβαρυντικά, δεν ερμηνεύουν την ασθένεια) ενισχύει αυτή τη θέση. Επίσης, έρευνες σε μονοζυγωτικούς διδύμους έχουν δείξει ότι οι επιγενετικοί μηχανισμοί είναι όμοιοι στους διδύμους μικρής ηλικίας, ενώ μεταβάλλονται αισθητά στους γηραιότερους διδύμους [48]. Αυτό υπαινίσσεται κάποια σχέση ανάμεσα στην επιγενετική και στη γήρανση, που ως γνωστόν είναι ο κατεξοχήν επιβαρυντικός παράγοντας για τη ΝΑ. Υπάρχουν άλλωστε ενδείξεις ότι η παθολογία της ΝΑ οφείλεται στην αλληλεπίδραση γονιδίων και περιβάλλοντος, όπως αυτή εκφράζεται μέσω επιγενετικών μηχανισμών.

3.7.2 Επιγενετικοί μηχανισμοί

Οι δύο καλύτερα τεκμηριωμένοι τέτοιοι μηχανισμοί είναι η μεθυλίωση του DNA και οι τροποποιήσεις των ιστονών. Οι αλλαγές στις πρωτεΐνες των ιστονών ή η αλλαγή στον τρόπο που το DNA τυλίγεται γύρω από τις ιστόνες, μπορεί να μεταβάλλει σημαντικά την πρόσβαση του μεταγραφικού μηχανισμού σε κάποια γονίδια, ενώ αφήνει άλλα ανεπηρέαστα [44]. Οι τροποποιήσεις στις ιστόνες εκφράζονται με πολλούς τρόπους, ο χαρακτηριστικότερος ίσως από αυτούς είναι η ακετυλίωση, που μειώνει ή εμποδίζει εντελώς τη μεταγραφή συγκεκριμένων γονιδίων, μέσω της ουδετέρωσης του θετικού φορτίου των πρωτεϊνών των ιστονών. Η μεθυλίωση του DNA είναι ένας άλλος μηχανισμός της επιγενετικής που συνδέεται στενά και με την τροποποίηση των ιστονών και αφορά την ομοιοπολική προσθήκη μεθυλομάδων στις κυτοσίνες, με στόχο επίσης να μειωθούν τα επίπεδα μεταγραφής ενός γονιδίου [49]. Οι παραπάνω μηχανισμοί είναι απαραίτητοι στον οργανισμό προκειμένου για παράδειγμα τα κύτταρα του εγκεφάλου, που δεν χρειάζεται να παράγουν αιμοσφαιρίνη, όντως να μην παράγουν. Ωστόσο οι επιγενετικοί μηχανισμοί θεωρείται ότι είναι και το μέσο με το οποίο το περιβάλλον επηρεάζει την έκφραση του ανθρώπινου γονιδιώματος, όπως σημειώνουν πολλές έρευνες που καταγράφουν αλλαγές στην επιγενετική έπειτα από έκθεση του οργανισμού σε ιούς, ακτινοβολία, ορμόνες, μεταβολή στη διατροφή, αλλά ακόμη και συναισθηματικούς και κοινωνικούς παράγοντες.

3.7.3 Γήρανση, επιγενετική και νόσος του Αλτσχάιμερ

Η ύπαρξη σχέσης ανάμεσα στην επιγενετική και τη ΝΑ είναι αρκετά καλά τεκμηριωμένη, όπως δείχνουν τα αποτελέσματα μιας έρευνας που βασιζόταν στη νεκροψία εγκεφάλων διδύμων με ΝΑ και χωρίς ΝΑ. Ο εγκεφαλος του διδύμου που είχε ΝΑ εμφάνιζε διαφορές στις πρωτεΐνες των ιστονών [51]. Άλλες μελέτες σε ασθενείς εν ζωή και πειραματόζωα επιβεβαιώνουν αυτά τα αποτελέσματα. Σε αυτό το στάδιο πάντως δεν γνωρίζουμε αν οι επιγενετικοί μηχανισμοί συμμετέχουν ενεργά στην ανάπτυξη της ΝΑ ή απλώς εκφυλίζονται εξαιτίας της ΝΑ [50]. Η υπόθεση αυτή στηρίζεται κυρίως στην καλά τεκμηριωμένη συσχόρευση των επιγενετικών αλλαγών με τη γήρανση του οργανισμού, καθώς και στον ρόλο της επιγενετικής σε νευρολογικές ασθένειες όπως η σχιζοφρένεια ή η κατάθλιψη. Για να εξετασθεί αν και πώς η επιγενετική είναι αιτία ανάπτυξης ΝΑ θα πρέπει να γίνουν μελέτες στα πρώιμα στάδια της νόσου, και ιδιαίτερα στη μετάβαση από MCI σε ΝΑ, για παράδειγμα με χάρτες τροποποίησης των ιστονών ή καταγραφή της επιγενετικής κατάστασης των γονιδίων που σχετίζονται με την εμφάνιση ΝΑ. Οι δυσκολίες που υπάρχουν στην επανάληψη των αποτελεσμάτων των GWAS, ή η χαμηλή στατιστική συσχέτιση ορισμένων SNPs ίσως να οφείλεται στους επιγενετικούς μηχανισμούς. Αν και είναι δύσκολο να συγκεντρωθούν τα σχετικά στοιχεία, προσφέρουν έναν νέο ερευνητικό στόχο για τη ΝΑ.

Table 1
Evidence on altered DNA methylation in AD and AD model systems.

DNA methylation pattern alterations	Sample	Reference
APP promoter hypomethylation in AD patient	Human brains	West et al. (1995)
APP promoter hypomethylation with aging, significant above 70 years of age	Human parietal cortex	Tohgi et al. (1999a,b)
Hypomethylation of promoters of ribosomal genes with aging	Human lymphocytes	de Carvalho et al. (2000)
APOE, MTHFR promoter hypermethylation	Human prefrontal cortex and lymphocytes	Wang et al. (2008)
Higher "epigenetic distance from norm" in AD vs. normal, TFAM in brain and NCSN in lymphocytes the biggest difference	Human prefrontal cortex and lymphocytes	Wang et al. (2008)
Decrements in DNA methylation	Human entorhinal cortex	Mastroeni et al. (2008)
Differences in DNA methylation in a twin pair discordant for AD	Human temporal neocortex	Mastroeni et al. (2009)
Increase in SORBS3 promoter methylation and decrease in S100A2 promoter methylation	Human cerebral cortex	Siegmund et al. (2007)
APP promoter methylation influenced by sex steroids and aging	Intact and gonadectomized mice brains	Mani and Thakur (2006)
Hypomethylation of APP, BACE1, Sp1 with a decrease in DNMT1 activity after infantile exposure to Pb	Non-human primate various cortical areas	Wu et al. (2008)
PSEN1 is regulated by promoter methylation in response to metabolic stimuli	TgCRND8 mice brains	Fuso et al. (2009)
SAM downregulates PSEN1 expression and A β production	Neuronal cell cultures	Scarpa et al. (2003)
BACE, PSEN1 expression regulated by promoter methylation in response to metabolic stimuli	Neuronal cell cultures	Fuso et al. (in press)
Differential expression of four genes implicated in AD in response to SAM administration	Neuronal cell cultures	Cavallaro et al. (2006)
APP, PSEN1 hypomethylation after SAH administration, prevented with betaine	Microglial cell cultures	Lin et al. (2009)
A β induces aberrant methylation, i.e. hypermethylation of NEP promoter	Cerebral endothelial cultures	Chen et al. (2009)

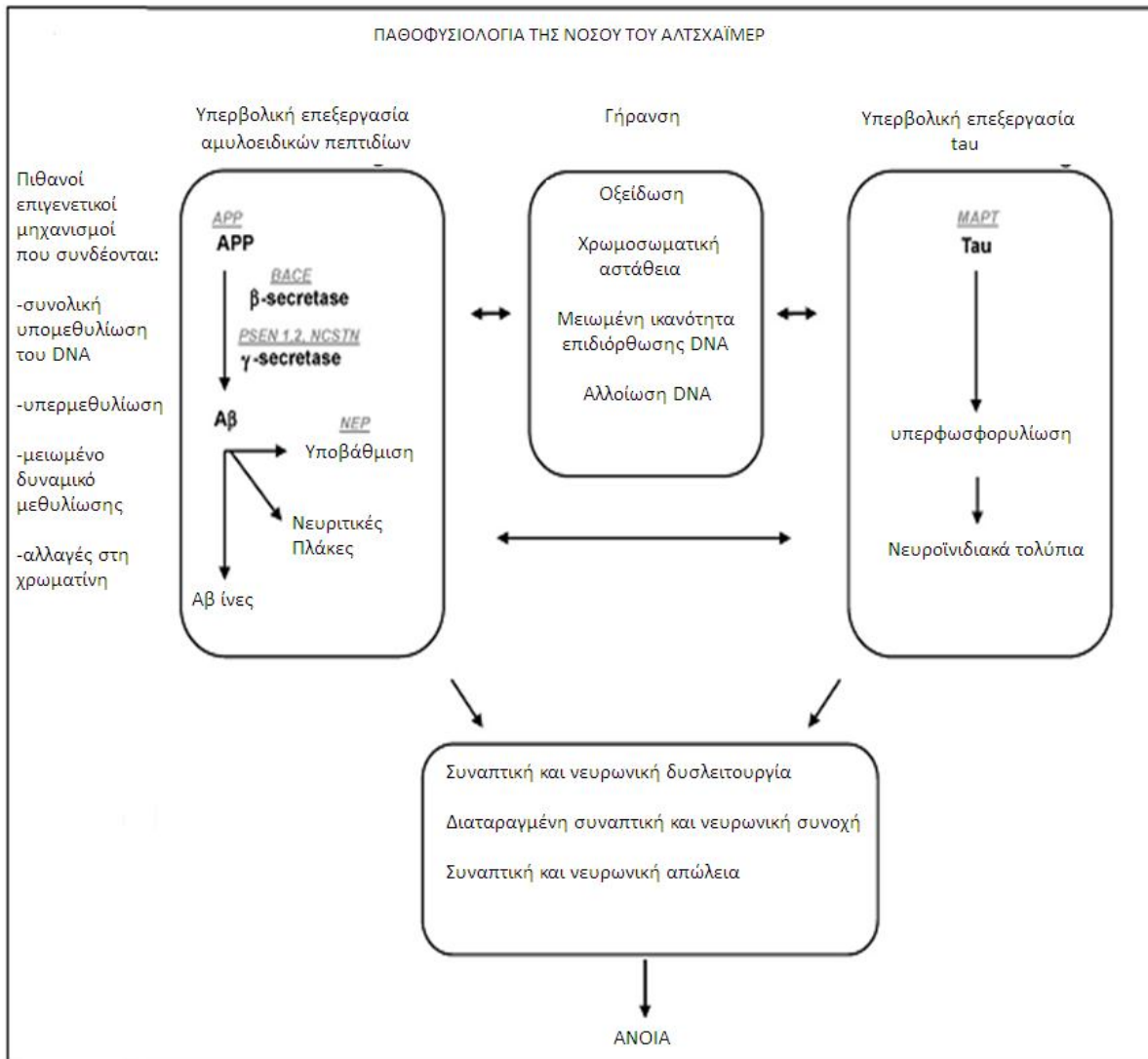
Table 2
Evidence on histone modifications in AD and AD model systems.

Histone modifications	Effect	Sample	Reference
HDAC1 trichostatin A	Attenuates hyperphosphorylation of tau	Neuronal cell cultures	Ding et al. (2008)
HDAC6	Elevated levels in AD patients compared to controls	Human cerebral cortical and hippocampal homogenates	Ding et al. (2008)
HDAC1 nicotinamide	Prevention of cognitive deficits through decreasing hyperphosphorylated tau	3x-Tg-AD mice	Green et al. (2008)
HDAC1 sodium 4-phenylbutyrate	Memory deficit and spatial learning reverse	Tg2576 mice	Ricobaraza et al. (2009)
Environmental enrichment and HDAC1 sodium butyrate	Learning behavior reinstatement and long-term memory recovery	CK-p25 mice	Fischer et al. (2007)
HDAC2	Memory impairment	HDAC10E and HDAC20E mice	Guan et al. (2009)
HDAC6	Tau decreases its activity leading to decreased degradation of tau aggregates	Tau(-/-) mice and neuronal cell cultures	Perez et al. (2009)
H2AX phosphorylation	Evidence of DNA damage in astrocytes	Hippocampal regions of AD patients' brains	Myung et al. (2008)
HDAC1 TSA	Restored decreased H4 acetylation after a learning task with in parallel improvement of memory deficit and synaptic dysfunction	APP/PS1 mice	Francis et al. (2009)

Πίνακας 11: Αλλαγές σε επιγενετικούς μηχανισμούς σε ασθενείς με ΝΑ και οι αντίστοιχες μελέτες [44]

3.7.4 Επιγενετική, νόσος του Αλτσχάιμερ και θεραπεία

Πέρα από την ερμηνεία της ΝΑ, η επιγενετική μπορεί να είναι και στόχος θεραπείας ή πρόληψης της ΝΑ, καθώς θεωρητικά τα γονίδια μπορούν να απομεθυλωθούν, μέσω φαρμάκων ή αλλαγών στον τρόπο ζωής, σε αντίθεση με τις μεταλλάξεις στα γονίδια. Για παράδειγμα, σε μελέτες σε ζώα, τροποποιήσεις στις ιστόνες μέσω φαρμάκων έχουν βοηθήσει τη μνήμη των πειραματόζωων να επανέλθει [52]. Συμπερασματικά είναι σαφές ότι οι επιγενετικοί μηχανισμοί, είτε ως υποψήφιος βιοδείκτης, είτε ως αιτία ή ακόμα και ως θεραπευτικός στόχος εκπροσωπεί μία νέα μεταβλητή στην έρευνα για τη ΝΑ, που έχει τη δυναμική να μεταβάλει ουσιαστικά τον τρόπο που αντιμετωπίζουμε την ασθένεια.



Εικόνα 16: Παθολογία της ΝΑ και οι επιγενετικοί μηχανισμοί που θεωρείται ότι συνδέονται με αυτή [44]

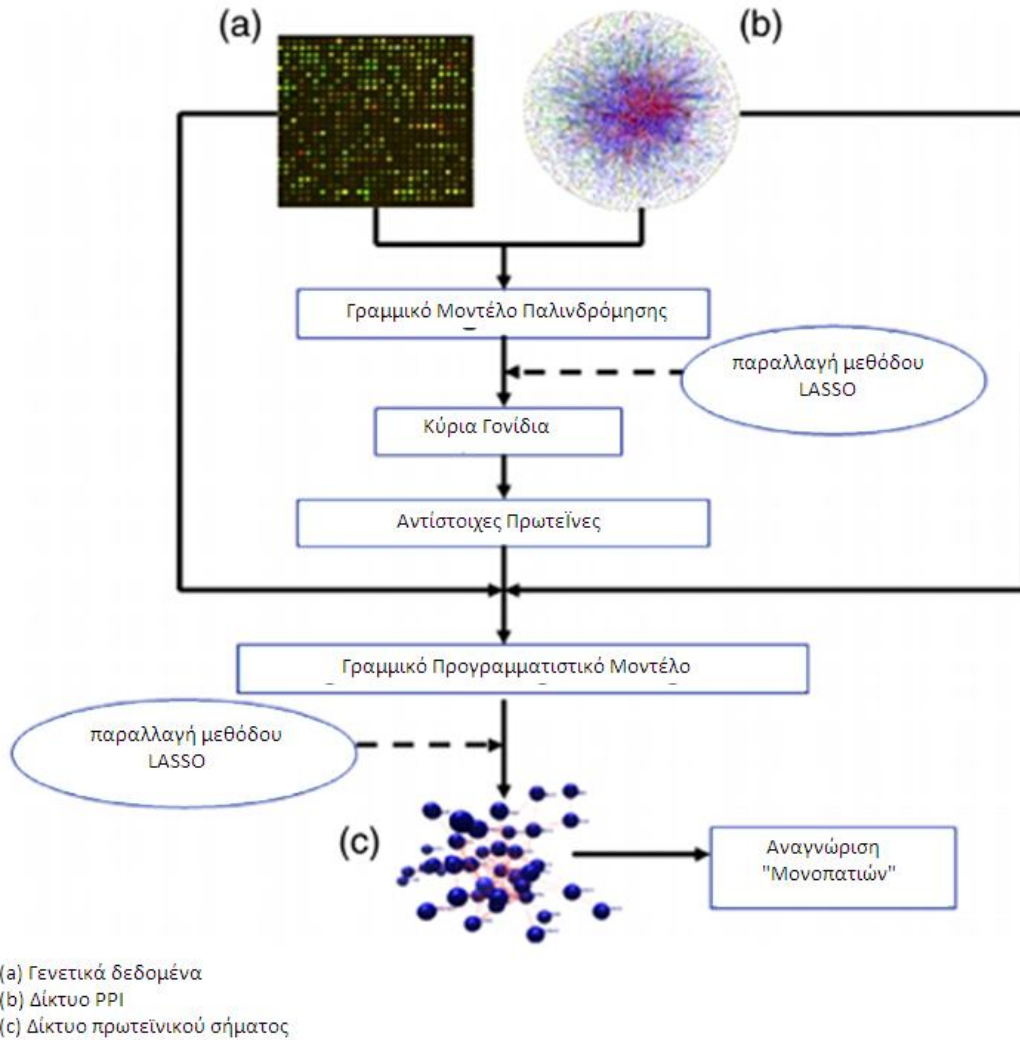
3.8 Πρωτεομική και γενετική στη νόσο του Αλτσχάιμερ

Όπως έχει ήδη αναφερθεί η αναγνώριση γονιδίων που επιβαρύνουν τη ΝΑ βοηθάει στην κατανόηση της νόσου. Ο πιο άμεσος τρόπος είναι η μελέτη των πρωτεϊνών που αποτελούν το μέσο έκφρασης αυτών των γονιδίων και των μεταξύ τους συσχετίσεων. Χαρακτηριστικό παράδειγμα

αποτελεί η ακόλουθη μελέτη, όπου οι ερευνητές εξέτασαν την έκφραση των γονιδίων της NA μέσω μιας βάσης δεδομένων αλληλεπίδρασης πρωτεϊνών (protein-protein interaction, PPI) [53].

Τα PPI δίκτυα περιλαμβάνουν τις πρωτεΐνες και τα μεταξύ τους “μονοπάτια” και συνιστούν έτσι έναν ενδιαφέροντα τρόπο ταξινόμησης των βιολογικών λειτουργιών. Στην περίπτωση της NA, ο καθορισμός των πρωτεϊνών που συμβάλλουν στην ανάπτυξή της, οι μεταξύ τους αλληλεπιδράσεις και ο τρόπος που αυτές μεταβάλλονται με την εξέλιξη της ασθένειας μπορούν να διαφωτίσουν σημαντικά την αιτιολογία της νόσου [54].

Η διαδικασία που ακολουθήθηκε ήταν σχετικά απλή: οι ερευνητές πήραν τα γονίδια που συνδέθηκαν με τη NA σε δύο προηγούμενα πειράματα, ένα μοντέλο σε πειραματόζωα και μία έρευνα σε νοσούντες από NA. Στη συνέχεια εφήρμοσαν ένα γραμμικό μοντέλο παλινδρόμησης προκειμένου να εξαχθούν τα γονίδια που είναι βασικά για τη NA. Οι πρωτεΐνες που εκφράζονται μέσω αυτών των γονιδίων αποτέλεσαν τον πυρήνα του δικτύου των πρωτεϊνών. Τα μεταξύ τους “μονοπάτια” και οι επιπλέον πρωτεΐνες που αλληλεπιδρούν με αυτές χαρτογραφήθηκαν μέσω μιας βάσης δεδομένων PPI.



Εικόνα 17: Παραγωγή γράφου πρωτεϊνικού σήματος από γενετικά δεδομένα για NA και πρωτεϊνική βάση δεδομένων έπειτα από στατιστική επεξεργασία [53]

Ο χάρτης που προέκυψε αποκαλύπτει επιπλέον αλληλεπιδράσεις μεταξύ πρωτεϊνών και την αντίστοιχη επίδρασή τους στη βιολογία του οργανισμού. Παράλληλα επιτρέπει την αξιολόγηση άλλων μοντέλων για τη NA. Στην προκειμένη για παράδειγμα τα αποτελέσματα του μοντέλου από τα ποντίκια προσέγγιζε πολύ το μοντέλο από τους νοσούντες από NA. Δεδομένου ότι η μέθοδος είναι σχετικά απλή, μπορεί όντως να χρησιμοποιηθεί για την επιβεβαίωση των αποτελεσμάτων άλλων ερευνών.

Παρεμφερείς έρευνες έχουν γίνει και άλλες, όπως η έρευνα των Liu et al [54], όπου οι ερευνητές χαρτογράφησαν το πρωτεϊνικό αποτύπωμα της NA σε διαφορετικά στάδια της νόσου. Οι χάρτες που παρήγαγαν γίνονταν όλο και πιο περίπλοκοι με την εξέλιξη της νόσου, δηλαδή όλο και περισσότερες

πρωτεΐνες συμμετείχαν στις διαδικασίες της ΝΑ. Έτσι προκύπτει ένας ενδιαφέρων υποψήφιος δείκτης για το ρυθμό εξέλιξης της νόσου, ενώ και ο μηχανισμός συντονισμένης δράσης των γονιδίων μπορεί να διαλευκανθεί. Τέλος, η εργασία επιβεβαιώνει το ρόλο σημαντικών για τη ΝΑ πρωτεϊνών όπως η APP.

ΚΕΦΑΛΑΙΟ 4: ΣΥΝΔΥΑΣΤΙΚΕΣ ΜΕΘΟΔΟΙ ΒΙΟΔΕΙΚΤΩΝ ΓΙΑ ΤΗ ΝΟΣΟ ΤΟΥ ΑΛΤΣΧΑΪΜΕΡ

4.1 Συνδυαστικές μέθοδοι

Τα πολυτροπικά (multimodal) συστήματα είναι ο χώρος όπου όλα τα πορίσματα από τις παραπάνω έρευνες εισάγονται σε ένα κοινό σύστημα που επιδιώκει να επεξεργαστεί τα δεδομένα με μία κοινή μέθοδο και να εξάγει νέες συσχετίσεις. Έτσι, στο ίδιο διαγνωστικό σύστημα συνυπάρχουν δεδομένα από γενετικές έρευνες, νευροαπεικόνιση, δείκτες CSF αλλά και επιδημιολογικά χαρακτηριστικά, όπως ηλικία και χρόνια εκπαίδευσης. Η έρευνα σε αυτόν τον τομέα είναι εκτεταμένη και προσπαθεί να απαντήσει κατεξοχήν στα ακόλουθα ερωτήματα:

- Βελτιώνει τη διαγνωστική μας ικανότητα η διαχείριση πολλών βιοδεικτών έναντι ενός; Παρότι η απάντηση σε αυτό το ερώτημα διαισθητικά είναι καταφατική, μέχρι πρόσφατα δεν ήταν δεδομένο ότι τα αποτελέσματα δικαιολογούσαν το επιπλέον κόστος. Πολλές σύγχρονες μελέτες όμως, έχουν απαντήσει στο παραπάνω ερώτημα υπέρ της ταυτόχρονης επεξεργασίας πολλών βιοδεικτών.
- Ποιος συνδυασμός βιοδεικτών είναι ο βέλτιστος; Πολλοί δείκτες προσφέρουν πληροφορία που είναι παρεμφερής με αυτή που προσφέρουν άλλοι και άρα η προσθήκη τους είναι περιττή. Επίσης αυτό το ερώτημα συναρτάται και άλλων παραγόντων εκτός του ιδεατού περιβάλλοντος του εργαστηρίου, όπως το κόστος των εξετάσεων και την ανάγκη εξειδίκευσης του συστήματος στον ασθενή.
- Ποια υπολογιστική προσέγγιση ταξινομεί καλύτερα τα δεδομένα; Αξίζει να χρησιμοποιηθούν πολυπαραγοντικές τεχνικές όπου τα δεδομένα επεξεργάζονται ταυτόχρονα και αλληλοεπηρεάζονται (για παράδειγμα ο αλγόριθμος paraICA που είδαμε πιο πάνω) ή απλούστερα συστήματα, όπως ένα SVM που ενοποιεί τα αποτελέσματα της επεξεργασίας των δεικτών, είναι επαρκή;

Γενικά η έρευνα σε αυτόν τον τομέα μπορεί να διαχωριστεί σε δύο στάδια: στη δημιουργία ενός συνολικού βιοδείκτη και στην πρακτική ένταξή του στη διάγνωση ως ένα σύστημα υποστήριξης κλινικών αποφάσεων (Clinical Decision Support System, CDSS). Ως εκ τούτου οι μελέτες που θα εξετάσουμε δίνουν βάρος σε έναν από τους παραπάνω τομείς. Συστήματα όπως τα παραπάνω είναι προφανώς περίπλοκα και αφορούν τη διαχείριση γιγαντιαίου όγκου πληροφορίας. Απαιτείται συνεπώς η χρήση των πλέον πρόσφατων τεχνικών εξόρυξης δεδομένων.

Ιστορικά, η ιδέα της πολυτροπικής ταξινόμησης στη ΝΑ προέκυψε από τη διαπίστωση ότι βιοδείκτες, όπως για παράδειγμα οι CSF δείκτες, προσφέρουν διαφορετική πληροφορία από τις μορφολογικές αλλαγές στην MRI και συνεπώς ο συνδυασμός τους ίσως προσφέρει κάποια βελτίωση στην ταξινόμηση [55]. Καθώς το μεγαλύτερο ποσοστό δεικτών για τη ΝΑ έχει προκύψει σχετικά πρόσφατα, όπως επίσης και η υπολογιστική τεχνογνωσία για τη διαχείριση πολλών πηγών δεδομένων, η πολυτροπική ταξινόμηση έχει μεγάλο ερευνητικό ενδιαφέρον.

Ο συνδυασμός βιοδεικτών όντως βελτιώνει την ακρίβεια της πρόβλεψης, όπως θα δούμε στην πορεία. Ωστόσο το πόσο μπορεί να βελτιώσει την πρόβλεψη και από ποιο σημείο και μετά η πληροφορία που προσφέρει ένας επιπλέον δείκτης καλύπτεται από κάποιον άλλο είναι ανοιχτά θέματα.

Η αξιολόγηση των παρακάτω μεθόδων γίνεται κυρίως με βάση τις επιδόσεις του κάθε συστήματος στην ταξινόμηση. Ο κύριος δείκτης είναι η ακρίβεια ενώ συχνά χρησιμοποιείται και η υψηλή ειδικότητα και η ευαισθησία. Όπως έχουμε αναφέρει στην περίπτωση της ΝΑ είναι πολύ σημαντικότερη η υψηλή ευαισθησία, παρά η υψηλή ειδικότητα, γιατί το κόστος λανθασμένης ταξινόμησης ενός υγιούς είναι μικρό. Η αξιολόγηση των μεθόδων περιπλέκεται από το γεγονός ότι δεν θέτουν όλες τα ίδια πειράματα. Σχεδόν παντού υπάρχει η σύγκριση ΝΑ με ομάδα ελέγχου, όπου υπάρχουν ποσοστά ακρίβειας πάνω από 90% με τις συνδυαστικές μεθόδους. Το αμέσως πιο κοινό πρόβλημα είναι η διάκριση ασθενών με MCI και ομάδας ελέγχου, που το ποσοστό ακρίβειας κινείται στο 75%. Ταυτόχρονα επιχειρούνται και άλλα πιο ενδιαφέροντα πειράματα όπως η διάκριση ασθενών ΝΑ από ασθενείς MCI, με ακρίβεια γύρω στο 70%, ασθενών ΝΑ από ασθενείς με στάσιμο MCI και ασθενών με στάσιμο MCI από ασθενείς με εξελισσόμενο MCI. Τα τελευταία είναι τα δυσκολότερα πειράματα με αρκετά χαμηλή ακρίβεια ακόμη. Δυστυχώς ο αριθμός των ερευνών που τα περιλαμβάνουν δεν είναι αρκετά μεγάλος.

Σχεδόν σε όλες τις μεθόδους επιχειρείται κάποια σύγκριση της εκάστοτε πολυτροπικής μεθόδου με μεμονωμένες πηγές δεδομένων ή με συνδυασμό μίας κατηγορίας πηγών δεδομένων, για παράδειγμα απεικονιστικά δεδομένα. Δεδομένου ότι εξελίσσονται στο πλαίσιο της ίδιας έρευνας είναι ευκολότερο να τα συγκρίνουμε και συνολικά επιβεβαιώνουν ότι η ταξινόμηση με πολυτροπικά συστήματα έχει πολύ μεγαλύτερη ακρίβεια, ως και 10% καλύτερη, από την ταξινόμηση με τους καλύτερους μεμονωμένους βιοδείκτες.

Συνολικά πάντως όλα τα παρακάτω συστήματα είναι αρκετά πρόσφατα, από το 2010 και μετά, και άρα δεν έχουν επαληθευθεί στον επιθυμητό βαθμό τα αποτελέσματα. Ιδιαίτερα στην εφαρμογή των πολυτροπικών συστημάτων στη διαγνωστική διαδικασία τα περισσότερα συστήματα είναι αρκετά πρώιμα.

4.1.1 Αξιολόγηση συστημάτων συνδυασμού πηγών δεδομένων στη νόσο του Αλτσχάιμερ: Πλατφόρμα πολυπύρηνης εκμάθησης

Χαρακτηριστικό παράδειγμα αποτελεί η σχετική εργασία του Hinrichs [56]. Οι ερευνητές επιδίωξαν να προβλέψουν τη μετάβαση σε ΝΑ των ασθενών με MCI, με τη δημιουργία ενός πολυτροπικού δείκτη για την ασθένεια (multi-modality disease marker) μέσω μιας πλατφόρμας μάθησης υπολογιστών (machine learning framework). Η κεντρική ιδέα είναι ότι κάποιες παθογένειες δεν είναι εμφανείς σε μία εξέταση αλλά φαίνονται στις υπόλοιπες και άρα, η επεξεργασία τους μέσω ενός μόνο συστήματος ίσως βελτιώσει τις συνολικές επιδόσεις [57].

Η βάση της μεθόδου ήταν η πολυπύρηνη εκμάθηση (Multi-Kernel Learning, MLK), που είναι μία πλατφόρμα σχεδιασμένη για το χειρισμό πολλών πηγών δεδομένων με τη μείωση της πολυπλοκότητας του προβλήματος. Ουσιαστικά η ιδέα πίσω από την MLK είναι κάθε πηγή δεδομένων να ταξινομείται μέσω ενός πυρήνα (kernel function), στην πράξη ένα SVM που λειτουργεί ως ταξινομητής. Μετά οι πυρήνες συνδυάζονται γραμμικά (κάθε συνάρτηση πολλαπλασιάζεται με βάρη - παράγοντες βελτιστοποίησης) ώστε να προκύψει ο συνολικός πυρήνας. Οι παράγοντες βελτιστοποίησης επιλέγονται έτσι ώστε η ομάδα εκπαίδευσης να διαχωρίζεται με τον καλύτερο δυνατό τρόπο, ενώ ταυτόχρονα να μην υπάρχει υπερπαραμετροποίηση (overfitting). Δεδομένου ότι η παθολογία της ΝΑ είναι ετερογενής, πολλοί διαφορετικοί δείκτες πρέπει να χρησιμοποιηθούν για την πρόγνωση της πορείας της ασθένειας, χωρίς όμως να χαθεί η ικανότητα του συστήματος να γενικεύει επιτυχώς στην πορεία άλλων ασθενών. Το παραπάνω πλαίσιο λοιπόν είναι ένας ταιριαστός αλγόριθμος για τη διαχείριση των δεδομένων της ΝΑ.

Για την πραγμάτωση του πειράματος επιλέχθηκαν από τη βάση της ADNI άτομα που έκαναν 2 εξετάσεις MRI με διαφορά περίπου 2 ετών και αντίστοιχα 2 εξετάσεις FDG-PET. Αφού αφαιρέθηκαν οι εικόνες με εμφανή θόρυβο (προτού ξεκινήσει το πείραμα) έμειναν 233 άτομα (48 με ΝΑ, 66 υγιείς και 119 με MCI) που αποτέλεσαν το δείγμα του πειράματος.

Οι εξετάσεις που ήταν διαθέσιμες για το ανωτέρω δείγμα ήταν MRI και FDG-PET απεικονίσεις, CSF βιοδείκτες που έχουν συνδεθεί με τη ΝΑ, όπως Αβ-42 και t-tau, νευροψυχολογικές δοκιμασίες και πληροφορίες για το APOE γονίδιο. Με βάση αυτά τα δεδομένα εκπαίδευσαν το σύστημα στις εξετάσεις της ομάδας ελέγχου και της ομάδας που νοσούσε από ΝΑ και μετά εφήρμοσαν το δείκτη που προέκυψε στην ομάδα που είχε MCI, για να διαγνώσει ποιοι θα ανέπτυσαν ΝΑ. Για να αξιολογήσουν την επίδοση του συστήματος εφήρμοσαν το ίδιο πείραμα και στους ταξινομητές που προέκυψαν από μία μόνο εξέταση, καθώς και σε ένα απλό SVM που άθροιζε τους προηγούμενους ταξινομητές.

Ακολουθεί πίνακας με τις επιδόσεις της ταξινόμησης στο βασικότερο πείραμα, δηλαδή στη διάκριση ασθενών με ΝΑ από την ομάδα ελέγχου (υγιείς), οργανωμένα σε τέσσερις διαφορετικές ομάδες δεδομένων (απεικονιστικά, βιολογικά χαρακτηριστικά -δηλαδή CSF και πλάσμα- και γνωσιακές δοκιμασίες) και στις συνολικές επιδόσεις τους.

Modalities used	Accuracy	Sensitivity	Specificity	Area under ROC
Imaging modalities	0.876	0.789	0.938	0.944
Biological measures	0.704	0.581	0.794	0.767
Cognitive scores	0.912	0.892	0.926	0.983
All modalities	0.924	0.867	0.966	0.977

Πίνακας 12: Οι επιδόσεις της ταξινόμησης του συνολικού απεικονιστικού δείκτη, του συνολικού βιολογικού δείκτη, του συνολικού γνωσιακού δείκτη και του ολικού δείκτη στην ταξινόμηση ασθενών με ΝΑ και υγιών [56]

Συνολικά βλέπουμε ότι ο ολικός δείκτης έχει τις καλύτερες επιδόσεις, με ακρίβεια 92%. Ο γνωσιακός δείκτης έχει επίσης πολύ καλές επιδόσεις, κάτι αναμενόμενο δεδομένου ότι έχουμε το απλό πείραμα διάκρισης ΝΑ από υγιείς. Παρατηρούμε ότι ο απεικονιστικός δείκτης επίσης παράγει ακρίβεια κοντά στο 88%, δηλαδή 4% χαμηλότερα από τον ολικό. Οι βιολογικοί δείκτες με ακρίβεια 70%, δείχνουν γιατί μέχρι τώρα οι μορφομετρικοί προτιμούνται στην έρευνα για τη ΝΑ.

Στην πρόβλεψη της μετάβασης από MCI σε ΝΑ η εικόνα αλλάζει: Η καλύτερη πρόβλεψη έγινε από το δείκτη που συνδυάζει όλες τις απεικονιστικές τεχνικές, χωρίς όμως να περιλαμβάνει τους υπόλοιπους δείκτες. Η επίδοση του συστήματος που περιείχε όλους τους δείκτες ήταν λίγο χαμηλότερη, ενώ το ίδιο ίσχυε και για το δείκτη που περιείχε τις νευροψυχολογικές δοκιμασίες. Το συμπέρασμα που εξάγεται από το παραπάνω αποτέλεσμα είναι ότι η ΝΑ μπορεί να ανιχνευθεί σε πρώιμο στάδιο καλύτερα από τις απεικονιστικές τεχνικές παρά από τις νευροψυχολογικές εξετάσεις, καθώς οι γνωσιακές αλλαγές δεν έχουν πραγματοποιηθεί ακόμη. Ενδιαφέρον είναι παράλληλα ότι οι βιολογικοί και γενετικοί δείκτες είχαν κακές επιδόσεις στην πρόβλεψη της νόσου. Το τελικό συμπέρασμα ωστόσο είναι ότι οι δείκτες που περιλάμβαναν μεμονωμένες εξετάσεις είχαν συστηματικά χειρότερες επιδόσεις από όσες πέρασαν από το MLK σύστημα, προβάλλοντας έτσι την πολυτροπική διαχείριση των δεδομένων ως ένα πολύ καλό εργαλείο ταξινόμησης.

Παρεμφερή και καλύτερα ακόμη αποτελέσματα έχουν προκύψει και από άλλες έρευνες πάνω σε πολυτροπικά συστήματα. Η σημαντική διαφορά όμως είναι ότι στη συγκεκριμένη μελέτη οι ερευνητές δημιούργησαν έναν “παγκόσμιο” δείκτη και απέφυγαν έτσι την επιρροή άλλων παραγόντων, όπως την

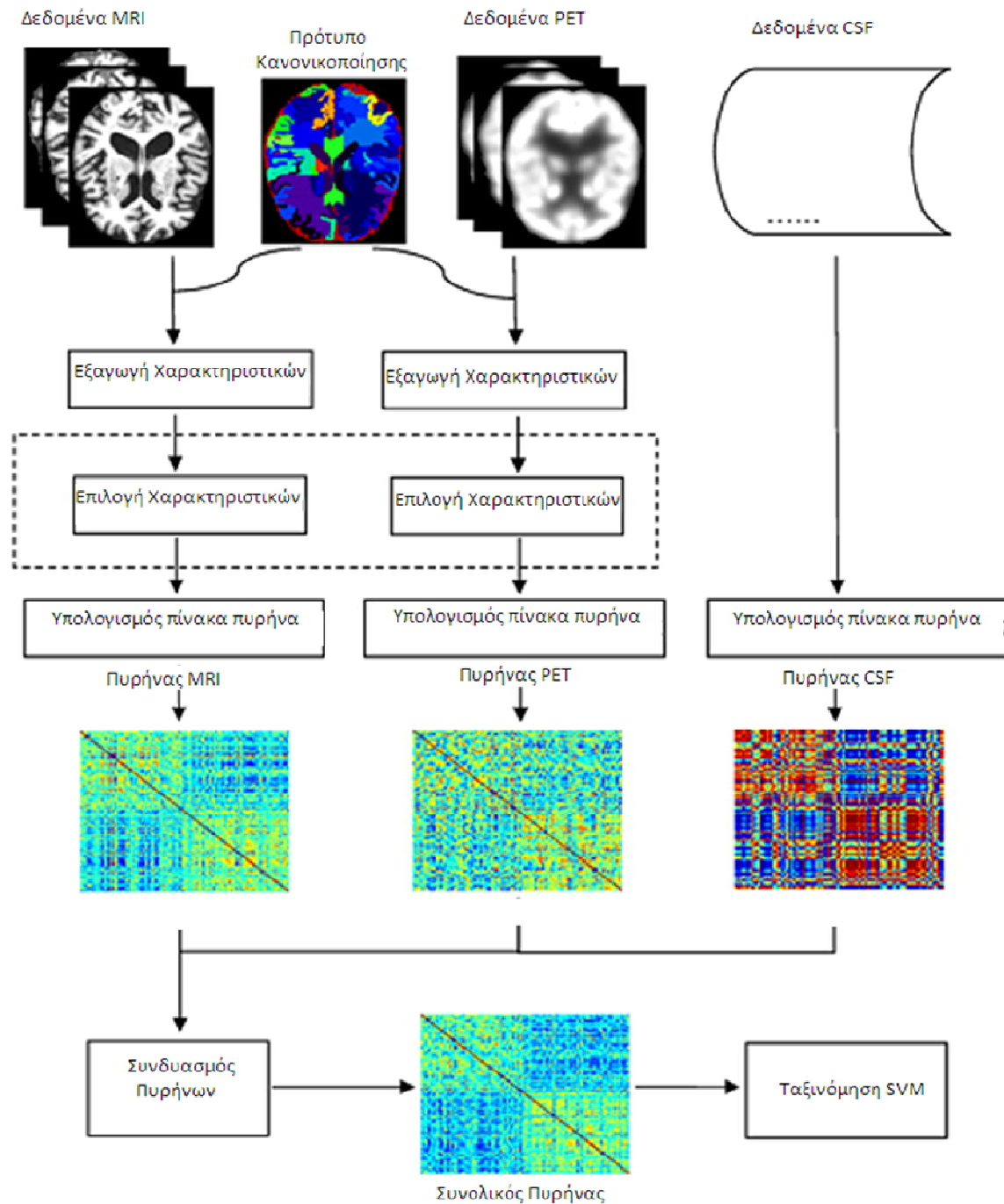
έλλειψη κάποιων εξετάσεων. Τέλος, το σύστημα λόγω της δομής του, μπορεί εύκολα να μεταβληθεί ώστε να συμπεριλάβει άλλες πιο αποτελεσματικές μεθόδους προεπεξεργασίας δεδομένων ή νέους δείκτες. Η MLK πλατφόρμα λοιπόν αποτελεί μία σοβαρή πρόταση στην πολυτροπική διαχείριση των εξετάσεων των ασθενών με ΝΑ ακόμη και σήμερα.

4.1.2 Αξιολόγηση συστημάτων συνδυασμού πηγών δεδομένων στη νόσο του Αλτσχάιμερ: Χρήση διανυσμάτων υποστήριξης

Όπως έχει προαναφερθεί, η χρήση των πολυτροπικών τεχνικών βελτιώνει αισθητά την απόδοση της ταξινόμησης του δείγματος. Σαφή ένδειξη προς αυτή την κατεύθυνση μας παρέχει και η έρευνα των Zhang et al., όπου οι ερευνητές μέτρησαν ικανή βελτίωση ακόμη και με ένα σχετικά απλό πολυπύρηνο σύστημα ταξινόμησης [58].

Οι ερευνητές επέλεξαν να αξιοποιήσουν τρεις διαφορετικές πηγές δεδομένων: MRI, PET και CSF. Το δείγμα προέκυψε από την ADNI με τη απόρριψη όσων δεν είχαν και τα τρία παραπάνω στοιχεία διαθέσιμα.

Η επιστημονική ομάδα εκπαίδευσε ένα πολυπύρηνο SVM. Ουσιαστικά πρόκειται για ένα απλό SVM όπου η συνάρτηση χαρτογράφησης k είναι η σύνθεση των πυρήνων για κάθε διαφορετική μορφή δεδομένων, αφού ένα απλό SVM έχει εκπαιδευτεί σε κάθε πηγή δεδομένων.



Εικόνα 18: Σχηματικό διάγραμμα της πολυπύρηνης πολυτροπικής επεξεργασίας και ταξινόμησης δεδομένων [58]

Τα προβλήματα που έθεσε η ομάδα στο σύστημα είναι δύο: ο διαχωρισμός υγιών από ασθενείς με ΝΑ και ο διαχωρισμός υγιών από ασθενείς με ΜCΙ. Τα αποτελέσματα ενισχύουν τα ήδη γνωστά για τα

πολυτροπικά συστήματα: το πολυπύρηνο σύστημα έχει καλύτερη απόδοση από οποιαδήποτε μεμονωμένη πηγή δεδομένων, με αντίστοιχη απόδοση με τα άλλα συστήματα τέτοιας φύσης.

Σύγκριση των επιδόσεων της ταξινόμησης της πολυτροπικής μεθόδου και των απλών μεθόδων. Οι αριθμοί μέσα σε παρενθέσεις υποδεικνύουν την χαμηλότερη και υψηλότερη επίδοση που πέτυχε κάθε μέθοδος στο σύνολο των πειραμάτων

Methods	AD vs. HC			MCI vs. HC		
	ACC (%)	SEN (%)	SPE (%)	ACC (%)	SEN (%)	SPE (%)
MRI	86.2 (82.9–89.0)	86 (82.7–88.7)	86.3 (83.1–89.1)	72.0 (68.4–74.7)	78.5 (75.6–80.6)	59.6 (55.1–63.7)
CSF	82.1 (80–84.9)	81.9 (80–84.7)	82.3 (80–85.1)	71.4 (68.2–73.3)	78 (75.6–79.4)	58.8 (54.3–61.7)
PET	86.5 (82.9–90.5)	86.3 (82.7–90.3)	86.6 (83.1–90.6)	71.6 (67.4–74.7)	78.2 (75–80.6)	59.3 (52.9–63.7)
Combined	93.2 (89.0–96.5)	93 (88.7–96.3)	93.3 (89.1–96.6)	76.4 (73.5–79.7)	81.8 (79.4–84.4)	66.0 (62.6–70.3)
Baseline	91.5 (88.5–96.5)	91.4 (88.3–96.3)	91.6 (88.6–96.6)	74.5 (71.9–78.2)	80.4 (78.3–83.3)	63.3 (59.7–68.3)

AD = Alzheimer's Disease, MCI = Mild Cognitive Impairment, HC = Healthy Control, ACC = classification ACCuracy, SEN = SENSitivity, SPE = SPEcificity.

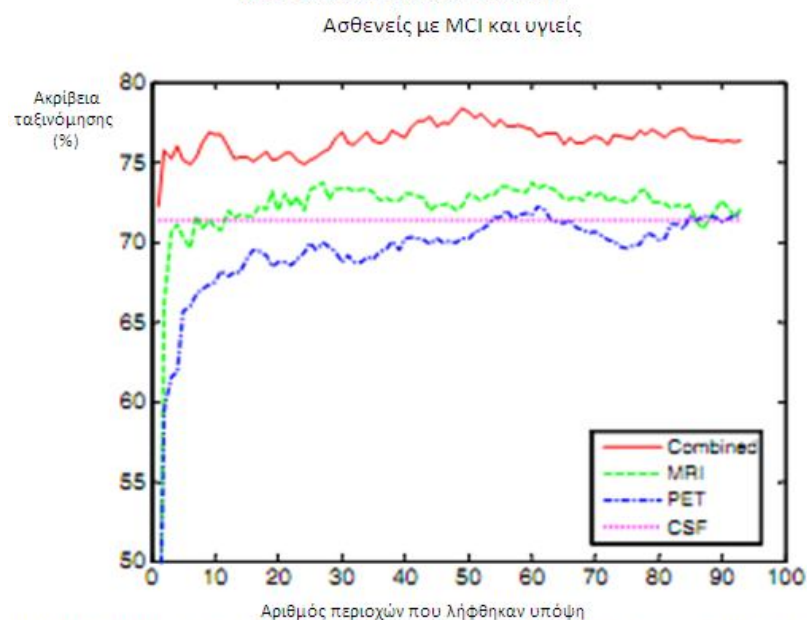
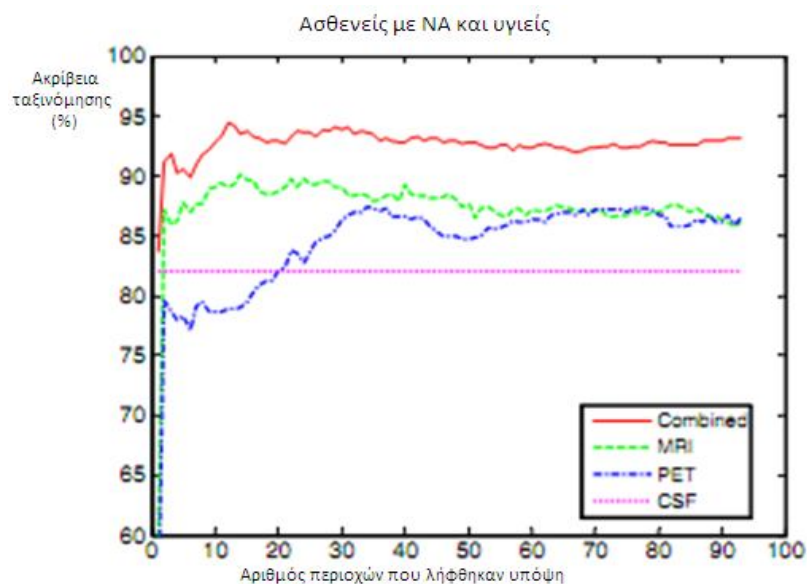
Πίνακας 13: Σύγκριση των επιδόσεων της ταξινόμησης των διαφορετικών πηγών δεδομένων και των ενοποιημένων πηγών [58]

Βλέπουμε ότι ο συνολικός δείκτης στη διάκριση NA από υγιείς είχε επίδοση 93.2%, δηλαδή λίγο υψηλότερη από αυτή που έγινε εφικτή με το MLK στην προηγούμενη εργασία. Η ταξινόμηση MCI από υγιείς έγινε με ακρίβεια 76.4%, που είναι επίσης συγκρίσιμη με τα αποτελέσματα της έρευνας στο χώρο, αν και λίγο χαμηλή.

Ωστόσο οι ερευνητές σε αυτό το σημείο έθεσαν ένα επιπλέον πρόβλημα: διαχώρισαν την MCI ομάδα σε αυτούς που έμειναν σταθεροί και σε αυτούς που ανέπτυξαν NA στην πορεία, δηλαδή έθεσαν πρακτικά το πρόβλημα διαχωρισμού στάσιμου MCI από υγιείς και εξελισσόμενου MCI από υγιείς. Διαπίστωσαν ότι ενώ η ταξινόμηση των ασθενών με MCI γενικά είχε ακρίβεια 76.4% , η ομάδα με εξελισσόμενο MCI διαγνώστηκε με MCI με ακρίβεια 91.5%, ενώ η ομάδα με στάσιμο MCI με ακρίβεια 73.4%, δηλαδή το σύστημα κατέταξε σωστά όσους ανέπτυξαν αργότερα NA. Αυτό είναι ένα σημαντικό αποτέλεσμα, γιατί σε μια πιθανή εφαρμογή διάγνωσης το κόστος της λανθασμένης ταξινόμησης ενός ασθενούς με στάσιμο MCI είναι μικρό, ενώ η ορθή ταξινόμηση ασθενών με εξελισσόμενο MCI είναι σαφώς σημαντικότερο θέμα. Ένας διαχωρισμός όπως ο παραπάνω δεν έχει εφαρμοστεί σε πολλές έρευνες και θα είχε ενδιαφέρον να διαπιστωθεί τι απόδοση έχουν άλλα συστήματα ως μία επέκταση του κριτηρίου της ευαισθησίας (αν και οι πολύ μικρές ομάδες δείγματος ενέχουν κάποιους στατιστικούς κινδύνους).

Τέλος, οι ερευνητές δοκιμάζουν πόσα χαρακτηριστικά (εικόνες περιοχών του εγκεφάλου) απαιτούνται για βέλτιστη απόδοση. Εκεί διαπιστώνεται ότι το σύστημα σταματά να βελτιώνεται με την

προσθήκη επιπλέον περιοχών από ένα σημείο και πέρα. Αυτό δεν ήταν αναμενόμενο, γιατί η παθολογία της ΝΑ εκδηλώνεται με πολλούς τρόπους και το σύστημα θα κληθεί να διαγνώσει για παράδειγμα είτε τον εγκέφαλο με μεγάλη φθορά στον ιππόκαμπο είτε τον εγκέφαλο με μικρότερες φθορές σε πολλές περιοχές. Ταυτόχρονα όμως υπάρχει και ένα ανώτατο στατιστικό όριο με τα συγκεκριμένα δείγματα που περιορίζει τις δυνατότητές μας.



Η ακρίβεια ταξινόμησης για τις δύο διαφορετικές δοκιμασίες και τις τέσσερις διαφορετικές μεθόδους (χρώμα) σε σχέση με τον αριθμό των περιοχών. Παρατηρούμε ότι ο συνδυασμός τους αποδίδει πάντα καλύτερα αποτελέσματα και ότι η αύξηση του αριθμού των περιοχών δεν βελτιώνει το αποτέλεσμα από ένα σημείο και πέρα.

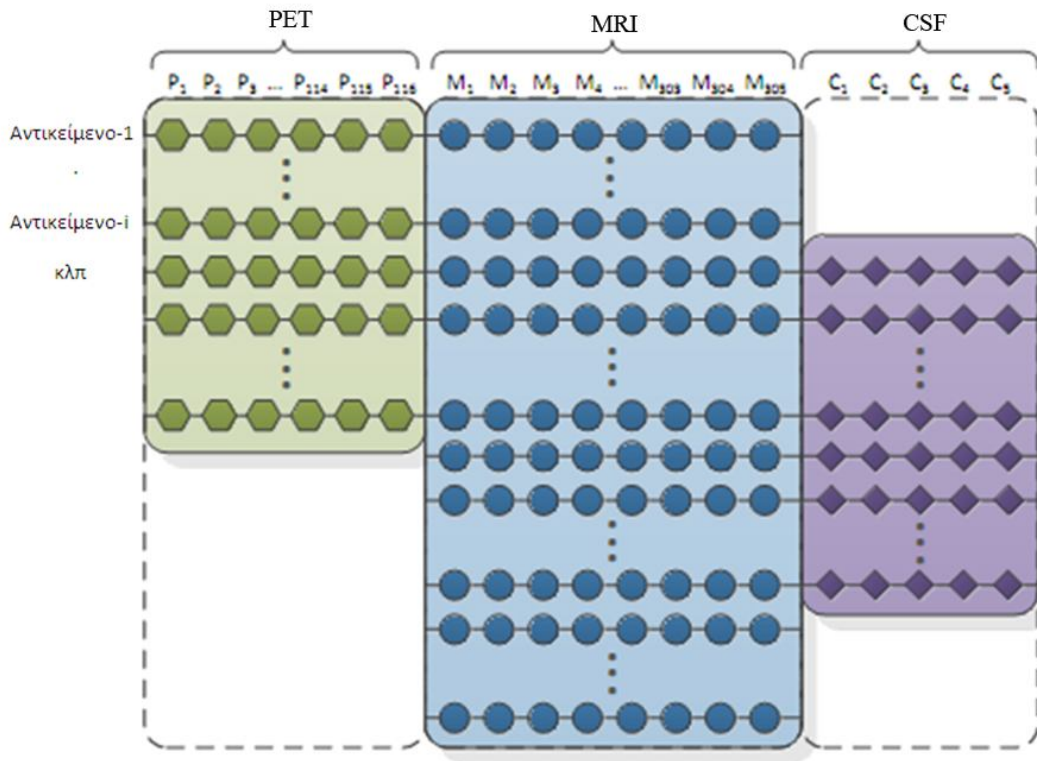
Διάγραμμα 4: Επίδοση των διαφόρων πηγών δεδομένων στην ταξινόμηση ασθενών με Αλτσχάιμερ έναντι υγιών και ασθενών με ΜCI έναντι υγιών όσον αφορά την ακρίβεια και με διαφορετικό αριθμό περιοχών

Καταλήγουμε συνεπώς ότι η έρευνα προσφέρει κάποιες ενδιαφέρουσες προοπτικές, που επιβεβαιώνουν ότι η πολυτροπική ταξινόμηση προσφέρει μεγάλες δυνατότητες στη ΝΑ. Τα αποτελέσματα που παράγονται αν και δεν είναι τα απόλυτα καλύτερα είναι παραπάνω από αποδεκτά με 93% ακρίβεια στην ταξινόμηση ΝΑ και υγιών και υψηλή ευαισθησία σε όλες τις περιπτώσεις, δηλαδή ελαφρώς καλύτερα από αυτά της MLK πλατφόρμας. Ωστόσο, νεότερες έρευνες συνδυάζουν περισσότερες πηγές δεδομένων ή αντιμετωπίζουν προβλήματα όπως ο διαχωρισμός στάσιμου και εξελισσόμενου MCI, που πραγματικά είναι στην αιχμή της έρευνας για τη ΝΑ. Τέλος, η έρευνα είχε μικρό δείγμα, αφού αφαιρέθηκαν εκείνοι οι ασθενείς της ADNI για τους οποίους δεν ήταν διαθέσιμη κάποια από τις τρεις εξετάσεις που χρησιμοποιήθηκαν (MRI, PET, CSF).

4.1.3 Εισαγωγή της συμπλήρωσης δεδομένων σε συστήματα συνδυασμού πηγών δεδομένων

Μία άλλη ερευνητική πρόταση που λύνει θεαματικά ένα καίριο πρόβλημα έρχεται πάλι από την ομάδα του Thompson [59]. Η ομάδα διαπίστωσε ότι στις περισσότερες πολυπαραγοντικές έρευνες για τη ΝΑ ένα μεγάλο μέρος των υπαρχόντων δεδομένων απορρίπτεται, γιατί δεν υπάρχουν όλοι οι δείκτες για έναν ασθενή. Συνεπώς, καθώς εφαρμόζονται μέθοδοι που απαιτούν την αλληλεπίδραση μεταξύ των δεδομένων, απορρίπτεται το σύνολο των εξετάσεων για το συγκεκριμένο ασθενή. Στην περίπτωση της ADNI για παράδειγμα, το δείγμα μειώνεται στο μόλις 25% του αρχικού έπειτα από την παραπάνω αφαίρεση.

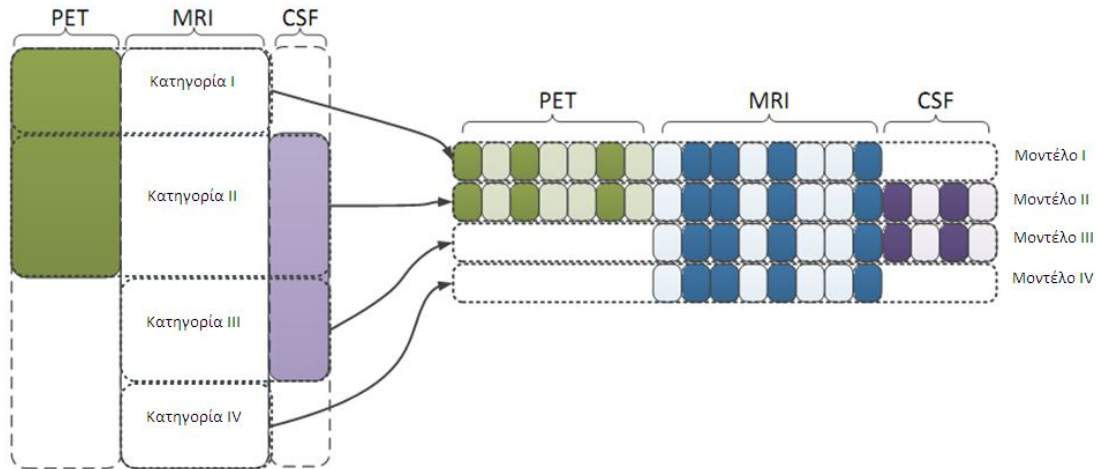
Αντί να απορρίπτεται ένα τόσο μεγάλο μέρος του δείγματος, προτείνεται η χρήση αλγορίθμων συμπλήρωσης δεδομένων. Ο συγκεκριμένος τομέας έχει πρόσφατα επιδείξει εξαιρετικές λύσεις. Όμως στις βάσεις δεδομένων για ΝΑ συνήθως δεν λείπουν στοιχεία σε τυχαίες θέσεις, όπως θα συνέβαινε αν για παράδειγμα μια εικόνα απορριφθεί λόγω θορύβου. Αντίθετα λείπουν ολόκληρα μπλοκ από πληροφορίες γιατί δεν υπήρχαν οι σχετικοί πόροι για κάποια εξέταση ή κρίθηκε ότι κάποιος βιοδείκτης δεν χρειαζόταν. Και η ίδια η ADNI για παράδειγμα θα δημοσιεύσει σύντομα δεδομένα για 3 νέους δείκτες, αλλά θα πάρει τιμές για τον κάθε έναν μόνο από το 1/3 του πληθυσμού που χρησιμοποιεί. Απαιτείται λοιπόν κάποιο σύστημα που θα συμπληρώνει ολόκληρα μπλοκ από πληροφορίες, κάτι που είναι ακόμη δυσκολότερο από τη συμπλήρωση τυχαίων θέσεων.



Εικόνα 19: Κατάλληλη δομή δεδομένων για «block-wise» συμπλήρωση δεδομένων, δηλαδή με ολόκληρα τμήματα δεδομένων να λείπουν [59]

Για να αντιμετωπιστεί αυτό το πρόβλημα οι ερευνητές προτείνουν δύο διαφορετικές μεθόδους: Η πρώτη βασίζεται στην ατελή εκμάθηση με πολλές πηγές (incomplete multi-source feature learning, iMSF) και προσεγγίζει το πρόβλημα της πρόβλεψης τιμών ως ένα πρόβλημα εκμάθησης πολλών κατηγοριών (multi-task learning). Αρχικά διασπά τα δεδομένα σε κατηγορίες ανάλογα με το συνδυασμό πηγών δεδομένων που είναι διαθέσιμες για κάθε ασθενή. Θεωρώντας ότι όλοι οι ασθενείς έχουν κάνει εξετάσεις MRI, υπάρχουν άρα τα εξής μπλοκ πληροφορίας:

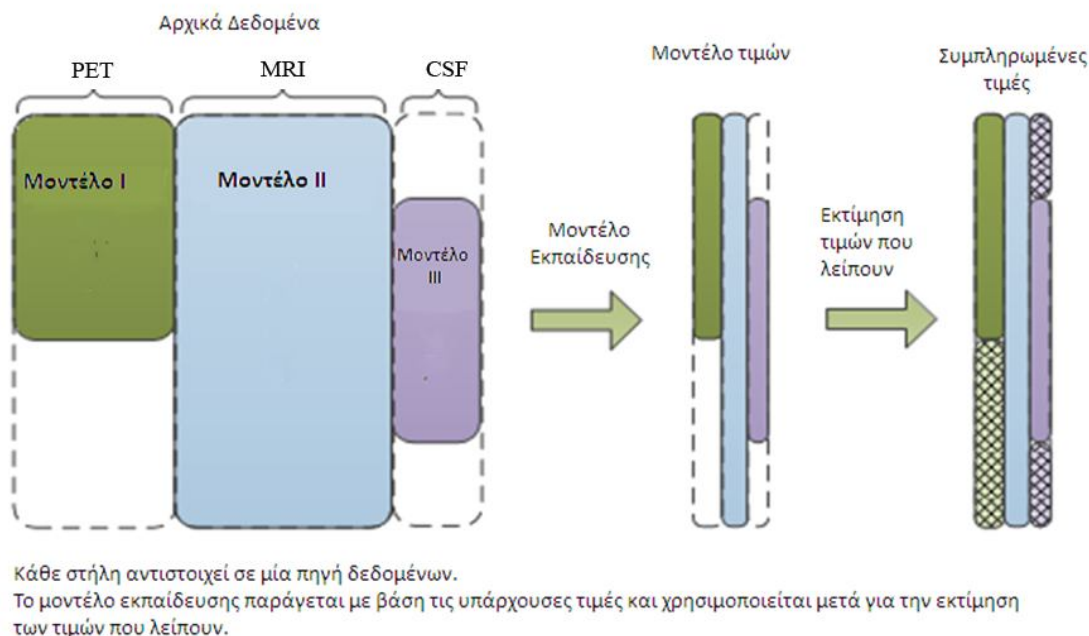
1. MRI, PET
2. MRI, PET, CSF
3. MRI, CSF
4. MRI



Εικόνα 20: Τα τέσσερα διαφορετικά μοντέλα που δημιουργήθηκαν για εκτίμηση τιμών από τον iMSF αλγόριθμο [59]

Για κάθε ένα από τα παραπάνω μπλοκ οι ερευνητές έφτιαξαν ένα μοντέλο πρόβλεψης τιμών με κοινή όμως εκμάθηση των διαφόρων χαρακτηριστικών, ώστε να αξιοποιείται η φύση των δεδομένων (πολλές πηγές). Ουσιαστικά πρόκειται για ένα γραμμικό μοντέλο με κανονικοποίηση, όπου δημιουργούνται έστω m μοντέλα πρόβλεψης και εισάγονται N_i αντικείμενα σε κάθε μοντέλο με $i=1, \dots, m$, $j=1, \dots, N_i$. Το j -οστό αντικείμενο του i -οστού μοντέλου ορίζεται ως (x_j^i, y_j^i) και για κάθε μοντέλο έχουμε $f^i(x_j) = (\beta^i)^T x_j^i$, όπου β^i είναι ο πίνακας βαρών για το i -μοντέλο και περιλαμβάνει τις παραμέτρους για κάθε πηγή δεδομένων.

Η δεύτερη μέθοδος εισάγει έναν εντελώς διαφορετικό αλγόριθμο συμπλήρωσης δεδομένων, τη συμπλήρωση επίδοσης (score completion, ScoreComp). Οι ερευνητές πρώτα εκπαιδεύουν ένα μοντέλο για κάθε πηγή δεδομένων με βάση τα διαθέσιμα δείγματα και παράγουν προβλέψεις τιμών για κάθε πηγή. Έτσι σχηματίζεται ένας ενιαίος πίνακας προβλέψεων με κενά στα μπλοκ που λείπουν. Στη συνέχεια κάθε γραμμή αυτού του πίνακα χρησιμοποιείται για να εκπαιδεύσει έναν τελευταίο ταξινομητή.



Εικόνα 21: Δημιουργία μοντέλου συμπλήρωσης επίδοσης στα ελλιπή δεδομένα [59]

Όπως προηγουμένως, και αυτός ο αλγόριθμος χρησιμοποιεί το σύνολο της πληροφορίας για να εξάγει την προσέγγιση. Ωστόσο ο αλγόριθμος αποφεύγει τη συμπλήρωση ολόκληρων μπλοκ πληροφορίας με το να συμπληρώνει τιμές σε κάθε σειρά του ενοποιημένου πίνακα. Παράλληλα είναι αρκετά απλός και όχι ιδιαίτερα απαιτητικός υπολογιστικά.

Η εκτίμηση των τιμών που λείπουν δοκιμάζεται με τρεις διαφορετικούς τρόπους:

- τον αλγόριθμο μηδέν (zero) που αρχικά βάζει μηδέν στα δεδομένα που λείπουν και αφού κανονικοποιήσει τα δεδομένα, εισάγει ως τιμή τον μέσο όρο της απόδοσης (imputation).
- τον αλγόριθμο μεγιστοποίησης προσδοκίας (expectation-maximization) που υπολογίζει τη μέση τιμή και την απόκλιση με βάση τα διαθέσιμα δεδομένα, συμπληρώνει τον πίνακα και μετά καλεί πάλι τον εαυτό του για τις νέες τιμές
- τον αλγόριθμο των k-πιο κοντινών γειτόνων (k-nearest neighbor, KNN) που ουσιαστικά για κάθε στοιχείο που λείπει βρίσκει το πιο όμοιο στοιχείο του πίνακα από τις τιμές που υπάρχουν και το εισάγει σαν τιμή

Οι παραπάνω αλγόριθμοι συγκρίθηκαν με βάση την απόδοσή τους στη βάση της ADNI. Οι ερευνητές έθεσαν τρία διαφορετικά προβλήματα, το διαχωρισμό ασθενών NA από την ομάδα ελέγχου,

ασθενών με MCI που ανέπτυξαν στη συνέχεια NA από την ομάδα ελέγχου και ασθενών NA από ασθενείς με MCI που έμειναν στάσιμοι . Τα παραπάνω προβλήματα είναι δηλαδή τα συνηθισμένα σε τέτοιου τύπου έρευνες, με το τελευταίο να είναι το δυσκολότερο από τα τρία. Σύγκριση εξελισσόμενου MCI με στάσιμο MCI δεν έγινε στο πλαίσιο της συγκεκριμένης έρευνας.

Σε σύγκριση οι ερευνητές δοκίμασαν και άλλους αλγόριθμους συμπλήρωσης δεδομένων, καθώς και έναν ταξινομητή που χρησιμοποιούσε μόνο τα MRI στοιχεία, ενώ αντιπαρέβαλαν τα αποτελέσματα ερευνών που δεν χρησιμοποίησαν κάποια συμπλήρωση δεδομένων με τα δικά τους. Τα αποτελέσματα είναι ομολογουμένως εντυπωσιακά.

Από τους αλγόριθμους συμπλήρωσης δεδομένων, οι δύο μέθοδοι που προτείνουν οι ερευνητές αποδίδουν καλύτερα από τους υπόλοιπους ταξινομητές. Η πρώτη μέθοδος, που χρησιμοποιεί την iMSF πλατφόρμα, παράγει τα σταθερότερα αποτελέσματα για τα δύο από τα τρία προβλήματα, δηλαδή το διαχωρισμό των ασθενών με MCI από τους υγιείς και ασθενών με NA από τους ασθενείς με στάσιμο MCI. Η δεύτερη μέθοδος, που χρησιμοποιεί τον αλγόριθμο ScoreComp σε συνδυασμό είτε με τον KNN είτε με τον αλγόριθμο μεγιστοποίησης προσδοκίας, διαχωρίζει καλά τους ασθενείς με NA από τους υγιείς, αλλά δυσκολεύεται στα άλλα δύο προβλήματα. Σε σύγκριση με αυτές τις μεθόδους, οι απλοί αλγόριθμοι συμπλήρωσης δεδομένων έχουν αρκετά χαμηλότερες επιδόσεις, επιβεβαιώνοντας ότι η συμπλήρωση ολόκληρων μπλοκ πληροφορίας απαιτεί πιο εξειδικευμένες λύσεις.

Ακολουθεί ο σχετικός πίνακας. Οι ερευνητές συνέκριναν τις επιδόσεις των αλγορίθμων για διαφορετικά μεγέθη της ομάδας εκπαίδευσης.

Σύγκριση των επιδόσεων ταξινόμησης του αλγορίθμου iMSF, του αλγορίθμου score completion και μεθόδων εκτίμησης τιμών που λείπουν. Το ποσοστό της ομάδας εκπαίδευσης σε σχέση με το δείγμα κινείται από 50% ως 75%

Training Size 50%	AD vs. Normal			AD vs. Non-converter			Converter vs. Normal		
	Accuracy	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity
iMSF-Best	86.37%	85.64%	87.65%	78.05%	57.34%	89.91%	91.20%	94.11%	89.59%
iMSF-Average	83.00%	81.56%	84.80%	77.51%	55.97%	88.64%	90.34%	93.31%	84.18%
ScoreComp-Zero	87.31%	87.61%	87.35%	81.25%	53.26%	95.68%	83.15%	99.32%	48.12%
ScoreComp-EM	89.74%	94.13%	85.32%	81.05%	53.35%	95.31%	88.99%	98.73%	68.30%
ScoreComp-KNN	90.13%	94.87%	85.37%	81.46%	56.02%	94.58%	87.60%	98.89%	63.49%
Zero	85.38%	87.74%	83.23%	77.25%	44.10%	94.14%	89.48%	97.93%	71.36%
EM	87.20%	88.76%	85.75%	78.11%	47.25%	93.80%	89.40%	97.97%	70.95%
KNN	85.79%	88.10%	83.66%	75.15%	36.78%	94.59%	86.40%	97.88%	61.41%
SVD	83.89%	84.94%	82.96%	75.73%	43.00%	92.37%	87.57%	97.09%	67.07%
SVT	55.75%	61.31%	50.70%	81.15%	56.82%	93.48%	71.69%	87.81%	36.34%
Training Size 66.7%	Accuracy	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity
iMSF-Best	88.77%	86.77%	91.25%	80.55%	65.83%	89.03%	93.17%	95.36%	90.52%
iMSF-Average	86.07%	83.23%	89.23%	80.12%	63.70%	88.16%	92.40%	94.72%	87.37%
ScoreComp-Zero	88.18%	86.48%	90.10%	83.48%	59.90%	94.96%	84.48%	99.53%	50.13%
ScoreComp-EM	90.28%	93.64%	86.61%	83.25%	60.85%	94.12%	90.82%	99.16%	72.07%
ScoreComp-KNN	90.19%	94.48%	85.51%	83.30%	62.25%	93.56%	89.56%	99.24%	67.66%
Zero	86.79%	88.80%	84.61%	80.87%	52.21%	94.56%	92.30%	98.84%	77.32%
EM	88.46%	88.40%	88.52%	80.90%	53.28%	93.96%	91.69%	98.76%	75.75%
KNN	87.20%	87.60%	86.75%	78.46%	44.24%	94.75%	89.07%	98.68%	66.84%
SVD	85.25%	85.19%	85.36%	78.52%	48.89%	92.51%	90.44%	98.13%	72.75%
SVT	57.14%	61.08%	53.39%	83.25%	60.71%	94.01%	75.30%	91.66%	38.03%
Training Size 75%	Accuracy	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity
iMSF-Best	88.79%	87.16%	90.92%	81.33%	66.02%	89.06%	93.19%	95.94%	90.58%
iMSF-Average	85.81%	82.77%	89.38%	80.78%	65.20%	88.70%	92.98%	95.03%	88.87%
ScoreComp-Zero	88.53%	87.67%	89.75%	84.51%	61.48%	96.00%	84.96%	99.60%	52.74%
ScoreComp-EM	89.91%	92.97%	86.77%	83.80%	61.24%	94.95%	89.93%	99.30%	70.21%
ScoreComp-KNN	90.35%	93.89%	86.74%	84.35%	63.79%	94.60%	88.89%	99.11%	66.91%
Zero	86.49%	88.06%	85.13%	81.76%	55.54%	94.74%	91.48%	97.68%	78.49%
EM	88.96%	88.54%	89.54%	81.25%	55.72%	93.83%	90.74%	98.44%	74.38%
KNN	87.32%	88.20%	86.66%	77.73%	44.97%	93.66%	88.89%	98.43%	68.12%
SVD	85.93%	85.78%	86.23%	78.55%	50.13%	92.29%	89.19%	97.48%	70.92%
SVT	56.58%	59.38%	53.73%	81.92%	56.93%	94.17%	75.56%	90.23%	42.59%

Πίνακας 14: Σύγκριση των επιδόσεων των διαφόρων μεθόδων συμπλήρωσης δεδομένων [59]

Στη σύγκριση με τον ταξινομητή που χρησιμοποιούσε μόνο τα MRI δεδομένα, οι δύο αλγόριθμοι ήταν πολύ καλύτεροι, επιβεβαιώνοντας ότι η χρήση πολλών πηγών δεδομένων βελτιώνει την απόδοση του συστήματος.

Στη συνέχεια οι ερευνητές διαχώρισαν τα πλήρη δεδομένα της ADNI από τα υπόλοιπα και συνέκριναν την ταξινόμηση που προκύπτει αν απέρριπταν τα ελλιπή δεδομένα, με την ταξινόμηση για τα πλήρη δεδομένα από το δικό τους σύστημα. Η βελτίωση είναι πολύ εντυπωσιακή.

Επιδόσεις απλής πολυτροπικής ταξινόμησης και ταξινόμησης με αλγόριθμο iMSF και αλγόριθμο score completion.

	AD vs. Normal			AD vs. Non-converter			Converter vs. Normal		
	Accuracy	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity
Baseline	87.72%	100.00%	63.16%	76.19%	65.79%	84.78%	83.02%	97.06%	57.89%
iMSF-Average	95.79%	97.37%	92.63%	84.05%	79.47%	87.83%	88.68%	90.59%	85.26%
ScoreComp-EM	98.25%	97.37%	100.00%	83.33%	65.79%	97.83%	84.91%	97.06%	63.16%

Πίνακας 15: Σύγκριση των επιδόσεων πολυτροπικών συστήματος με τον «score-completion» τρόπο συμπλήρωσης δεδομένων και τον iMSF [59]

Βλέπουμε ότι το σύστημα με τον score completion αλγόριθμο και τον αλγόριθμο μεγιστοποίησης της προσδοκίας στη διάκριση NA από υγιείς πέτυχε ακρίβεια ταξινόμησης 98.2%. Αυτό αντιστοιχεί σε βελτίωση πάνω από 5% σε σχέση με τους προηγούμενους αλγορίθμους. Ακόμη μεγαλύτερη είναι η βελτίωση στην ειδικότητα και στην ευαισθησία με ποσοστά 100% και 97.4% αντίστοιχα. Η διάκριση NA από στάσιμο MCI έχει ακρίβεια 83.3% με τον ίδιο αλγόριθμο και λίγο μεγαλύτερη με τον iMSF (84%). Τέλος, στη διάκριση εξελισσόμενου MCI από υγιείς ο iMSF έχει την καλύτερη ακρίβεια με επίδοση 88.6%. Σε αυτό το πρόβλημα, η πλατφόρμα MLK που είδαμε στην προηγούμενη παράγραφο [58] πέτυχε λίγο υψηλότερη ακρίβεια (91%), ωστόσο, δεδομένου ότι η συμπλήρωση δεδομένων μπορεί να γίνει σε πρώτο στάδιο και στη συνέχεια να εφαρμοστεί οποιοσδήποτε αλγόριθμος ταξινόμησης, υπάρχει δυνατότητα βελτίωσης των αποτελεσμάτων.

Οι ερευνητές καταλήγουν ότι η συμπλήρωση των ελλιπών δεδομένων βελτιώνει θεαματικά την ταξινόμηση, δηλαδή σε ποσοστό που αγγίζει το +5% στην ακρίβεια, ενώ παράλληλα είναι σε θέση να προβλέψει τιμές αν για παράδειγμα στο μέλλον εισαχθούν νέα δεδομένα, κάτι που μπορεί να είναι χρήσιμο σε κλινικό επίπεδο. Η παραπάνω έρευνα αντιμετωπίζει ένα σημαντικό ερευνητικό πρόβλημα στην έρευνα της NA. Όντως, σε πολλές μελέτες έχουμε δει ότι η ελλιπής πληροφορία απορρίπτεται και το μέγεθος του δείγματος μειώνεται σημαντικά με τις στατιστικές δυσκολίες που αυτό συνεπάγεται. Για παράδειγμα, στην περίπτωση της ADNI από 648 δείγματα μένουν μόλις 153 που έχουν διαθέσιμες όλες τις εξετάσεις. Μελλοντικές έρευνες ίσως θα πρέπει να εφαρμόσουν τέτοιες προσεγγίσεις, που επιτρέπουν την αξιοποίηση του συνόλου της πληροφορίας.

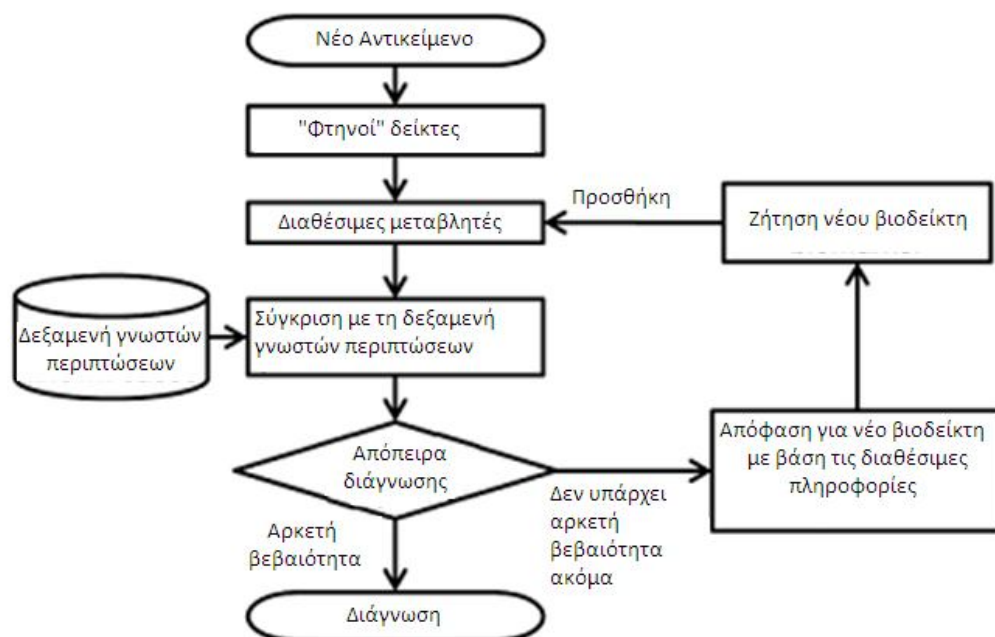
4.2 Εφαρμογή στη διαγνωστική διαδικασία

4.2.1 Εξατομίκευση και κόστος στα συστήματα υποστήριξης κλινικών αποφάσεων

Η επόμενη έρευνα που θα εξετάσουμε ασχολείται περισσότερο με την ένταξη ενός πολυτροπικού βιοδείκτη στη διαγνωστική διαδικασία [60]. Η διαφοροποίηση αυτής της έρευνας από τις προηγούμενες έγκειται στο ότι η πρακτική εφαρμογή ενός βιοδείκτη περιορίζεται από τις ειδικές ανάγκες κάθε ασθενούς και τις διαθέσιμες εξετάσεις. Συνεπώς οι ερευνητές προτείνουν ένα σύστημα που συμβουλεύει τον ιατρό για το ποιος βιοδείκτης θα προσφέρει την περισσότερη πληροφορία στο επόμενο στάδιο και αποτιμά την πιθανότητα να έχει ο ασθενής ΝΑ. Αντίθετα, τα περισσότερα συστήματα που είδαμε μέχρι τώρα θεωρούν διαθέσιμες όλες τις εξετάσεις από την αρχή.

Η ερευνητική ομάδα αποφάσισε να χρησιμοποιήσει τη βάση δεδομένων της ADNI για να εκπαιδεύσει τον ταξινομητή. Αρχικά επέλεξαν τα δείγματα για τα οποία όλες οι εξετάσεις και τα επιδημιολογικά δεδομένα ήταν διαθέσιμα. Σε κάθε ένα από τα παραπάνω στοιχεία εισήγαγαν ένα εκτιμώμενο κόστος. Για παράδειγμα πληροφορίες για φύλο, ηλικία και χρόνια εκπαίδευσης θεωρήθηκαν ότι έχουν μηδενικό κόστος. Μία εξέταση MRI αντίθετα θεωρήθηκε ότι κοστίζει 163 δολάρια ενώ μία εξέταση PET 844 δολάρια.

Όταν εισάγεται ένας νέος ασθενής, ο αλγόριθμος συγκρίνει τα διαθέσιμα στοιχεία (για παράδειγμα επιδημιολογικά δεδομένα ή εξετάσεις που έχει ήδη κάνει) με αυτά μιας βάσης ασθενών, για τους οποίους έχει ήδη ληφθεί κάποια διαγνωστική απόφαση, προκειμένου να βρεθούν οι πλέον όμοιες περιπτώσεις και να εισαχθούν τα αντίστοιχα βάρη. Η προσέγγιση αυτή είναι η κλασική προσέγγιση της τοπικά επιβαρυσμένης εκπαίδευσης (locally weighted training) και επιτρέπει την εξατομίκευση του ταξινομητή στον κάθε ασθενή. Στη συνέχεια το σύστημα προσπαθεί να βγάλει κάποια διάγνωση. Αν μπορεί να καταλήξει σε διάγνωση με αρκετή βεβαιότητα, η διαδικασία λήγει. Αλλιώς το σύστημα επιλέγει ποιος θα είναι ο επόμενος δείκτης. Οι ερευνητές δοκίμασαν δύο διαφορετικά συστήματα απόφασης για την επιλογή του επόμενου δείκτη: το πρώτο σύστημα επέλεγε το δείκτη που θα προσφέρει την περισσότερη πληροφορία (δηλαδή τους δείκτες που προσέφεραν την περισσότερη πληροφορία στις όμοιες περιπτώσεις που εξέτασε το σύστημα στο παρελθόν), ενώ το δεύτερο σύστημα αποσκοπούσε στην ελαχιστοποίηση του συνολικού κόστους των εξετάσεων.



Εικόνα 22: Εξατομίκευση της επιλογής βιοδεικτών για ημιαυτοματοποιημένη διάγνωση της ΝΑ με k -πιο κοντινούς γείτονες και συνυπολογισμό του κόστους και της αποτελεσματικότητας [60]

Η προσέγγιση της τοπικά επιβαρυνμένης εκπαίδευσης επιλέγεται έναντι άλλων προσεγγίσεων όπως οι k -πιο κοντινοί γείτονες, γιατί είναι λιγότερο ευαίσθητη στο θόρυβο, ενώ επιτρέπει την εξατομίκευση του ταξινομητή, καθώς οι υπολογισμοί γίνονται όταν εισάγεται το δεδομένο.

Η προσαρμογή των δεδομένων στον ασθενή οφείλεται στη δεξαμενή πληροφοριών για προηγούμενους ασθενείς. Για να λειτουργήσει αποτελεσματικά το σύστημα είναι απαραίτητη μία αρκετά μεγάλη βάση. Αφού βρεθούν οι παρεμφερείς περιπτώσεις μοιράζονται βάρη για την επιλογή των επιθυμητών χαρακτηριστικών. Αν ζητάμε μέγιστη αποτελεσματικότητα, θα επιλεχθούν χαρακτηριστικά που σε ανάλογες περιπτώσεις έδωσαν τη μέγιστη πληροφορία (δηλαδή ελαχιστοποίηση του αριθμού των δεικτών). Αν λάβουμε υπόψη και το κόστος, θα επιλεχθούν τα πλέον αποτελεσματικά χαρακτηριστικά ανά μονάδα κόστους.

Τα προβλήματα που καλείται να λύσει ο ταξινομητής είναι δύο, τη σχετικά εύκολη διαφοροποίηση ανάμεσα στους υγιείς και στους ασθενείς με ΝΑ και την πολύ δυσκολότερη ανάμεσα στους ασθενείς με ΜCI που παρέμειναν στάσιμοι και σε αυτούς με ΜCI που εντός 12 μηνών εμφάνισαν ΝΑ. Τα αποτελέσματα συγκρίθηκαν με κλασικούς ταξινομητές που έχουν διαθέσιμα όλα τα δεδομένα.

Τα αποτελέσματα είναι ενθαρρυντικά. Όταν ζητείται απλώς η ελαχιστοποίηση του αριθμού των δεικτών μέχρι να υπάρξει ικανοποιητική διάγνωση (δηλαδή επιλογή δεικτών με κριτήριο τη μεγιστοποίηση της πληροφορίας), η απόδοση είναι λίγο χαμηλότερη από τους κλασικούς ταξινομητές (0,88 έναντι 0,9 για διαχωρισμό NA από ομάδα ελέγχου), ενώ χρησιμοποιούνται λιγότεροι από τους μισούς δείκτες κατά μέσο όρο. Αντίστοιχα αποτελέσματα λαμβάνονται και για το διαχωρισμό στάσιμου MCI από εξελισσόμενο MCI (0,63 έναντι 0,66). Το σύστημα απόφασης με στόχο την ελαχιστοποίηση του κόστους είχε πολύ χειρότερα αποτελέσματα (μόλις 0,75 ακρίβεια για διαχωρισμό NA από ομάδα ελέγχου), συνεπώς απαιτείται κάποια επαναπροσέγγιση αυτού του αλγορίθμου (οι ερευνητές αναφέρουν ότι θα επανεξετάσουν τα κριτήρια με τα οποία λαμβάνει απόφαση σε επόμενη μελέτη). Ο αλγόριθμος ελαχιστοποίησης του αριθμού των δεικτών είναι, όμως, μια μάλλον αξιόλογη προσπάθεια, ιδιαίτερα γιατί το σύστημα έχει μεγάλες δυνατότητες βελτίωσης (οι ερευνητές αξιοποίησαν έναν μόνο δείκτη για τα MRI και τα PET, ενώ είναι σαφές ότι υπάρχουν πιο προηγμένες λύσεις διαθέσιμες). Παράλληλα, αν η βάση δεδομένων χρησιμοποιεί τοπικά στοιχεία, δηλαδή ασθενείς από ένα μόνο νοσοκομείο, μπορεί να βελτιστοποιηθεί ο αλγόριθμος όσον αφορά τα επιδημιολογικά χαρακτηριστικά της NA (πχ χρόνια εκπαίδευσης, κα).

Συνολικά η έρευνα αποτελεί μια απόπειρα πρακτικής ένταξης ενός πολυτροπικού βιοδείκτη στην κλινική διαδικασία. Τα αποτελέσματα είναι αρκετά καλά, δεδομένου ότι μειώνεται η ταλαιπωρία του ασθενούς (λιγότερες εξετάσεις) και το κόστος. Ταυτόχρονα τίθεται και ένα ενδιαφέρον ερώτημα: είναι σαφές ότι περισσότεροι δείκτες βελτιώνουν τη διάγνωση, αλλά από ποιο σημείο και μετά οι επιπλέον δείκτες αποφέρουν μικρή βελτίωση; Και πώς αυτή η πληροφορία θα αξιοποιηθεί σε ένα διαγνωστικό εργαλείο για τη NA;

4.2.2 Συνδυασμός βιοδεικτών: Αλληλοεπικάλυψη πληροφορίας;

Μία απάντηση σε αυτό το ερώτημα έδωσαν οι ερευνητές στην ακόλουθη εργασία [61]. Εδώ ο στόχος ήταν να βρεθεί ο βέλτιστος (και ελάχιστος) συνδυασμός διαφόρων δεικτών (που περιλαμβάνουν και νευροψυχολογικές δοκιμασίες σε αντίθεση με προηγούμενες εργασίες) στην πρόγνωση της NA μέσω του ελέγχου της επανάληψης της πληροφορίας σε διαφορετικούς δείκτες (για παράδειγμα κάποια χαρακτηριστικά MRI μπορεί να μας δίνουν παρεμφερή πληροφορία). Η αποτελεσματικότητα των συνδυασμών αποτιμάται από το πόσο καλά διαχώρισαν τις διάφορες ομάδες του δείγματος, που περιλαμβάνει ασθενείς με NA, ασθενείς με MCI με σταθερή πρόγνωση, ασθενείς με εξελισσόμενο MCI και υγιείς ως ομάδα ελέγχου.

Οι ερευνητές αντιμετώπισαν δύο κατεξοχήν προβλήματα, που είναι κοινά στις πολυτροπικές μελέτες: ο βέλτιστος συνδυασμός μπορεί να είναι διαφορετικός για κάθε πληθυσμό και άρα προκύπτει η ανάγκη για προσαρμογή των βιοδεικτών στον εκάστοτε ασθενή. Παράλληλα, κάθε βιοδείκτης παίζει διαφορετικό ρόλο σε κάθε στάδιο της ασθένειας, συνεπώς διαφορετικοί συνδυασμοί είναι αποτελεσματικότεροι ανάλογα με τη σοβαρότητα των συμπτωμάτων.

Στην έρευνα χρησιμοποιήθηκαν δύο διαφορετικές βάσεις δεδομένων, αυτή της ADNI και η Kuorio L-MCI βάση δεδομένων. Η τελευταία είναι μια βάση ενός πανεπιστημίου της Φινλανδίας, που, όμοια με την ADNI, είχε πενταετή διάρκεια, προκειμένου να καταγραφεί η πορεία των ασθενών. Αυτή η βάση δεν είναι ακόμη πλήρης (για παράδειγμα δεν είναι διαθέσιμα όλα τα στοιχεία MRI, που είναι ένα πιθανό πρόβλημα για την έρευνα). Ο αριθμός του διαθέσιμου δείγματος ήταν 821 από την ADNI και 977 από την Kuorio L-MCI. Το παραπάνω δείγμα και οι διαθέσιμοι δείκτες ενοποιήθηκαν σε μια ενιαία βάση.

Η επιλογή των δεικτών έγινε με 3 τρόπους. Την επιλογή-προς-τα-εμπρός, (forward selection) που ξεκινά με τον πιο χρήσιμο δείκτη και προσθέτει ανάλογα και άλλους, την επιλογή-προς-τα-πίσω (backward selection), που ξεκινά με όλα τα χαρακτηριστικά και αφαιρεί αυτά που επηρεάζουν την ταξινόμηση το λιγότερο και την επιλογή-με-βήμα (stepwise selection), που σε κάθε στάδιο αφαιρεί τα ν-λιγότερο αποτελεσματικά και προσθέτει τα μ-υποψήφια καλύτερα. Η τελευταία επιλογή είναι και η αποτελεσματικότερη γιατί αποφεύγει το πρόβλημα του φωλιάσματος, τον εγκλωβισμό δηλαδή σε κάποια μη βέλτιστη “κοιλιάδα” δεδομένων.

Για την αξιολόγηση των χαρακτηριστικών, πέρα από το κλασικό σφάλμα ταξινόμησης, χρησιμοποιήθηκε και το κριτήριο ελάχιστης επανάληψης μέγιστης συνάφειας (minimum redundancy maximum relevance - χρησιμοποιείται για την επιλογή χαρακτηριστικών), προκειμένου να μην επαναλαμβάνεται πληροφορία από διαφορετικούς δείκτες. Η αξιολόγηση των διαφόρων συνδυασμών έγινε με την εφαρμογή ενός SVM ταξινομητή.

Το πόρισμα της έρευνας είναι ότι μόλις το 30% των χαρακτηριστικών που δίνονται από την ADNI και την L-MCI αλλάζει σημαντικά για τις διάφορες ομάδες του δείγματος και άρα συνιστούν τον βέλτιστο συνδυασμό δείκτη για τα διάφορα στάδια της ασθένειας. Οι διάφορες νευροψυχολογικές δοκιμασίες είχαν και το σημαντικότερο ρόλο στο διαχωρισμό των ομάδων, όπως βλέπουμε στον πίνακα 16, κάτι που προφανώς γνωρίζουμε και από τα κλινικά πρότυπα για την άνοια. Ωστόσο είχαν μειωμένο ρόλο στον πιο ενδιαφέροντα διαχωρισμό, αυτόν ανάμεσα σε εξελισσόμενο και στάσιμο MCI. Το ίδιο ισχύει και για άλλους δείκτες όπως ηλικία και επίπεδο μόρφωσης. Αντίθετα οι δείκτες του MRI και του APOE διαφοροποιούνταν αισθητά για όλες τις ομάδες. Τέλος, συγκρίνεται και ο ρόλος του βέλτιστου

δείκτη σε σχέση με τον βέλτιστο συνδυασμό, όπου επιβεβαιώνεται ότι τα πολυτροπικά συστήματα βελτιώνουν πολύ την απόδοση της ταξινόμησης.

Ακρίβεια ταξινόμησης για διαφορετικά χαρακτηριστικά και το συνδυασμό τους

Classification task	Features from one single category	Combined features	Most important feature category
AD vs. P-MCI	84 %	94 %	CDR
AD vs. S-MCI	93 %	99 %	MMSE
AD vs. Control	100 %	100 %	CDR, NB
P-MCI vs. S-MCI	81 %	87 %	NB
P-MCI vs. Control	100 %	100 %	CDR
S-MCI vs Control	100 %	100 %	CDR

Πίνακας 16: Επιδόσεις στην ταξινόμηση για τις διάφορες ομάδες [61]. Τα πιο σημαντικά χαρακτηριστικά είναι οι νευροψυχολογικές δοκιμασίες

Η παραπάνω έρευνα εξερευνά μια ενδιαφέρουσα πτυχή της ανάπτυξης δεικτών. Το σύστημα αυτό όμως δεν είναι ανεξάρτητο από ανθρώπινη παρέμβαση, καθώς περιλαμβάνει το δείκτη κλινικής εκτίμησης της άνοιας (Clinical Dementia Rating, CDR), στον οποίο συμμετέχει ο ιατρός που παίρνει τη συνέντευξη. Η ακρίβεια που πετυχαίνουν είναι εντυπωσιακή (100% ακρίβεια στην ταξινόμηση ΝΑ και υγιών), αλλά ουσιαστικά βεβαιώνει ότι το CDR είναι αποτελεσματικό μέσο διάγνωσης ΝΑ. Η έρευνα δεν μπορεί να συγκριθεί αποτελεσματικά με άλλες μελέτες, καθώς οι περισσότερες έρευνες που εξετάσαμε αποσκοπούσαν σε ένα ανεξάρτητο από ανθρώπινη παρέμβαση σύστημα απόφασης και γι' αυτό δεν εισήγαγαν το CDR στους δείκτες που χρησιμοποίησαν. Ωστόσο οι ερευνητές σχεδιάζουν νέα μελέτη που δεν θα περιλαμβάνει το CDR. Πάντως η συμβολή της έρευνας στην αλληλεξάρτηση των διαφόρων δεικτών είναι σημαντική.

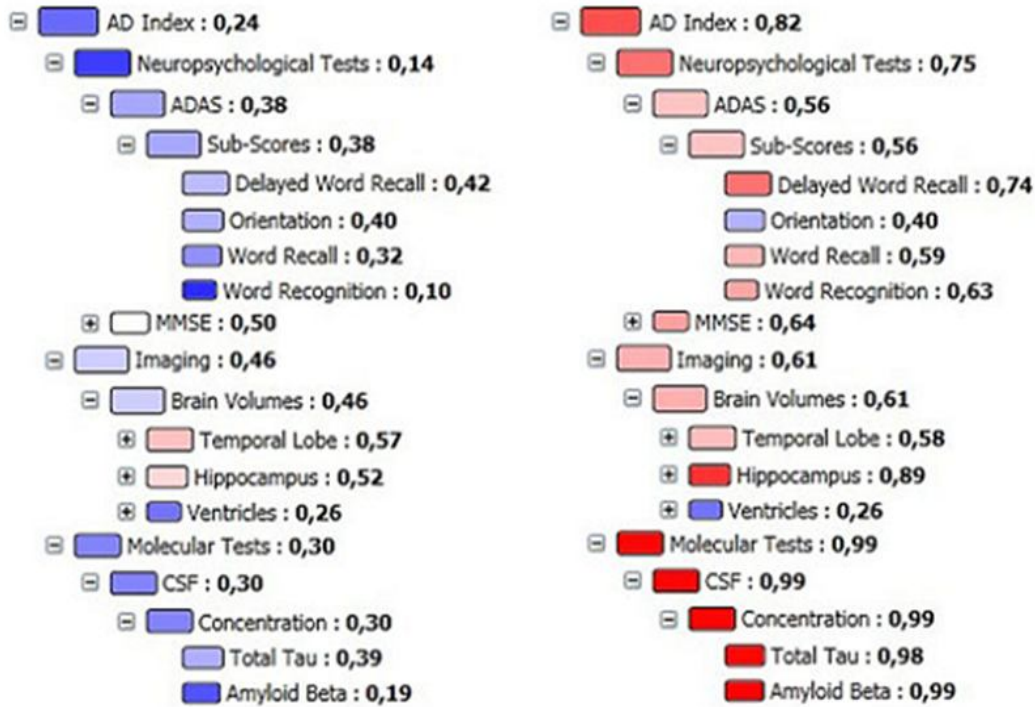
4.2.3 Η εισαγωγή ενός αγνωστικιστικού συστήματος ως προδρόμου ενός κλινικού εργαλείου

Η ίδια ομάδα που σχεδίασε το παραπάνω σύστημα, δοκίμασε στη συνέχεια να γενικεύσει την προσπάθειά της [62]. Οι ερευνητές διαπίστωσαν ότι τα περισσότερα CDSSs είναι εξαιρετικά εξειδικευμένα, καθώς περιορίζονται στη διάγνωση κάποιας συγκεκριμένης ασθένειας, ενώ ο τρόπος που τα δεδομένα μαζεύονται από διαφορετικές κλινικές διαφοροποιείται πολύ. Παράλληλα, ο όγκος της

απαιτούμενης πληροφορίας για κάθε ασθένεια μεγεθύνεται εξαιρετικά γρήγορα, δυσχεραίνοντας περαιτέρω την εφαρμογή τέτοιων συστημάτων στο κλινικό περιβάλλον.

Η ομάδα προσπάθησε να δημιουργήσει ένα αγνωστικιστικό ως προς τα δεδομένα που διαχειρίζεται σύστημα απόφασης, δηλαδή ένα σύστημα που επιτρέπει την εισαγωγή δεδομένων κάθε τύπου, τη στατιστική τους ανάλυση και την παρουσίασή τους στον ιατρό. Στη συνέχεια αξιολόγησαν την προσπάθειά τους με την επίδοσή της στη βάση της ADNI.

Ουσιαστικά η εφαρμογή προσπαθεί να ενοποιήσει τα δεδομένα και να εξάγει κάποια τιμή στην κλίμακα του $[0,1]$, που θα δείχνει την κατάσταση του αντικειμένου όσον αφορά κάποια ασθένεια, με βάση προηγούμενες περιπτώσεις. Εξυπακούεται ότι η εφαρμογή περιορίζεται στην ανάλυση ποσοτικών δεδομένων, δηλαδή ο χρήστης έχει ήδη εξάγει για παράδειγμα τιμές μορφομετρίας από το MRI. Ωστόσο, σχεδόν οποιαδήποτε τιμή υποστηρίζεται είτε αυτή είναι κατηγορική είτε βαθμωτή. Για κάθε τιμή εφαρμόζεται μία συνάρτηση προσαρμογής που προσπαθεί να καθορίσει σε τι σημείο βρίσκεται ο ασθενής σε σχέση με τον υπόλοιπο πληθυσμό. Πρακτικά το σύστημα μετατρέπει κάθε τιμή σε έναν κοινό τόπο ταξινόμησης και στη συνέχεια ενοποιεί το σύνολο των αποτελεσμάτων.



Απεικόνιση δένδρου για δύο περιπτώσεις του δείγματος, έναν υγιή (αριστερά) και έναν ασθενή με NA (δεξιά). Τα μεγαλύτερα μεγέθη κόμβων αντιπροσωπεύουν τη μεγαλύτερη σημασία για τη διάκριση ασθενών - υγιών. Τα πιο σκούρα χρώματα αντιπροσωπεύουν την ομοιότητα με άλλες περιπτώσεις του δείγματος (περιπτώσεις υγιών με γαλάζιο, περιπτώσεις ασθενών με κόκκινο).

Εικόνα 23: Περιπτώσεις ταξινόμησης του συστήματος [62] ανάμεσα σε ασθενή με NA και υγιή και τα αντίστοιχα κριτήρια

Η εισαγωγή των δεδομένων που χρησιμοποιούνται για την εκπαίδευση του συστήματος είναι συμβατή με τα πιο πολλά από τα σημερινά πρότυπα. Όταν αρχικοποιείται η βιβλιοθήκη, ο τύπος των δεδομένων και τα χαρακτηριστικά πρέπει να οριστούν. Στη συνέχεια πρέπει να εξαχθεί η συνάρτηση προσαρμογής. Για να επιτευχθεί αυτό η βιβλιοθήκη δημιουργεί ομάδες ελέγχου και ομάδες θετικές στην ασθένεια μέσω ενός γενικού συνόλου κανόνων. Ο χρήστης μπορεί να προσαρμόσει τους κανόνες περαιτέρω μέσω ενός γραφικού περιβάλλοντος. Για την εισαγωγή δεδομένων από διαφορετικές κλινικές χρησιμοποιείται μια βάση δεδομένων που ομαλοποιεί τα δεδομένα με βάση τον ορισμό που έχει δοθεί από το χρήστη.

Το σύστημα αξιολογήθηκε στη βάση δεδομένων της ADNI και η επίδοσή του συγκρίθηκε με κλασικούς ταξινομητές (SVM και λογιστική παλινδρόμηση). Το βασικό πρόβλημα που κλήθηκαν να αντιμετωπίσουν ήταν ο διαχωρισμός στάσιμων MCI και εξελισσόμενων MCI, που είναι και το πιο ενδιαφέρον θέμα γενικότερα. Η ακρίβεια της πρόβλεψης άγγιξε το 0.68 και ήταν ελαφρώς καλύτερη από την επίδοση των υπόλοιπων μεθόδων με τις οποίες συγκρίθηκε. Στο συγκεκριμένο πείραμα τα

πρόσφατα συστήματα πετυχαίνουν ακρίβεια περίπου 69%, δηλαδή το γενικευμένο CDSS πέτυχε επιδόσεις αντίστοιχες με των εξειδικευμένων συστημάτων στην πρόγνωση της ΝΑ. Εδώ να σημειώσουμε ότι οι μελετητές δοκίμασαν το σύστημά τους και σε τρεις άλλες ομάδες δεδομένων (για ηπατίτιδα, διαβήτη και καρδιακά νοσήματα) και είχαν εξίσου καλά αποτελέσματα. Βέβαια η έρευνα δεν εκτελεί τα πιο κλασικά πειράματα όπως διαχωρισμός ΝΑ από ομάδα ελέγχου, για τα οποία υπάρχουν πιο εκτεταμένα στατιστικά δεδομένα που θα επέτρεπαν την πιο έγκυρη αξιολόγησή της.

Επίδοση ταξινόμησης στην ομάδα MCI του δείγματος της ADNI

Method	AUC ^a	Accuracy	Sensitivity	Specificity
DSI	0.75 ± 0.08	0.68 ± 0.08	0.70 ± 0.12	0.66 ± 0.10
SVM	0.75 ± 0.08	0.67 ± 0.07	0.64 ± 0.11	0.69 ± 0.11
Bayes	0.76 ± 0.08	0.67 ± 0.07	0.65 ± 0.12	0.69 ± 0.11
LR	0.69 ± 0.09	0.62 ± 0.07	0.73 ± 0.10	0.53 ± 0.11

Η τιμή και η απόκλιση προέκυψαν έπειτα από 10 επαναλήψεις του πειράματος

Πίνακας 17: Επιδόσεις ταξινόμησης σε εξελισσόμενο MCI και στάσιμο MCI με τους διαφορετικούς αλγορίθμους [62]

Συνολικά, το σύστημα υλοποιεί έναν ακόμη κρίκο στην εισαγωγή τέτοιων συστημάτων στην κλινική διαδικασία. Απαιτεί πολύ μικρότερη αρχικοποίηση από το χρήστη, επιτρέπει τη σποραδική εισαγωγή δεδομένων (που είναι πιο κοντά στην κλινική πραγματικότητα) και δεν είναι ιδιαίτερα δύσκολο υπολογιστικά. Η ομαλοποίηση των διαφορετικών βάσεων δεδομένων και η κοινή διαχείρισή τους είναι επίσης ένα σημαντικό εργαλείο. Η εισαγωγή ενός τέτοιου γενικού συστήματος στην κλινική διαδικασία θα επέτρεπε την τυποποίηση των ερευνών στη ΝΑ, ενώ η εφαρμογή του σε πολλούς τύπους ασθενειών θα ενθάρρυνε την περαιτέρω βελτιστοποίησή του. Με την άνοδο των ηλεκτρονικών φακέλων ασθενείας ένα τέτοιο πρότυπο θα μπορούσε να λειτουργεί πλήρως αυτοματοποιημένα και να συλλέγει και να ερμηνεύει δεδομένα σε εθνικό ή παγκόσμιο επίπεδο. Προς το παρόν πάντως το σύστημα απαιτεί αρκετά μεγάλη παρέμβαση από το χρήστη, ενώ η κάθε βιβλιοθήκη χειρίζεται μόνο μία ασθένεια. Η επέκταση στην ταυτόχρονη διαχείριση πολλών ασθενειών θα υλοποιηθεί στο μέλλον.

Πάντως, η δημιουργία ενός διαγνωστικού εργαλείου για τη ΝΑ με πιο γενικές δυνατότητες σίγουρα επιτρέπει την ευκολότερη εισαγωγή του στην κλινική διαδικασία. Τέλος, έχει καλή απόδοση σε σχέση με πιο εξειδικευμένα CDSSs, κάτι από μόνο του ενθαρρυντικό.

4.3 Συμπεράσματα για τα πολυτροπικά συστήματα

Συμπερασματικά, η βελτίωση που προσφέρει η πολυτροπική ταξινόμηση είναι καλά τεκμηριωμένη. Με τη συνεχή εξέλιξη στατιστικών τεχνικών πρόβλεψης, η απόδοση αντίστοιχων συστημάτων θα βελτιωθεί.

Ωστόσο, υπάρχει πιθανότητα και κάποιο ανώτατο όριο στη δυνατή βελτίωση. Η πληροφορία που προσφέρουν επιπλέον δείκτες δεν βελτιώνει απαραίτητα την ταξινόμηση γιατί είναι εξαρτώμενοι από άλλους. Είδαμε για παράδειγμα ένα πολυτροπικό σύστημα όπου αριθμός μεγαλύτερος των δέκα πηγών δεδομένων δεν βελτιώνει τα αποτελέσματα ή ότι ο συνδυασμός τριών μόλις πηγών προσεγγίζει αρκετά την απόδοση συνδυασμών με πολύ περισσότερους δείκτες.

Η εισαγωγή περιβαλλοντικών χαρακτηριστικών σε αυτούς τους “παγκόσμιους” βιοδείκτες ίσως επιτρέψει την εισαγωγή νέας πληροφορίας και φέρει στο προσκήνιο την αλληλεπίδραση περιβαλλοντικών και γενετικών παραγόντων στη ΝΑ.

Τέλος, και σε πρακτικότερο επίπεδο, ο συνδυασμός βιοδεικτών μπορεί να παράξει ένα χρήσιμο προγνωστικό εργαλείο, υπό τον όρο ότι θα προσαρμοστεί στις πραγματικές ανάγκες των ασθενών με το βάρος να πέφτει στην καλύτερη επιλογή των εξετάσεων. Παράγοντες όπως το κόστος και ο αριθμός των εξετάσεων θα πρέπει να ληφθούν υπόψη προκειμένου τα πολυτροπικά συστήματα να εισαχθούν στην κλινική διαδικασία.

ΚΕΦΑΛΑΙΟ 5: ΣΥΜΠΕΡΑΣΜΑΤΑ – ΜΕΛΛΟΝΤΙΚΗ ΕΡΕΥΝΑ

5.1 Συνολικά συμπεράσματα

Από το σύνολο των ερευνών που εξετάσαμε είναι προφανές ότι στην έρευνα για τη ΝΑ έχουν εισαχθεί πάρα πολλές νέες τεχνολογίες τα τελευταία χρόνια. Καθώς ο αριθμός των νοσούντων αυξάνεται λόγω της γήρανσης του πληθυσμού, η έρευνα στη ΝΑ μεγεθύνεται.

Δυστυχώς μέχρι τώρα δεν έχει βρεθεί κάποιος καθοριστικός παράγοντας για την εμφάνιση της ΝΑ. Και από τις μέχρι τώρα μελέτες συμπεραίνει κανείς ότι δεν θα προκύψει ένας κύριος επιβαρυντικός παράγοντας, καθώς η ΝΑ έχει πολύπλοκη αιτιολογία και εμφανίζεται με ποικίλη νευροπαθολογία.

Όπως είδαμε οι περισσότερες έρευνες αξιοποίησαν τη βάση της ADNI. Αυτό καταδεικνύει τη σημασία μίας βάσης που παρέχει δωρεάν τα δεδομένα της και είναι κοινώς αποδεκτή από την επιστημονική κοινότητα. Καθώς οι περισσότερες έρευνες έχουν τα ίδια δεδομένα, δημιουργείται ένα κοινό επίπεδο με βάση το οποίο μπορούν να αξιολογηθούν τα αποτελέσματα. Ταυτόχρονα βέβαια αυξάνει και τον κίνδυνο σφαλμάτων (όπως υπερπαραμετροποίηση), ιδιαίτερα γιατί δεν υπάρχουν άλλες εκτεταμένες βάσεις για την επανάληψη των αποτελεσμάτων.

Εδώ πρέπει να αναφερθούμε στην έλλειψη μεγάλων και καλά ορισμένων βάσεων δεδομένων για τη ΝΑ. Δεδομένου ότι εξετάζεται ένα θέμα με τεράστιο όγκο δεδομένων, η στατιστική βεβαιότητα περιορίζεται από το σχετικά μικρό δείγμα. Σε πολλές μελέτες οι επιστήμονες συγκράτησαν για περαιτέρω μελέτη περιπτώσεις με επιδόσεις χαμηλότερες από το προβλεπόμενο στατιστικό όριο, με την πεποίθηση ότι σε μεγαλύτερα δείγματα θα υπήρχε μεγαλύτερη επιρροή.

Επιστρέφοντας στη διάγνωση της ΝΑ, η βασική αρχή που ακολουθείται είναι η εφαρμογή νευροψυχολογικών δοκιμασιών. Ωστόσο, η επιβεβαίωση με απεικονιστικές τεχνικές, κυρίως sMRI, έχει ενταχθεί στα πρωτόκολλα διάγνωσης ΝΑ. Ωστόσο εδώ πρέπει να διαφοροποιήσουμε για μια ακόμη φορά τη διάγνωση από την πρόγνωση, που είναι ο κύριος στόχος στην έρευνα για τη ΝΑ, προκειμένου να μην προκληθούν ανεπανόρθωτες βλάβες στον εγκέφαλο.

Παρότι λοιπόν η διάγνωση της ΝΑ με νευροψυχολογικές δοκιμασίες είναι καλά τεκμηριωμένη, δεν είναι αποτελεσματική στην πρόγνωση, ούτε βέβαια στην ερμηνεία της εμφάνισης της ΝΑ. Οι νευροαπεικονιστικοί δείκτες φαίνεται να διακρίνουν καλύτερα τα πρώτα στάδια της νόσου. Η αυτοματοποίηση της κατάτμησης απεικονίσεων του εγκεφάλου επέτρεψε την εισαγωγή τους σε μεγάλες

έρευνες για τη ΝΑ και συμπεριλαμβάνονται στις περισσότερες εργασίες. Το βάρος δίνεται σε μορφομετρικά χαρακτηριστικά συγκεκριμένων περιοχών του εγκεφάλου, όπως ο ιππόκαμπος, που πλήττονται ιδιαίτερα από τη νόσο. Ωστόσο, με την εξέλιξη των μεθόδων επεξεργασίας, πληθαίνουν οι περιπτώσεις όπου εξετάζεται το σύνολο του εγκεφάλου. Η αύξηση της αξιοπιστίας των απεικονιστικών δεικτών στη ΝΑ επέτρεψε το συνδυασμό τους με γενετικούς παράγοντες.

Ο συνδυασμός απεικονιστικών τεχνικών και γενετικών δεικτών αναζωογόνησε πρόσφατα την έρευνα στη ΝΑ. Οι έρευνες στη νόσο είτε έχουν σαν εφελκτήριο κάποιο συγκεκριμένο γονίδιο, είτε αφορούν το σύνολο του γονιδιώματος (GWAS). Εξάχθηκαν αρκετά ενδιαφέροντα αποτελέσματα, αλλά σε καμία περίπτωση δεν επιλύθηκε το LOAD. Τα στατιστικά μεγέθη που κλήθηκαν οι μελετητές να διαχειριστούν με τις GWAS σίγουρα περιόρισαν εν μέρει τις δυνατότητές τους. Για να λυθεί αυτό το πρόβλημα, αλλά και για την εγκυρότερη κατανόηση της νόσου, πρόσφατα άρχισαν να χρησιμοποιούνται μέθοδοι παραγοντοποίησης των γενετικών δεδομένων και ομαδικής αποτίμησης της έκφρασής τους στον εγκέφαλο. Τα πολυπαραγοντικά GWAS δίνουν πλέον αρκετά έγκυρα αποτελέσματα, που πείθουν όμως ότι δεν θα δοθεί λύση στη ΝΑ με αμιγώς γενετικούς δείκτες. Ωστόσο το πεδίο αυτό δεν έχει εξαντληθεί: νέες έρευνες στους γενετικούς δείκτες σχεδιάζονται, ενώ έχει επεκταθεί ο συνδυασμός τους με άλλες πηγές δεδομένων. Η μεταβολή των επιγενετικών μηχανισμών από εξωτερικούς παράγοντες είναι ίσως η πιο ελπιδοφόρα από αυτές.

Η τελευταία τάση στην έρευνα για τη ΝΑ είναι τα συστήματα διαχείρισης πολλών πηγών δεδομένων. Αυτά επιτρέπουν την ταυτόχρονη αξιοποίηση πολλών δεικτών για τη ΝΑ. Έτσι αξιολογήθηκε η εγκυρότητα κάθε προσέγγισης στην πρόωμη διάγνωση της ΝΑ, με τους μορφομετρικούς δείκτες να έχουν τα πρωτεία. Ταυτόχρονα, η διαγνωστική ακρίβεια αυξήθηκε κατά τουλάχιστον 5%. Ωστόσο αυτά τα συστήματα δεν έχουν εξαντλήσει τις δυνατότητές τους. Η πρόβλεψη τιμών δεδομένων και η εισαγωγή καλά ορισμένων δεδομένων θα βοηθήσει πολύ την αποδοτικότητά τους. Η εισαγωγή περιβαλλοντικών παραγόντων μπορεί επίσης να δώσει νέα ώθηση στα πολυτροπικά συστήματα. Τέλος, η εφαρμογή τους στη διαγνωστική διαδικασία έχει δειχθεί εφικτή, αν και ακόμη αρκετά μακρινή.

5.2 Νέες ερευνητικές προοπτικές και στόχοι

Όπως αναφέραμε πολλές φορές η έλλειψη αρκετά μεγάλων και καλά ορισμένων δειγμάτων αποτελεί μεγάλη τροχοπέδη για την έρευνα στη ΝΑ. Είναι εξαιρετικά σημαντικό να μεγαλώσουν τα σημερινά μεγέθη των διαθέσιμων βάσεων δεδομένων, αν και το κόστος είναι δυσβάσταχτο. Πολλές

έρευνες, που αξιολογούν γενετικούς ή απεικονιστικούς δείκτες, δεν καταλήγουν σε ασφαλή συμπεράσματα, ενώ μεγαλύτερο δείγμα θα τους έδινε αυτή τη δυνατότητα. Μία πιθανή λύση μπορεί να δοθεί με την ταυτόχρονη χρήση πολλών βάσεων, όπως για παράδειγμα έγινε στην εργασία που είδαμε για το CSF [39], αν και αυτό εγείρει θέματα κανονικοποίησης των δεδομένων.

Σε αυτό το πλαίσιο, η ADNI επεκτείνεται, αν και όχι για να αυξήσει το δείγμα της, αλλά για να συμπεριλάβει νέους δείκτες. Υπάρχουν αρκετοί νέοι δείκτες που θα εισαχθούν, ανάμεσα στους οποίους και εξωτερικοί παράγοντες, όπως ο πρότερος τραυματισμός στο κεφάλι ή η ύπαρξη διαταραχής μετατραυματικού στρες [63]. Η εισαγωγή νέων δεικτών θα προσφέρει διαφορετικούς ερευνητικούς στόχους.

Στις απεικονιστικές μελέτες νέοι δείκτες εισάγονται. Η αξιολόγηση του florbetapir-PET στη ΝΑ είναι ένα ακόμη ανοιχτό θέμα, που πιθανόν θα δια φωτίσει το σχηματισμό αμυλοειδικών πλακών ή θα προσφέρει νέους γενετικούς δείκτες. Ταυτόχρονα, μελετάται η δυνατότητα αξιολόγησης του συνόλου του εγκεφάλου σε συνδυασμό με γενετικά δεδομένα και τα πιθανά πλεονεκτήματα που προσφέρει (που θα βοηθήσει για παράδειγμα στην εισαγωγή της εκτεταμένης παθολογίας της ΝΑ, αλλά θα δυσκολέψει το ερευνητικό έργο λόγω του όγκου των δεδομένων).

Στις γενετικές μελέτες εξετάζεται η εισαγωγή τεχνικών εκτεταμένης αναδιάταξης του γονιδιώματος, που πιθανόν να αποκαλύψουν κάποιον μη προφανή δείκτη για τη ΝΑ. Αυτό είναι ίσως το πιο ενδιαφέρον θέμα για τη μελλοντική έρευνα στη ΝΑ, αν και έχει τεράστιο υπολογιστικό κόστος και θα προκαλέσει ακόμη περισσότερα εμπόδια στη στατιστική επεξεργασία (το μέγεθος του γονιδιώματος για παραγοντοποίηση θα αυξηθεί κατακόρυφα). Σε αυτό το πλαίσιο, η παραγοντοποίηση του γονιδιώματος σε συνάρτηση με γενετικούς χάρτες έκφρασης στον εγκέφαλο δεν έχει εξαντληθεί ακόμη. Τα περισσότερα πολυπαραγοντικά GWAS που έχουν πραγματοποιηθεί βασίστηκαν στην απλή γεινίαση στο γονιδίωμα, ενώ η βιολογική έκφραση των ομάδων SNPs στις λειτουργίες του εγκεφάλου μπορεί να δια φωτίσει τον τρόπο ανάπτυξης της ΝΑ [19]. Τέλος, αρκετές νέες έρευνες επικεντρώνονται στους επιγενετικούς μηχανισμούς στη ΝΑ και στη σχέση τους με περιβαλλοντικούς παράγοντες. Δεδομένου ότι η σχέση της ΝΑ με άλλες ασθένειες ή εξωτερικούς παράγοντες είναι καλά τεκμηριωμένη, η περαιτέρω διερεύνηση αυτού του θέματος, όπως και ο τρόπος μεταβολής των επιγενετικών μηχανισμών με τη γήρανση, θα απασχολήσει τους μελλοντικούς ερευνητές. Σχεδιάζονται για παράδειγμα μελέτες που θα επικεντρωθούν στην επιγενετική κατάσταση των επιβαρυντικών γονιδίων κατά τη μεταβολή από υγιή σε MCI και από MCI σε ΝΑ.

Στη διαχείριση διαφορετικών πηγών πληροφορίας σημαντικό ρόλο θα παίξουν τα πολυτροπικά συστήματα. Ουσιαστικά σε αυτό το σημείο γνωρίζουμε ότι οι πολυτροπικές έρευνες προσφέρουν πολλά πλεονεκτήματα, αλλά δεν έχει γίνει κάποια πραγματικά εκτεταμένη έρευνα σε αυτή την κατεύθυνση, γιατί δεν υπάρχουν μεγάλες βάσεις δεδομένων που συμπεριλαμβάνουν περιβαλλοντικούς και επιδημιολογικούς παράγοντες και εξετάσεις σε μεγάλο βάθος χρόνου. Οι πραγματικές δυνατότητες αυτών των συστημάτων μας διαφεύγουν, καθώς δεν έχουν εξαντληθεί οι πιθανές συσχετίσεις περιβαλλοντικών παραγόντων με την εμφάνιση της ΝΑ και την έκφρασή τους σε αντίστοιχους βιοδείκτες. Η αποτελεσματική εισαγωγή τέτοιων “παγκόσμιων” δεικτών στην κλινική διαδικασία επίσης είναι ένα ανοιχτό ερευνητικό θέμα που εγείρει πολλά ζητήματα. Πρώτα από όλα, απαιτείται ένα κοινά αποδεκτό CDSS με στιβαρή οντολογία. Αν και αυτό είναι ένα θέμα με το οποίο δεν ασχοληθήκαμε στο πλαίσιο αυτής της διπλωματικής, η εισαγωγή πολλών διαφορετικών δεικτών και η ακόλουθη διάγνωση, προϋποθέτει ότι θα δοθεί βάρος στην οντολογική οργάνωση και στους κανόνες που θα τη διέπουν [64][65].

Η έρευνα για τη ΝΑ -όπως για κάθε περίπλοκη ασθένεια- είναι εφελτήριο για την οργάνωση εκτεταμένων βάσεων δεδομένων, προκειμένου να διευρυνθεί η δεξαμενή από όπου αντλείται το δείγμα μας. Η υλοποίηση ομολογουμένως μακρινών στόχων, όπως η διάθεση αναλυτικών ηλεκτρονικών φακέλων σε εθνικό επίπεδο, θα παρείχε εκτεταμένα επιδημιολογικά χαρακτηριστικά για τη νόσο και θα βοηθούσε την αντιμετώπισή της. Αυτό είναι το πλαίσιο στο οποίο εγείρονται και ορισμένα νομικά και ηθικά ζητήματα, όπως η χρήση του γενότυπου ενός ανθρώπου στην έρευνα, που θα πρέπει να επιλυθούν [66].

Τέλος, υπάρχουν αρκετά ακόμη σημεία όπου η τεχνολογία μπορεί να προσφέρει λύσεις πέρα από την έρευνα για τη νόσο: αυτά περιλαμβάνουν την αυτοματοποιημένη χορήγηση νευροψυχολογικών δοκιμασιών, ακόμη και μέσω τηλεφώνου, για τη συλλογή επιδημιολογικών χαρακτηριστικών και την έγκαιρη πρόληψη [67][68] ή την ανάπτυξη εργαλείων που διευκολύνουν ασθενείς με ΝΑ [69].

Βιβλιογραφία

- [1] Brookmeyer et al.: “Forecasting the global burden of Alzheimer's disease”, 2007, Alzheimer's and Dementia
- [2] Brookmeyer et al.: “Projections of Alzheimer's Disease in the United States and the Public Health Impact of Delaying Disease Onset”, 1998, American Journal of Public Health
- [3] Gray: “Βασικές αύλακες και λοβοί του εγκεφαλικού φλοιού υπό οριζόντια άποψη”, 2008, Anatomy of the Human Body, φιγούρα 728, Bartleby
- [4] Sutton Amy: “Alzheimer’s Disease Sourcebook 5th Edition”, 2011, Omnigraphics
- [5] Ramanan et al.: “APOE and BCHE as modulators of cerebral amyloid deposition: a florbetapir PET genome-wide association study”, 2013, Molecular Psychiatry
- [6] Petersen et al.: “Mild cognitive impairment: clinical characterization and outcome”, 1999, Arch. Neurol. 56
- [7] Morris et al.: “Mild cognitive impairment represents early-stage Alzheimer’s Disease”, 2001, Archives of Neurology 58
- [8] Liu et al.: “Neuropsychological prediction of conversion to Alzheimer’s Disease in patients with mild cognitive impairment”, 2006, Arch. Gen. Psychiatry 63
- [9] Eli Vakil: “Neuropsychological assessment: Principles, rationale, and challenges”, 2012, Psychology Press, 34 (2), 135–150
- [10] Jeffrey Cummings: “Theories Behind Existing Scales for Rating Behavior in Dementia”, 1996, International Psychogeriatrics. Vol. 8, Suppl. 3
- [11] Perneczky R et al.: “Mapping scores onto stages: mini-mental state examination and clinical dementia rating”, 1994, Acta Neurol Scand.
- [12] Χατζηαδαμίδου Θέκλα: “Νοητικές διαταραχές και η εξέλιξή τους σε υπερήλικες με σακχαρώδη διαβήτη τύπου ΙΙ”, 2010, Αριστοτέλειο Πανεπιστήμιο

- [13] Braskie et al.: “Recent Advances in Imaging Alzheimer’s Disease”, 2012, Journal of Alzheimer Disease 30
- [14] Anoop et al.: “CSF Biomarkers for Alzheimer's Disease Diagnosis”, 2010, International Journal of Alzheimer's Disease
- [15] Saykin et al.: “Alzheimer’s Disease Neuroimaging Initiative biomarkers as quantitative phenotypes: Genetics core aims, progress, and plans”, 2010, Elsevier
- [16] Scott Heutel: “An introduction to functional MRI”, Duke University
- [17] Shmueli et al.: “Low-frequency fluctuations in the cardiac rate as a source of variance in the resting-state fMRI BOLD signal”, 2007, Neuroimage
- [18] Willem de Haan, et al.: “Functional neural network analysis in frontotemporal dementia and Alzheimer's disease using EEG and graph theory”, 2009, BMC Neuroscience
- [19] Calhoun et al.: “A large scale multivariate parallel ICA method reveals novel imaging-genetic relationships for Alzheimer’s disease in the ADNI cohort”, 2012, Neuroimage
- [20] Nuno Vasconcelos, “Intelligent Data Analysis and Probabilistic Inference Lecture 15”, Imperial College
- [21] Vince D. Calhoun, Jingyu Liu, Tülay Adalı : “A review of group ICA for fMRI data and ICA for joint inference of imaging, genetic, and ERP data”, 2008, Elsevier
- [22] Jingyu Liu, Vince Calhoun: ”Parallel Independent Component Analysis for Multimodal Analysis: Application to fMRI and EEG Data”, 2007, IEEE
- [23] Vounou, et al.: “Discovering Genetic Associations with High-Dimensional Neuroimaging Phenotypes: A sparse reduced rank regression approach”, 2011, Neuroimage
- [24] Hennings et al: “Improving the Prediction of Clinical Outcomes from Genomic Data Using Multiresolution Analysis” 2012, IEEE
- [25] Βλαχάβας et al.: “Τεχνητή Νοημοσύνη Γ΄έκδοση”, 2006, Β. Γκιούρδας

- [26] Morra et al.:“Validation of a fully automated 3D hippocampal segmentation method using subjects with Alzheimer's disease mild cognitive impairment, and elderly controls”, 2008, Neuroimage 43
- [27] Morra et al.:“Comparison of AdaBoost and Support Vector Machines for Detecting Alzheimer’s Disease Through Automated Hippocampal Segmentation”, 2010, IEEE Transactions on Medical Imaging
- [28] David Glenn Clark:“Residual Vectors for Alzheimer’s disease diagnosis and prognostication”, 2011, Brain and Behavior
- [29] Gerard D. Schellenberg, Thomas J. Montine “The genetics and neuropathology of Alzheimer’s disease”, 2012, Acta Neuropathol. 124:305–323
- [30] Reiman et al.: “Linking Brain Imaging and Genomics in the Study of Alzheimer Disease and Aging”, 2007, New York Academy of Sciences
- [31] Reiman et al.: “Brain imaging and fluid biomarker analysis in young adults at genetic risk for autosomal dominant Alzheimer disease in the presenilin 1 E280A kindred: a case-control study”, 2012, The Lancet
- [32] Braskie et al.: “Common Alzheimer Disease Risk Variant within the CLU gene affects white matter microstructure in young adults”, 2011, The Journal of Neuroscience
- [33] Hibar et al.: “Alzheimer Disease Risk Gene, GAB2, is Associated with Regional Brain Volume Differences in 755 Young Healthy Twins”, 2012, Twin Research and Human Genetics
- [34] A. Corvin, N. Craddock and P. F. Sullivan: “Genome-wide association studies: a primer”, 2010, Psychological Medicine 40
- [35] Paul M. Thompson et al.: “Imaging Genomics”, 2010, Current Opinion in Neurology, 23:368–373
- [36] Bertram et al.: “The Genetics of Alzheimer Disease: Back to the Future”, 2010, Neuron
- [37] Liu et al.: “A Parallel Independent Component Analysis Approach to Investigate Genomic Influence on Brain Function”, 2008, IEEE

- [38] Stein et al.: “Genome Wide Analysis Reveals Novel Genes Influencing Temporal Lobe Structure With Relevance to Neurodegeneration in Alzheimer’s Disease”, 2010, Neuroimage
- [39] Cruchaga et al.: “GWAS of Cerebrospinal Fluid Tau Levels Identifies Risk Variants for Alzheimer’s Disease”, 2013, Neuron
- [40] Hibar et al.: “Multilocus genetic analysis of brain images”, 2011, Frontiers in Genetics Vol.2
- [41] Hibar, et al.: “Boosting Power to Detect Genetic Associations in Imaging Using Multi-Locus, Genome-Wide Scans and Ridge Regression”, 2011, ISBI
- [42] Vounou et al.: “Sparse reduced-rank regression detects genetic associations with voxel-wise longitudinal phenotypes in Alzheimer’s disease”, 2012, Neuroimage
- [43] Judith Miklossy “Emerging roles of pathogens in Alzheimer disease”, 2011, Expert Reviews in Molecular Medicine
- [44] Leonidas Chouliaras, et al:“Epigenetic regulation in the pathophysiology of Alzheimer’s disease”, 2011, Progress in Neurobiology
- [45] Tang MX, et al:.”Effect of age, ethnicity, and head injury on the association between APOE genotypes and Alzheimer’s disease”, 1996, Annual NY Academic Science; 802:6e15.
- [46] Mastroeni, et, al: “ Epigenetic differences in cortical neurons from a pair of monozygotic twins discordant for Alzheimer’s disease”, 2009, PLoS One 4, e6617.
- [47] Kannayiram Alagiakrishnan, et al.“Genetics and epigenetics of Alzheimer’s disease”, 2012, Post Graduation Medicine Journal
- [48] Fraga, Esteller: “Epigenetics and aging: the targets and the marks”, 2007, Trends Genet. 23, 413–418.
- [49] Diego Mastroeni, et al:“Epigenetic mechanisms in Alzheimer’s disease”, 2011, Neurobiology of Aging
- [50] Christian Behl, “Brain aging and late-onset Alzheimer’s disease: Many open questions”, 2012, International Phychogeriatrics

- [51] Ryu et al.: “Epigenetic modification in a monozygotic twin with Alzheimer's disease”, 2008, Elsevier
- [52] Fischer A, et al.:”Recovery of learning and memory is associated with chromatin remodeling”, 2007, Nature; 447:178e82
- [53] Yue Huang, Xuezhi Sun, Guangshu Hu: “An integrated genetics approach for identifying protein signal pathways of Alzheimer’s disease”, 2012, Computer Methods in Biomechanics and Biomedical Engineering
- [54] Zhi-Ping Liu et al: “Dynamically dysfunctional protein interactions in the development of Alzheimer’s disease”, 2009, IEEE
- [55] Polikar et al:“Multimodal EEG, MRI and PET Data Fusion for Alzheimer Disease Diagnosis”, 2010, IEEE
- [56] Hinrichs et al.: “Predictive markers for AD in a multi-modality framework: An analysis of MCI progression in the ADNI population”, 2011, Neuroimage
- [57] Jieping Ye, et al:”Heterogeneous Data Fusion for Alzheimer Disease Study”, 2008, KDD ‘08 August 24
- [58] Zhang, Wang, et al: “Multimodal Classification of Alzheimer Disease and Mild Cognitive Impairment”, 2011, Neuroimage
- [59] Thompson et al.: “Multi-Source Learning for Joint Analysis of Incomplete Multi-Modality Neuroimaging Data”, 2012, KDD’12, August 12–16
- [60] Escudero, Zajicek, et al: “Machine LearningBased Method for Personalized and Cost-Effective Detection of Alzheimer’s Disease”, 2012, IEEE
- [61] Van Gills et al: “Discovery and use of efficient biomarkers for objective disease state assessment in Alzheimer’s disease”, 2010, IEEE
- [62] Van Gills, Mattila, et al: “Design and Application of a Generic CDSS for Multiscale Data”, 2012, IEEE

- [63] ADNI, “New ADNI studies seek volunteers with memory concerns, and Vietnam veterans”, 2013, <http://www.nia.nih.gov/alzheimers/announcements/2013/03/new-adni-studies-seek-volunteers-memory-concerns-and-vietnam>
- [64] Pedraza et al.: “Robust and Expanded Norms for the Dementia Rating Scale”, 2010, *Archives of Clinical Neuropsychology* 25
- [65] Eklund et al.: “Early diagnosis of cognitive disorder based on ground categories for decision support structures”, 2009, *IEEE*
- [66] Marshall B. Kapp: “Physicians’ legal duties regarding the use of genetic tests to predict and diagnose Alzheimer Disease”, 2000, *Journal of Legal Medicine*
- [67] Hodges et al.: “Automatic Assessment of Cognitive Impairment through Electronic Observation of Object Usage”, 2010, *Pervasive*
- [68] Wolfson et al.: “Telephone-Administered Cognitive Tests as Tools for the Identification of Eligible Study Participants for Population-Based Research in Aging”, 2009, *Canadian Journal on Aging*
- [69] Sugihara et al.: “Gaps between Assistive Technologies and Dementia Care”, 2012, *IEEE*