



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

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

ΤΙΤΛΟΣ ΠΑΡΑΔΟΤΕΟΥ: Ανάλυση ιατρικών και γενετικών δεδομένων

ΕΝΟΤΗΤΑ ΕΡΓΑΣΙΑΣ 4: Ανάπτυξη κυτταρικών και ζωικών μοντέλων, καθώς και νέων βιοδεικτών για τις νευροεκφυλιστικές ασθένειες Alzheimer και Parkinson.

ΥΠΕΥΘΥΝΗ ΕΡΕΥΝΗΤΙΚΗ ΟΜΑΔΑ (ΦΟΡΕΑΣ): ΔΗΜΗΤΡΙΟΣ ΦΩΤΙΑΔΗΣ (ΙΤΕ)

Ανάλυση ιατρικών και γενετικών δεδομένων

ΕΙΣΑΓΩΓΗ

Η Πρόκληση των Νευροεκφυλιστικών Νόσων

Οι νευροεκφυλιστικές νόσοι συνιστούν μείζον βιοϊατρικό και κοινωνικοοικονομικό πρόβλημα, καθώς χαρακτηρίζονται από **προοδευτική, μη αναστρέψιμη απώλεια νευρωνικών κυκλωμάτων**, μακρά **προκλινική/προσυμπτωματική φάση**, και συχνά ετερογενή κλινική έκφραση που αντανακλά την πολυπαραγοντική τους αιτιοπαθογένεια [1], [2]. Μεταξύ αυτών, η νόσος Αλτσχάιμερ (Alzheimer's disease, AD) και η νόσος Πάρκινσον (Parkinson's disease, PD) αποτελούν τις συχνότερες και κλινικά πλέον επιβαρυντικές οντότητες, με σημαντική επιβάρυνση σε λειτουργικότητα, ποιότητα ζωής, φροντιστικό φορτίο και δαπάνες υγείας [1], [3].

Η βασική πρόκληση έγκειται στο ότι η κλινική διάγνωση βασίζεται σε **συμπτωματολογία** που συνήθως εμφανίζεται **αφού έχει ήδη εγκατασταθεί σημαντική νευροβιολογική βλάβη**. Στην AD, η παθολογία αμυλοειδούς-β (Aβ) και tau μπορεί να προηγείται της κλινικής έκπτωσης κατά πολλά έτη, ενώ η βιολογική ταξινόμηση των βιοδεικτών (π.χ. A/T/N) αναδεικνύει την ανάγκη να αντιμετωπίζεται η νόσος ως συνεχές βιολογικών μεταβολών και όχι ως δυαδική (νόσος/μη νόσος) κατάσταση [2]. Αντίστοιχα, στην PD, η νευροεκφυλιστική διεργασία δεν περιορίζεται στο νιγροραβδωτό σύστημα αλλά περιλαμβάνει ευρύτερα δίκτυα, με αποτέλεσμα η κλινική εικόνα να εκτείνεται πέραν των κινητικών συμπτωμάτων (μη κινητικά συμπτώματα, γνωστική έκπτωση, αυτόνομες δυσλειτουργίες) [3], [4].

Τα τελευταία χρόνια, η διεθνής βιβλιογραφία υποστηρίζει με αυξανόμενη ισχύ ότι ένα σημαντικό ποσοστό του συνολικού κινδύνου άνοιας σχετίζεται με **τροποποιήσιμους παράγοντες** (π.χ. αγγειακοί/μεταβολικοί παράγοντες, φυσική αδράνεια, αισθητηριακές απώλειες, κοινωνική απομόνωση, περιβαλλοντικοί παράγοντες), και ότι η μείωση της επιβάρυνσης από αυτούς μπορεί να συμβάλει σε πρόληψη ή καθυστέρηση της κλινικής έκφρασης [1]. Ωστόσο, η μετάφραση αυτής της γνώσης σε κλινικά αξιοποιήσιμα εργαλεία απαιτεί:

- (α) ακριβή **στρωματοποίηση κινδύνου** σε επίπεδο ατόμου,
- (β) έγκαιρη **ανίχνευση βιολογικών μεταβολών** πριν την κλινική έκπτωση, και
- (γ) αξιόπιστη **πρόγνωση εξέλιξης** με βάση ετερογενείς πηγές δεδομένων [1], [2], [5].

Στο πλαίσιο αυτό, το έργο **Brain Precision** στοχεύει στην επιστημονικά τεκμηριωμένη μετάβαση προς **Ιατρική Ακριβείας** στις νευροεκφυλιστικές νόσους, αξιοποιώντας **δεδομένα πολλαπλών τύπων** (ιατρικά/κλινικά/νευροψυχολογικά και — κρίσιμα — γενετικά/γενομικά). Η κεντρική παραδοχή είναι ότι η ενοποίηση κλινικών φαινοτύπων με μοριακή πληροφορία και η εφαρμογή σύγχρονων μεθόδων **βιοπληροφορικής και μηχανικής μάθησης** μπορεί να αποκαλύψει: (i) νέους μοριακούς δείκτες, (ii) βιολογικά ερμηνεύσιμα υποσύνολα ασθενών, και (iii) υπογραφές κινδύνου/πρόγνωσης που υπερβαίνουν τους παραδοσιακούς μονοδιάστατους δείκτες [1], [2], [5], [6].

Κλινική εικόνα και παθοφυσιολογία των νόσων Αλτσχάιμερ και Πάρκινσον

Νόσος Αλτσχάιμερ (AD)

Κλινικά, η AD εκδηλώνεται τυπικά με **προοδευτική διαταραχή επεισοδιακής μνήμης**, ακολουθούμενη από έκπτωση εκτελεστικών λειτουργιών, γλώσσας, οπτικοχωρικών δεξιοτήτων και τελικά γενικευμένη γνωστική και λειτουργική επιβάρυνση. Σε πολλά άτομα προηγείται στάδιο **ήπιας γνωστικής διαταραχής (MCI)**, το οποίο αντιπροσωπεύει ενδιάμεση κατάσταση μεταξύ φυσιολογικής γήρανσης και άνοιας, με αυξημένο αλλά ετερογενή κίνδυνο μετατροπής σε άνοια [2], [5].

Σε παθοφυσιολογικό επίπεδο, η AD σχετίζεται με:

1. **απόθεση Αβ** (πλάκες),
 2. **συσσώρευση παθολογικής tau** (νευροϊνιδιακά δεμάτια),
 3. **νευροεκφύλιση** και **απώλεια** συναπτικών συνδέσεων.
- Το ερευνητικό πλαίσιο NIA-AA και η περιγραφική ταξινόμηση A/T/N έχουν εδραιώσει μια βιολογική θεώρηση όπου οι βιοδείκτες αμυλοειδούς και tau, σε συνδυασμό με δείκτες νευροεκφύλισης, χρησιμοποιούνται για την τεκμηρίωση παθολογικών διεργασιών ανεξάρτητα από το αν έχει εμφανιστεί ήδη πλήρης κλινική άνοια [2]. Παράλληλα, αναγνωρίζεται ότι η παρουσία βιοδεικτών δεν ισοδυναμεί πάντα με άμεση κλινική έκπτωση (ιδίως σε μεγαλύτερες ηλικίες), γεγονός που καθιστά την **πρόγνωση ατομικού κινδύνου** και τον χρόνο μετατροπής σε κλινική νόσο ένα σύνθετο πρόβλημα που απαιτεί ολοκληρωμένα μοντέλα [1], [5].

Νόσος Πάρκινσον (PD)

Η PD χαρακτηρίζεται κλινικά από τον συνδυασμό **βραδυκινησίας** με τρόμο ηρεμίας και/ή μυϊκή δυσκαμψία, με προοδευτική εξέλιξη και συχνές επιπλοκές όπως διακυμάνσεις κινητικής απόκρισης και δυσκινησίες σε προχωρημένα στάδια. Ιδιαίτερη σημασία έχουν τα **μη κινητικά συμπτώματα** (διαταραχές ύπνου, κατάθλιψη/άγχος, αυτόνομες δυσλειτουργίες, ανοσμία), τα οποία μπορεί να προηγούνται των κλασικών κινητικών εκδηλώσεων και να λειτουργούν ως ενδείξεις προσυμπτωματικής φάσης [3], [4].

Παθοφυσιολογικά, η PD σχετίζεται με εκφύλιση ντοπαμινεργικών νευρώνων και παθολογική συσσώρευση **α-συνουκλεΐνης** (συνουκλεϊνοπάθεια). Η νόσος θεωρείται πλέον διαταραχή δικτύων με πολυσυστηματική συμμετοχή (όχι μόνο κινητικά κυκλώματα), γεγονός που εξηγεί την κλινική ετερογένεια και την ανάγκη ολοκληρωμένων βιολογικών/κλινικών δεικτών για ακριβή στρωματοποίηση [3], [4].

Κοινά στοιχεία και επιπτώσεις για μοντελοποίηση κινδύνου

Και στις δύο νόσους, η κρίσιμη επιστημονική δυσκολία είναι η **ετερογένεια**: διαφορετικές τροχιές εξέλιξης, συννοσηρότητα, και αλληλεπίδραση γενετικών παραγόντων (π.χ. πολυγονδιακή επιβάρυνση) με τροποποιήσιμους κινδύνους και βιολογικές διεργασίες [1], [2], [6]. Επομένως, προσεγγίσεις που συνδυάζουν **κλινικό φαινότυπο** με **γενομικά δεδομένα** και εφαρμόζουν ισχυρές μεθόδους βιοπληροφορικής/ML έχουν σαφή αιτιολόγηση: μπορούν να αποτυπώσουν πολύπλοκες μη γραμμικές σχέσεις, να αναδείξουν υποομάδες και να εντοπίσουν μοριακές υπογραφές που συνδέονται με έναρξη, ρυθμό εξέλιξης ή μετατροπή σε κλινικά έκδηλη νόσο [1], [5], [6].

1.1.2 Η ανάγκη για έγκαιρη πρόγνωση: Από την κλινική διάγνωση στους μοριακούς και ψηφιακούς δείκτες

Η κλινική διάγνωση των νευροεκφυλιστικών νόσων, παρότι απαραίτητη, παραμένει δομικά περιορισμένη ως προς την **έγκαιρη ανίχνευση** και την **προγνωστική ακρίβεια**. Ο κύριος λόγος είναι ότι τα κλασικά διαγνωστικά κριτήρια βασίζονται κυρίως σε φαινοτυπικά συμπτώματα (γνωστικά, κινητικά, λειτουργικά), τα οποία εμφανίζονται όταν οι υποκείμενες βιολογικές διεργασίες έχουν ήδη εξελιχθεί σημαντικά. Στη νόσο Αλτσχάιμερ, η αλληλουχία παθολογικών μεταβολών (αμυλοειδές, tau, νευροεκφύλιση) μπορεί να προηγείται της κλινικής έκπτωσης επί σειρά ετών, ενώ αντίστοιχα στη νόσο Πάρκινσον παρατηρείται προσυμπτωματικό στάδιο με μη-κινητικές εκδηλώσεις και προοδευτική δικτυακή νευροεκφύλιση [7].

Η παραπάνω πραγματικότητα έχει οδηγήσει σε μια σαφή μετατόπιση: από τη διάγνωση «μετά την εγκατάσταση συμπτωμάτων» προς τη **βιολογικά τεκμηριωμένη πρόγνωση** και τη **στρωματοποίηση κινδύνου**. Στο πλαίσιο αυτό, το σχήμα A/T/N (Amyloid/Tau/Neurodegeneration) εισάγει ένα ουδέτερο, βιοδείκτη-κεντρικό σύστημα περιγραφής των βιολογικών υποστρωμάτων της AD, ανεξάρτητα από το κλινικό στάδιο, επιτρέποντας ορισμό της νόσου ως συνεχές βιολογικής εξέλιξης [7]. Η λογική αυτή δεν είναι απλώς ταξινομική: καθιστά εφικτή την ανάπτυξη προγνωστικών μοντέλων που εκτιμούν **πιθανότητα μετατροπής** (π.χ. από γνωστικά φυσιολογική κατάσταση ή MCI προς άνοια) και **χρονικό ορίζοντα εξέλιξης**, βάσει αντικειμενικών δεικτών.

Οι **μοριακοί/βιοχημικοί δείκτες** (σε ENY ή/και αίμα) ενισχύουν σημαντικά την πρόωμη ανίχνευση και την πρόγνωση. Ειδικά οι δείκτες πλάσματος, όπως GFAP, p-tau, NfL, καθώς και συνδυασμοί βιοδεικτών, έχουν αναδειχθεί ως ιδιαίτερα υποσχόμενοι για κλινική κλιμάκωση, λόγω της πρακτικότητας και της δυναμικής ευρείας εφαρμογής τους σε πληθυσμιακό επίπεδο [8]–[10]. Ενδεικτικά, η GFAP στο πλάσμα έχει συσχετισθεί με πρώιμες φάσεις αμυλοειδικής παθολογίας, πριν την εκδήλωση σαφούς κλινικής άνοιας [9], ενώ πολυβιοδεικτικοί συνδυασμοί (CSF/πλάσμα) έχουν δείξει αυξημένη ικανότητα πρόβλεψης γνωστικής έκπτωσης και κλινικής μετατροπής, ειδικά σε ετερογενείς κοόρτες [10]. Συνεπώς, η πρόγνωση μετατρέπεται από «εκτίμηση βάσει συμπτωμάτων» σε **ποσοτικοποιημένη εκτίμηση κινδύνου** με βιολογική ερμηνεία.

Παράλληλα, η επιτάχυνση των ψηφιακών τεχνολογιών εισάγει μια δεύτερη κατηγορία δεικτών: τους **ψηφιακούς βιοδείκτες**. Μετρήσεις από κινητές συσκευές, wearables και παθητική παρακολούθηση συμπεριφοράς/κινητικότητας/ύπνου, καθώς και μετρικές αλληλεπίδρασης (π.χ. typing, speech, gait proxies), μπορούν να αποτυπώσουν λεπτές, συνεχείς μεταβολές λειτουργικής κατάστασης που συχνά διαφεύγουν από σποραδικές κλινικές επισκέψεις [11]. Επιπλέον, η ευρεία διάδοση εφαρμογών κινητής υγείας και εργαλείων ψηφιακού screening για γνωστική έκπτωση έχει δημιουργήσει ένα οικοσύστημα όπου η πρόγνωση μπορεί να υποστηριχθεί από συνδυασμό κλασικών κλινικών δεδομένων και ψηφιακών σημάτων, με τρόπο κλιμακώσιμο και οικονομικά βιώσιμο [12].

Η ουσία της ανάγκης για έγκαιρη πρόγνωση είναι διττή: (α) να αναγνωρίζονται άτομα σε **υψηλό κίνδυνο** πριν την μη αναστρέψιμη βλάβη, και (β) να επιτρέπεται **εξατομίκευση παρακολούθησης και παρέμβασης** με βάση βιολογικά και λειτουργικά προφίλ. Αυτό απαιτεί συστηματική ενοποίηση μοριακών και ψηφιακών δεικτών με κλινικούς φαινοτύπους, ώστε η πρόγνωση να στηρίζεται σε πολυδιάστατη τεκμηρίωση και όχι σε μεμονωμένες, αποσπασματικές μετρήσεις [8], [11].

1.1.3 Η μετάβαση στην Ιατρική Ακριβείας (Precision Medicine): Πολυτροπική (Multimodal) ανάλυση δεδομένων

Η **Ιατρική Ακριβείας** προτείνει μια θεμελιώδη αλλαγή παραδείγματος: από την «μέση» θεραπευτική/διαγνωστική προσέγγιση προς εξατομικευμένες αποφάσεις που λαμβάνουν υπόψη τη **βιολογική ετερογένεια** και το **ατομικό προφίλ κινδύνου**. Στις νευροεκφυλιστικές νόσους, η ανάγκη αυτή είναι ιδιαίτερα έντονη, διότι τόσο η AD όσο και η PD εμφανίζουν πολλαπλές τροχιές εξέλιξης, διαφορετικές υποκείμενες βιολογικές διαδρομές, καθώς και ισχυρές αλληλεπιδράσεις μεταξύ γενετικών/μοριακών παραγόντων και κλινικο-περιβαλλοντικών εκθέσεων [7], [8]. Επομένως, η ακριβής στρωματοποίηση δεν μπορεί να επιτευχθεί επαρκώς με μία μόνο κατηγορία δεδομένων.

Η **πολυτροπική (multimodal) ανάλυση** αποτελεί λειτουργικό πυρήνα της Ιατρικής Ακριβείας, καθώς συνδυάζει:

- **Κλινικά και δημογραφικά δεδομένα** (ηλικία, φύλο, συννοσηρότητες, φαρμακευτική αγωγή),
- **Νευροψυχολογικά/λειτουργικά δεδομένα** (σκορ γνωστικών δοκιμασιών, δείκτες λειτουργικότητας),
- **Μοριακούς/βιοχημικούς δείκτες** (Aβ, tau-related markers, GFAP, NfL κ.ά.),
- **Γενετικά/γενομικά δεδομένα** (παραλλαγές κινδύνου, πολυγονιδιακά προφίλ, μοριακές υπογραφές),
- **Ψηφιακούς βιοδείκτες** (παθητική συμπεριφορά, κινητικότητα, ύπνος, αλληλεπίδραση με συσκευές).

Η αξία της ανάλυσης πολλαπλών τύπων δεδομένων δεν είναι απλώς αθροιστική. Επιτρέπει την ταυτόχρονη μοντελοποίηση βιολογικού υποστρώματος και λειτουργικού αποτελέσματος, μειώνοντας την αβεβαιότητα που προκύπτει όταν ένα μοντέλο στηρίζεται μονοδιάστατα σε κλινικό φαινότυπο ή σε έναν μόνο βιοδείκτη [8]–[11].

Κρίσιμο στοιχείο της μετάβασης αυτής είναι ότι τα δεδομένα πολλαπλών τύπων είναι υψηλής διαστασιμότητας και ετερογένειας, συνεπώς απαιτούν **συστηματική εναρμόνιση (data harmonization)**, αυστηρή προεπεξεργασία, και αναλυτικές προσεγγίσεις ικανές να χειριστούν μη γραμμικότητα, αλληλεπιδράσεις και θόρυβο. Επιπλέον, οι στόχοι της Ιατρικής Ακριβείας στις νευροεκφυλιστικές νόσους είναι τυπικά προγνωστικοί:

- (α) πρόβλεψη μετατροπής/εξέλιξης,
- (β) εκτίμηση χρονικού ορίζοντα επιδείνωσης,
- (γ) αναγνώριση υποομάδων με διακριτό βιολογικό προφίλ, και
- (δ) ανάδειξη βιολογικά ερμηνεύσιμων χαρακτηριστικών (candidate biomarkers) που συνδέονται με έκβαση [7], [10].

Επιπρόσθετα, η πολυτροπική προσέγγιση επιτρέπει τη σύνδεση της «βιολογίας» με τη «ζωή»: οι ψηφιακοί βιοδείκτες μπορούν να λειτουργήσουν ως συνεχής, οικολογικά έγκυρη παρατήρηση της καθημερινής λειτουργικότητας, συμπληρώνοντας τις περιοδικές κλινικές μετρήσεις και προσφέροντας πλούσια χρονική πληροφορία που ενισχύει την ευαισθησία πρόγνωσης [11], [12]. Έτσι, η Ιατρική Ακριβείας δεν περιορίζεται στην επιλογή θεραπείας, αλλά επεκτείνεται στη συνεχή παρακολούθηση κινδύνου, στην έγκαιρη ανίχνευση μεταβολών και στη δυναμική προσαρμογή στρατηγικών πρόληψης/παρέμβασης.

Συνολικά, η μετάβαση στην Ιατρική Ακριβείας στις νευροεκφυλιστικές νόσους προϋποθέτει ένα ολοκληρωμένο αναλυτικό πλαίσιο που: (i) ενώνει κλινικά, μοριακά, γενετικά και ψηφιακά δεδομένα, (ii) υποστηρίζει προγνωστική μοντελοποίηση με αυστηρή επικύρωση, και (iii) παράγει αποτελέσματα ερμηνεύσιμα και αξιοποιήσιμα για κλινική λήψη αποφάσεων. Αυτή η προσέγγιση είναι απολύτως εναρμονισμένη με την τρέχουσα κατεύθυνση της βιοδείκτη-κεντρικής θεώρησης της AD, καθώς και με τη ραγδαία ανάπτυξη ψηφιακών μεθοδολογιών παρακολούθησης και screening [7], [8], [11].

1.2 Το Πλαίσιο Υλοποίησης: Συνέργεια και Υποδομές

Η υλοποίηση του Brain Precision προϋποθέτει ένα συνεκτικό πλαίσιο **εθνικής συνέργειας, τυποποιημένων ψηφιακών υποδομών και διακυβέρνησης δεδομένων** που να επιτρέπει την αξιόπιστη συλλογή, ενοποίηση, ανάλυση και αξιοποίηση βιοϊατρικών πληροφοριών από δεδομένα πολλαπλών τύπων για τις νευροεκφυλιστικές νόσους, με έμφαση στις Alzheimer και Parkinson. Το έργο οργανώνεται ως εθνικό δίκτυο με στόχο τη γενετική/γενομική διερεύνηση, την ανίχνευση βιοδεικτών και την ανάπτυξη καινοτόμων υπολογιστικών τεχνολογιών στη βάση της Ιατρικής Ακριβείας [13], [14]. Η αρχιτεκτονική υλοποίησης υπηρετεί δύο αλληλένδετους άξονες: (α) την ομογενοποιημένη παραγωγή και διαχείριση δεδομένων υψηλής ποιότητας σε πολλαπλούς κόμβους και (β) την αναλυτική αξιοποίησή τους με βιοπληροφορική και τεχνικές μηχανικής μάθησης για προγνωστική μοντελοποίηση και ανακάλυψη μοριακών δεικτών [13].

1.2.1 Το Εθνικό Δίκτυο Ιατρικής Ακριβείας «Brain Precision»: Στόχοι για γονιδιακή χαρτογράφηση και ανάδειξη βιοδεικτών στον ελληνικό πληθυσμό

Το Brain Precision θεμελιώνεται στην έννοια του «εθνικού δικτύου» ως μηχανισμού που συνδέει κλινικούς/ερευνητικούς φορείς, εργαστηριακές και υπολογιστικές υποδομές, ώστε να επιτευχθεί **κλίμακα, αναπαραγωγιμότητα και διαλειτουργικότητα** στην παραγωγή βιοϊατρικών αποδεικτικών στοιχείων. Στο πλαίσιο αυτό, οι δηλωμένοι στόχοι του δικτύου περιλαμβάνουν: (i) ανάδειξη της γενετικής βάσης των νευροεκφυλιστικών νοσημάτων, (ii) εντοπισμό αξιόπιστων βιοδεικτών, και (iii) ανάπτυξη υπολογιστικών τεχνολογιών/στρατηγικών Ιατρικής Ακριβείας, με ειδική στόχευση στον ελληνικό πληθυσμό [13], [14]. Η έμφαση στον ελληνικό πληθυσμό είναι κρίσιμη, διότι η γενετική αρχιτεκτονική κινδύνου και οι μοριακές υπογραφές μπορεί να παρουσιάζουν διαφοροποιήσεις ανάλογα με πληθυσμιακές δομές, περιβαλλοντικές εκθέσεις και κλινικά μοτίβα, γεγονός που επηρεάζει τόσο την εξωτερική εγκυρότητα όσο και τη μεταφερσιμότητα διεθνών προγνωστικών μοντέλων.

Σε επίπεδο υποδομών, το δίκτυο απαιτεί ολοκληρωμένη στοίβα τεχνολογιών που να καλύπτει:

- (α) κλινική και εργαστηριακή ροή δεδομένων,
- (β) ενοποιημένη αποθήκευση/καταλογογράφηση,
- (γ) αναλυτική επεξεργασία μεγάλης κλίμακας, και
- (δ) ασφαλή πρόσβαση/διαμοιρασμό.

Η συμμόρφωση με κανονιστικά και δεοντολογικά πλαίσια αποτελεί δομική απαίτηση. Ειδικά σε δεδομένα υγείας και γενετικά/γενομικά δεδομένα, εφαρμόζονται αρχές προστασίας προσωπικών δεδομένων, όπως αυτές ορίζονται στον ΓΚΠΔ (GDPR) [15], καθώς και σύγχρονα ευρωπαϊκά θεσμικά πλαίσια ανταλλαγής/δευτερογενούς χρήσης δεδομένων υγείας, όπως ο Ευρωπαϊκός Χώρος Δεδομένων Υγείας (EHDS) [16]. Σε οργανωτικό επίπεδο, η πρακτική υλοποίηση προϋποθέτει πολιτικές ελαχιστοποίησης/σκοπού, ψευδωνυμοποίησης, ελεγχόμενης πρόσβασης,

ιγνηλασιμότητας και διαχείρισης κινδύνου ασφάλειας, σύμφωνα με διεθνείς προσεγγίσεις συστημάτων διαχείρισης ασφάλειας πληροφοριών (π.χ. ISO/IEC 27001) [17].

Παράλληλα, η διαχείριση των δεδομένων πρέπει να σχεδιαστεί ώστε να υποστηρίζει τη μηχανική αξιοποίηση και επαναχρησιμοποίηση. Η υιοθέτηση των αρχών FAIR (Findable, Accessible, Interoperable, Reusable) λειτουργεί ως κατευθυντήρια γραμμή για καταλογογράφηση, τεκμηρίωση μεταδεδομένων, και τυποποίηση εννοιών, ώστε τα δεδομένα να είναι λειτουργικά αξιοποιήσιμα από αναλυτικά pipelines και πλατφόρμες υποστήριξης αποφάσεων [18]. Συνολικά, το Brain Precision ως εθνικό δίκτυο επιδιώκει να μετασχηματίσει την ερευνητική και κλινική πληροφορία σε δομημένη υποδομή γνώσης, με δυνατότητα κλιμάκωσης και επικύρωσης σε πραγματικές συνθήκες.

1.2.2 Ενοποίηση Δεδομένων: Συνδυασμός κλινικών φαινοτύπων με προηγμένες γονιδιωματικές αναλύσεις

Η ενοποίηση δεδομένων αποτελεί τον λειτουργικό πυρήνα της Ιατρικής Ακρίβειας στις νευροεκφυλιστικές νόσους και αποσκοπεί στη γεφύρωση του χάσματος μεταξύ **κλινικού φαινοτύπου** (συμπτώματα, κλινική πορεία, νευροψυχολογική επίδοση, συννοσηρότητες) και **μοριακού υποστρώματος** (γενετικοί παράγοντες κινδύνου, γονιδιωματικές/ομικές υπογραφές, βιοχημικοί δείκτες). Η προσέγγιση αυτή είναι απαραίτητη διότι οι νευροεκφυλιστικές νόσοι είναι ετερογενείς και πολυπαραγοντικές: διαφορετικοί βιολογικοί μηχανισμοί μπορούν να καταλήγουν σε παρόμοιους κλινικούς φαινοτύπους, ενώ το ίδιο κλινικό στάδιο μπορεί να αντιστοιχεί σε διαφορετικές μοριακές καταστάσεις. Κατά συνέπεια, η προγνωστική μοντελοποίηση και η ανακάλυψη βιοδεικτών απαιτούν ενοποιημένες, πολυτροπικές αναπαραστάσεις που επιτρέπουν την ανάλυση αλληλεπιδράσεων, μη γραμμικοτήτων και υπο-ομάδων ασθενών (subphenotypes).

Σε τεχνικό επίπεδο, η ενοποίηση υλοποιείται ως διαδικασία (ETL/ELT) που μετασχηματίζει ετερογενείς ροές σε ένα συνεκτικό **κοινό σημασιολογικό και συντακτικό μοντέλο**, ώστε να είναι δυνατή η συστηματική ανάλυση. Για την κλινική πληροφορία, η διαλειτουργικότητα μπορεί να υποστηριχθεί με διεθνή πρότυπα ανταλλαγής δεδομένων υγείας, όπως το HL7 FHIR, το οποίο οργανώνει την πληροφορία σε «resources» και επιτρέπει συνεπή ανταλλαγή κλινικών οντοτήτων μεταξύ συστημάτων [19]. Για δευτερογενή ανάλυση παρατηρησιακών δεδομένων μεγάλης κλίμακας, πρότυπα κοινού μοντέλου δεδομένων όπως το OMOP Common Data Model διευκολύνουν τη δομική εναρμόνιση και την επαναχρησιμοποίηση αναλυτικών ρουτινών σε ομογενοποιημένες βάσεις [20]. Παράλληλα, για γενετικά/γενομικά δεδομένα και κανόνες χρήσης τους, τυποποιήσεις όπως το GA4GH Data Use Ontology (DUO) επιτρέπουν μηχανικά αναγνώσιμη περιγραφή περιορισμών/αδειών χρήσης, διευκολύνοντας υπεύθυνη πρόσβαση και διαμοιρασμό [21], [22].

Η ενοποίηση όμως δεν είναι μόνο τεχνικό ζήτημα. Απαιτεί επίσης **διακυβέρνηση δεδομένων**: ορισμούς μεταβλητών, λεξιλόγια, σαφείς κανόνες ποιότητας, τεκμηριωμένες διαδικασίες επεξεργασίας, και μηχανισμούς λογοδοσίας. Ειδικά σε κλινικά και γονιδιωματικά δεδομένα, η ποιότητα (πληρότητα, ακρίβεια, συνέπεια, χρονική ευθυγράμμιση) καθορίζει την αξιοπιστία τόσο των βιοπληροφορικών αναλύσεων όσο και των μοντέλων μηχανικής μάθησης. Η συστηματική ενοποίηση επιτρέπει: (i) βελτιωμένη ισχύ ανίχνευσης μοριακών συσχετίσεων, (ii) αναγνώριση πολυπαραγοντικών προφίλ κινδύνου, (iii) παραγωγή ερμηνεύσιμων υπογραφών για κλινική χρήση, και (iv) τελικά ανάπτυξη και επικύρωση προγνωστικών μοντέλων που είναι αναπαραγώγιμα και μεταφέρσιμα σε πολλαπλά κέντρα.

Συνοψίζοντας, το πλαίσιο υλοποίησης του Brain Precision εδράζεται σε εθνική συνέργεια και σε σύγχρονες υποδομές διαλειτουργικότητας/ασφάλειας, ενώ η ενοποίηση κλινικών φαινοτύπων με προηγμένες γονιδιωματικές αναλύσεις διαμορφώνει το αναγκαίο υπόβαθρο για την ανάδειξη αξιόπιστων βιοδεικτών και την παραγωγή κλινικά αξιοποιήσιμης προγνωστικής γνώσης [13], [14], [18]–[21].

1.3 Ορισμός Προβλήματος και Μεθοδολογικό Πλαίσιο

1.3.1 Διατύπωση του Προβλήματος Μηχανικής Μάθησης

Στο Brain Precision, η αναλυτική πρόκληση ορίζεται ως **πρόβλημα κλινικής πρόγνωσης** (clinical prediction modelling) όπου αξιολογούνται ετερογενή βιοϊατρικά χαρακτηριστικά (κλινικά/δημογραφικά, νευροψυχολογικά, εργαστηριακά/μοριακά και γενετικά/γενομικά) για την εκτίμηση ενός **εξατομικευμένου κινδύνου** με συγκεκριμένο χρονικό ορίζοντα. Η διατύπωση του προβλήματος ακολουθεί τη λογική των σύγχρονων κατευθυντήριων γραμμών για μοντέλα πρόγνωσης που χρησιμοποιούν παλινδρομικές ή/και ML προσεγγίσεις, δίνοντας έμφαση στη διαφάνεια, στην αναπαραγωγικότητα, στον έλεγχο μεροληψίας και στην κλινική αξιοποίηση [23]–[26].

Μοντελοποιητική οντότητα. Θεωρούμε σύνολο παρατηρήσεων $D = \{(x_i, y_i)\}_{i=1}^N$, όπου $x_i \in \mathbb{R}^P$ είναι διάνυσμα χαρακτηριστικών για το άτομο i και y_i είναι η έκβαση (ορισμένη παρακάτω ως δυαδική). Στόχος είναι η εκμάθηση συνάρτησης $f(\cdot)$ που προσεγγίζει είτε:

- **πιθανότητα κινδύνου** $\hat{p}_i = P(y_i = 1 | x_i)$ (probabilistic risk prediction), ή/και
- **κανόνα απόφασης** $\hat{y}_i = \mathbb{I}(\hat{p}_i \geq \tau)$ (classification rule με κατώφλι τ).

Η κλινική χρησιμότητα προϋποθέτει ότι το μοντέλο δεν αρκεί να «ταξινομεί» σωστά, αλλά πρέπει να παρέχει (όπου είναι εφικτό) **καλά βαθμονομημένες πιθανότητες** (calibration) και **καλή διακριτική ικανότητα** (discrimination), ώστε το \hat{p}_i είναι αξιόπιστο για λήψη αποφάσεων [27]. Επιπλέον, η ανάπτυξη/αναφορά πρέπει να τεκμηριώνει με σαφήνεια τον ορισμό συμμετεχόντων, χαρακτηριστικών, έκβασης, τον χειρισμό ελλειπόν τιμών, την αποφυγή διαρροής πληροφορίας (data leakage), καθώς και τη διαδικασία εσωτερικής/εξωτερικής επικύρωσης [23], [25].

Μεθοδολογικός έλεγχος ποιότητας και μεροληψίας. Σε προβλήματα πρόγνωσης, η μεροληψία μπορεί να προκύψει από: (i) επιλογή δείγματος και αντιπροσωπευτικότητα, (ii) μέτρηση/χρονισμό προβλεπών, (iii) ορισμό/μέτρηση έκβασης και (iv) αναλυτικές αποφάσεις (π.χ. υπερπροσαρμογή, ακατάλληλη επικύρωση, ατεκμηρίωτη επιλογή χαρακτηριστικών). Η συστηματική εκτίμηση κινδύνου μεροληψίας και εφαρμοσιμότητας (applicability) μπορεί να στηριχθεί στο PROBAST και στις νεότερες επεκτάσεις για AI/ML (PROBAST+AI), οι οποίες δομούν ένα συνεκτικό πλαίσιο αξιολόγησης σε τέσσερις άξονες (participants, predictors, outcome, analysis) [24], [26]. Αυτό είναι κρίσιμο για να διασφαλιστεί ότι η προγνωστική απόδοση δεν είναι τεχνητά διογκωμένη και ότι το μοντέλο είναι πραγματικά μεταφέρσιμο σε κλινικά περιβάλλοντα.

Επιλογή μοντελοποιητικής οικογένειας. Η επιλογή αλγορίθμων (π.χ. λογιστική παλινδρόμηση, SVM, δένδρα/ensembles, boosting) δεν πρέπει να αντιμετωπίζεται ως αυτοσκοπός. Η βιβλιογραφία δείχνει ότι, σε πολλά κλινικά σενάρια δυαδικής πρόγνωσης, η «υπεροχή» ML έναντι κλασικών παλινδρομικών μοντέλων δεν είναι δεδομένη και συχνά εξαρτάται από τη μεθοδολογική αυστηρότητα, την ποιότητα δεδομένων και την ορθή επικύρωση [28]. Συνεπώς, το πλαίσιο Brain Precision ορίζει το πρόβλημα πρωτίστως ως **κλινικό prediction problem** και δευτερευόντως ως

επιλογή αλγορίθμου, με έμφαση σε αναπαραγωγιμότητα, ερμηνευσιμότητα και κλινική χρησιμότητα [23], [27], [28].

1.3.1.1 Πρόβλημα Δυαδικής Ταξινόμησης (Binary Classification)

Στο επίπεδο της τυπικής διατύπωσης, το πρόβλημα καθορίζεται ως **δυαδική ταξινόμηση** όπου η έκβαση γλαμβάνει τιμές στο $\{0, 1\}$ και αντιστοιχεί σε *μη συμβάν* έναντι *συμβάντος* μέσα σε προκαθορισμένο χρονικό ορίζοντα. Στο πλαίσιο της πρόγνωσης νευροεκφυλιστικού κινδύνου, ένα συνεπές δυαδικό endpoint μπορεί να οριστεί ως «εμφάνιση/κλινικά σημαντική έκπτωση» έναντι «μη εμφάνισης/μη έκπτωσης» σε προκαθορισμένη χρονική περίοδο παρακολούθησης, με σαφή κριτήρια βασισμένα σε τυποποιημένες κλινικές/νευροψυχολογικές μετρήσεις. Η εννοιολογική σημασία της δυαδικοποίησης είναι ότι μετατρέπει μια συχνά συνεχόμενη βιολογική/γνωστική τροχιά σε **κλινικά ερμηνεύσιμο σημείο απόφασης** (decision endpoint), διευκολύνοντας τη διαχείριση κινδύνου και τη στρωματοποίηση.

Συνάρτηση στόχου και κανόνας απόφασης. Κατά κανόνα, τα κλινικά μοντέλα δυαδικής πρόγνωσης επιδιώκουν να εκτιμήσουν πιθανότητα κινδύνου \hat{p} . Η απόφαση «θετικό/αρνητικό» προκύπτει μέσω κατωφλίου τ , δηλαδή $\hat{y} = 1$ αν $\hat{p} \geq \tau$. Η επιλογή τ δεν είναι ουδέτερη: εκφράζει κλινική στάθμιση μεταξύ **ευαισθησίας (recall/sensitivity)** και **ειδικότητας**, ανάλογα με το κόστος των ψευδώς αρνητικών (missed high-risk) έναντι ψευδώς θετικών (unnecessary follow-up/intervention). Για νευροεκφυλιστικά σενάρια πρόγνωσης, συχνά προτιμάται προσανατολισμός σε υψηλότερη ευαισθησία όταν η έγκαιρη ανίχνευση θεωρείται κρίσιμη, υπό την προϋπόθεση ότι το downstream κλινικό μονοπάτι μπορεί να απορροφήσει τον αυξημένο αριθμό θετικών προβλέψεων [23], [27].

Αξιολόγηση μοντέλου για δυαδική έκβαση. Η απόδοση οφείλει να αξιολογείται σε τουλάχιστον δύο άξονες:

1. **Διακριτική ικανότητα** (π.χ. ROC-AUC), που αποτυπώνει την ικανότητα διάκρισης συμβάντος/μη συμβάντος ανεξαρτήτως κατωφλίου. Η ROC ανάλυση είναι καθιερωμένο εργαλείο για την οργάνωση και σύγκριση ταξινομητών [29]. Όταν απαιτείται σύγκριση AUC μεταξύ συσχετισμένων ROC καμπυλών (π.χ. ίδια δείγματα, διαφορετικά μοντέλα), η κλασική μη παραμετρική προσέγγιση DeLong χρησιμοποιείται ευρέως [30].
2. **Βαθμονόμηση** (calibration), δηλαδή κατά πόσο οι προβλεπόμενες πιθανότητες ανταποκρίνονται στις παρατηρούμενες συχνότητες κινδύνου. Στα κλινικά prediction models, η βαθμονόμηση είναι προϋπόθεση για τη χρήση του \hat{p} ως «ρίσκου» και όχι απλώς ως score [27]. (Η έμφαση στη βαθμονόμηση είναι επίσης βασική σε πλαίσια TRIPOD/TRIPOD+AI και στην αποτίμηση μεροληψίας PROBAST/PROBAST+AI) [23], [24], [26].

Ανισορροπία κλάσεων και κλινική ερμηνεία. Σε πολλά προγνωστικά προβλήματα, το συμβάν (θετική κλάση) είναι σχετικά σπάνιο, δημιουργώντας **class imbalance**. Σε τέτοιες περιπτώσεις, μετρικές όπως απλό accuracy μπορεί να είναι παραπλανητικές, και η επιλογή μετρικών (π.χ. balanced accuracy, F1-score, precision/recall) και στρατηγικών εκπαίδευσης (π.χ. στάθμιση κλάσεων, υποδειγματοληψία, συνθετική υπερδειγματοληψία) γίνεται κρίσιμη. Η τεχνική SMOTE αποτελεί κλασική προσέγγιση δημιουργίας συνθετικών παραδειγμάτων της μειονοτικής κλάσης για βελτίωση της ευαισθησίας, όταν εφαρμόζεται ορθά εντός του training loop ώστε να

αποφεύγεται διαρροή πληροφορίας [31]. Η μεθοδολογική αρχή εδώ είναι ότι η αντιμετώπιση ανισορροπίας δεν πρέπει να αλλοιώνει την κλινική ερμηνεία ή να οδηγεί σε υπερπροσαρμογή— και πρέπει να τεκμηριώνεται πλήρως, σύμφωνα με σύγχρονες οδηγίες αναφοράς και αξιολόγησης [23], [24], [26].

Συνοψίζοντας, η δυαδική ταξινόμηση στο Brain Precision αποτελεί τυπική αλλά αυστηρά κλινικά προσανατολισμένη διατύπωση πρόγνωσης: (i) εκμάθηση πιθανότητας κινδύνου, (ii) μετατροπή σε απόφαση με κατώφλι προσαρμοσμένο στο κλινικό κόστος σφαλμάτων, (iii) αξιολόγηση με έμφαση σε διακριτική ικανότητα και βαθμονόμηση, και (iv) συστηματικός έλεγχος μεροληψίας/αναφοράς σύμφωνα με TRIPOD+AI και PROBAST+AI [23], [24], [26], [27].

1.3.1.2 Στόχος: Πρόβλεψη Κινδύνου Γνωστικής Έκπτωσης σε βάθος 24 μηνών

Ο επιχειρησιακός στόχος του Brain Precision, στο επίπεδο της προγνωστικής μοντελοποίησης, είναι η **εκ των προτέρων (baseline) εκτίμηση του κινδύνου εμφάνισης γνωστικής έκπτωσης** σε συγκεκριμένο χρονικό ορίζοντα **24 μηνών**. Η επιλογή του ορίζοντα των 24 μηνών τεκμηριώνεται ως ένα πρακτικά και κλινικά ισορροπημένο διάστημα: είναι επαρκές ώστε να καταγραφούν **ουσιαστικές μεταβολές** στη συνολική νευροψυχολογική επίδοση, ενώ ταυτόχρονα παραμένει συμβατό με την έννοια της **έγκαιρης παρέμβασης** (δηλαδή πριν παγιωθούν μη αναστρέψιμες βλάβες ή πριν χαθεί το “παράθυρο” πρόληψης/μετριασμού).

Μεθοδολογικά, ο στόχος διατυπώνεται ως **προσομοίωση προσυμπτωματικού ελέγχου (screening)**: το μοντέλο «βλέπει» αποκλειστικά **πληροφορία διαθέσιμη κατά την αρχική αξιολόγηση (baseline)** και καλείται να προβλέψει αν ο συμμετέχων θα καταταχθεί σε κατηγορία αυξημένου κινδύνου γνωστικής έκπτωσης στα 2 έτη.

Η συγκεκριμένη διατύπωση έχει τρία κρίσιμα πλεονεκτήματα:

1. **Κλινικός ρεαλισμός**: ευθυγραμμίζεται με τον τρόπο που λαμβάνονται αποφάσεις σε κλινικά περιβάλλοντα, όπου η εκτίμηση κινδύνου γίνεται κατά την πρώτη επίσκεψη και καθοδηγεί ένταση παρακολούθησης ή/και εξατομίκευση παρεμβάσεων.
2. **Αποφυγή χρονικής διαρροής πληροφορίας (temporal leakage)**: η χρήση αποκλειστικά baseline προβλεπτών διασφαλίζει ότι ο αλγόριθμος δεν αξιοποιεί (άμεσα ή έμμεσα) πληροφορία που προκύπτει μετά την έναρξη της παρακολούθησης και σχετίζεται με την έκβαση.
3. **Ενίσχυση μεταφερσιμότητας**: το baseline-only σενάριο μειώνει την εξάρτηση από δεδομένα υψηλής συχνότητας ή δύσκολα κλιμακώσιμα, καθιστώντας πιο πιθανή την εφαρμογή του μοντέλου σε πολυκεντρικές ροές δεδομένων.

Η στόχευση στα 24 μήνες συνδέεται επίσης με τη δομή της αξιολόγησης: η γνωστική μεταβολή ορίζεται ως διαφορά μεταξύ baseline και 2-ετούς επανεξέτασης, και η ανάλυση περιορίζεται σε άτομα με πλήρη αξιολόγηση και στα δύο χρονικά σημεία, ώστε να εξασφαλιστεί σαφής και μη αμφίσημος ορισμός έκβασης.

Ενδεικτικά, σε σχετική διαδικασία επικύρωσης/ελέγχου γενίκευσης χρησιμοποιούνται μόνο συμμετέχοντες με πλήρες NTB baseline και 2-year follow-up (π.χ. N≈145), γεγονός που μειώνει τον θόρυβο από ετερογενείς/ελλιπείς μετρήσεις αλλά επιβάλλει αυστηρή τεκμηρίωση διαχείρισης απωλειών/ελλιπών δεδομένων .

Συνολικά, ο στόχος των 24 μηνών στο Brain Precision υπηρετεί μια σαφώς ορισμένη κλινική χρήση: **έγκαιρη ανίχνευση ατόμων υψηλού κινδύνου**, ώστε να κατευθυνθούν σε ενισχυμένη παρακολούθηση ή/και στοχευμένες παρεμβάσεις, με χρήση αναλυτικών μοντέλων που βασίζονται σε δεδομένα αρχικής αξιολόγησης.

1.3.2 Ορισμός Μεταβλητής-Στόχου (Target Variable): Μεταβολή της συνολικής νευροψυχολογικής επίδοσης (NTB total score)

Η μεταβλητή-στόχος ορίζεται με βάση τη **μεταβολή της συνολικής νευροψυχολογικής επίδοσης** όπως αποτυπώνεται από το **NTB total score** (Neuropsychological Test Battery – συνολικό σκορ). Η κεντρική ιδέα είναι ότι η γνωστική έκπτωση δεν προσεγγίζεται ως στατική κατάσταση, αλλά ως **μεταβολή (change score)** σε προκαθορισμένο χρονικό παράθυρο. Στο Brain Precision, η μεταβολή συνδέεται άμεσα με τον ορίζοντα των 2 ετών, δηλαδή με μετρήσεις baseline και 24μηνου follow-up.

(α) Ορισμός μεταβολής (ΔNTB)

Ο ποσοτικός ορισμός της γνωστικής αλλαγής δίνεται από:

$$\Delta NTB = NTB_{24m} - NTB_{bl}, \quad (1)$$

όπου NTB_{bl} είναι το συνολικό NTB στο baseline και NTB_{24m} στο follow-up των 24 μηνών. Σε τεχνικές ροές που χρησιμοποιούν τυποποιημένες κλίμακες, η μεταβολή μπορεί να εκφράζεται ως διαφορά **z-scores** συνολικής γνωστικής επίδοσης μεταξύ baseline και 2-ετούς follow-up [32].

Επιπλέον, όταν απαιτείται «ετήσια» ερμηνεία της μεταβολής, μπορεί να χρησιμοποιηθεί ετησιοποιημένη μορφή:

$$\Delta NTB_{annual} = \frac{NTB_{24m} - NTB_{bl}}{2}. \quad (2)$$

Η υλοποίηση στο έργο περιγράφεται ως **κωδικοποίηση της ετήσιας μεταβολής** του NTB total score, και στη συνέχεια μετατροπή της σε δυαδική κατηγορία κινδύνου.

(β) Δυαδική κωδικοποίηση: χαμηλός vs υψηλός κίνδυνος

Για να μετατραπεί η συνεχής μεταβολή σε επιχειρησιακό endpoint δυαδικής πρόγνωσης, η μεταβλητή-στόχος κωδικοποιείται ως:

- **Υψηλός κίνδυνος:** μη θετική ετήσια μεταβολή στη συνολική νευροψυχολογική επίδοση (δηλαδή $\Delta NTB_{annual} \leq 0$, ή ισοδύναμα $\Delta NTB \leq 0$ όταν η ετησιοποίηση δεν αλλάζει το πρόσημο).
- **Χαμηλός κίνδυνος:** θετική μεταβολή (βελτίωση/μη έκπτωση) της συνολικής επίδοσης ($\Delta NTB_{annual} > 0$).

Αυτός ο ορισμός (μη θετική μεταβολή ως κριτήριο αυξημένου κινδύνου) καταγράφεται ρητά ως βάση για την κατάταξη των συμμετεχόντων σε χαμηλό/υψηλό κίνδυνο.

Σημαντική συνέπεια του συγκεκριμένου ορισμού είναι ότι το μοντέλο δεν «κυνηγά» απαραίτητα μια μεγάλη πτώση, αλλά επιχειρεί να διακρίνει αυτούς που **δεν εμφανίζουν καθαρή θετική τροχιά** στη συνολική γνωστική επίδοση εντός 24 μηνών. Αυτό είναι συμβατό με σενάρια πρώιμης

πρόγνωσης, όπου οι μεταβολές μπορεί να είναι ήπιες και επηρεάζονται από θόρυβο, φαινόμενα εξοικείωσης (practice effects) και διακυμάνσεις μέτρησης.

(γ) Αναλυτικές επιπτώσεις του ορισμού της έκβασης

Ο ορισμός της ΔΝΤΒ ως target έχει συγκεκριμένες στατιστικές/μεθοδολογικές προεκτάσεις:

1. **Ανάγκη πλήρους παρακολούθησης:** Η αξιόπιστη εκτίμηση ΔΝΤΒ απαιτεί πλήρη NTB αξιολόγηση τόσο στο baseline όσο και στα 24 μήνες· για αυτό και, σε διαδικασίες αξιολόγησης, χρησιμοποιούνται μόνο πλήρεις περιπτώσεις (π.χ. $N \approx 145$).
2. **Έλεγχος συγχυτών baseline:** Επειδή το baseline NTB μπορεί να επηρεάζει το περιθώριο μεταβολής (ceiling/floor effects), είναι εύλογο να εξετάζεται η σχέση «πρόβλεψη κινδύνου → ΔΝΤΒ» και με προσαρμογή ως προς baseline NTB, π.χ. μέσω γραμμικού μοντέλου/OLS με robust standard errors.
3. **Μέτρια ανισορροπία κλάσεων:** Η δυαδικοποίηση οδηγεί σε κλάσεις που μπορεί να μην είναι ισομεγέθεις (ενδεικτικά ~66% χαμηλού και ~34% υψηλού κινδύνου), γεγονός που πρέπει να λαμβάνεται υπόψη σε εκπαίδευση/αξιολόγηση (επιλογή μετρικών, χειρισμός ανισορροπίας).

Συμπερασματικά, η μεταβλητή-στόχος στο Brain Precision είναι σαφώς ορισμένη, κλινικά ερμηνεύσιμη και άμεσα συνδεδεμένη με το σενάριο προσυμπτωματικού screening: **baseline δεδομένα → πρόβλεψη δυαδικής έκβασης κινδύνου**, όπου η έκβαση προκύπτει από τη **μεταβολή του NTB total score** μεταξύ baseline και 24 μηνών, με κριτήριο την **μη θετική** ετήσια μεταβολή ως ένδειξη αυξημένου κινδύνου.

1.3.3 Σενάριο Εφαρμογής: Προσομοίωση Προσυμπτωματικού Ελέγχου (Screening) με δεδομένα αναφοράς (Baseline)

Το σενάριο εφαρμογής του Brain Precision ορίζεται ως **προσομοίωση προσυμπτωματικού ελέγχου (screening)**, όπου η πρόβλεψη κινδύνου παράγεται **αποκλειστικά από δεδομένα αναφοράς (baseline)**, δηλαδή από πληροφορία που είναι διαθέσιμη στην αρχική κλινική/ερευνητική επαφή με το άτομο. Το σενάριο αυτό αντιστοιχεί στο πλέον ρεαλιστικό και κλινικά χρήσιμο πρότυπο ανάπτυξης προγνωστικών μοντέλων: μια **εξατομικευμένη εκτίμηση κινδύνου “στην είσοδο”** που μπορεί να καθοδηγήσει την ένταση παρακολούθησης, τη διενέργεια επιπλέον εξετάσεων, ή/και την ενεργοποίηση προληπτικών/παρεμβατικών στρατηγικών [31], [32].

(α) Χρονική αιτιότητα και αποφυγή διαρροής πληροφορίας

Η χρήση baseline δεδομένων ως μοναδικής εισόδου επιβάλλεται από την ανάγκη **χρονικής αιτιότητας (temporal ordering)**: οι προβλεπτικοί παράγοντες πρέπει να προηγούνται χρονικά της έκβασης. Αυτό μειώνει τον κίνδυνο **data leakage**, δηλαδή της άμεσης ή έμμεσης εισαγωγής πληροφορίας που σχετίζεται με την μελλοντική έκβαση (π.χ. μεταβλητές που μετρούνται μετά το baseline ή προκύπτουν ως συνέπεια πρώιμων συμπτωμάτων) [33]. Η αυστηρή διατήρηση της χρονικής διάταξης (baseline → follow-up outcome) αποτελεί βασικό κριτήριο αξιοπιστίας σε κλινικά prediction models και συνδέεται άμεσα με τη δυνατότητα πραγματικής κλινικής εφαρμογής [31].

(β) Επιχειρησιακή ροή χρήσης (operational workflow)

Σε επίπεδο επιχειρησιακής ροής, το screening σενάριο προϋποθέτει:

1. **Καθορισμό πληθυσμού-στόχου:** άτομα χωρίς διάγνωση άνοιας/σοβαρής έκπτωσης στο baseline, αλλά με επαρκές εύρος κινδύνου ώστε να έχει νόημα η πρόγνωση (π.χ. άτομα μέσης/μεγαλύτερης ηλικίας, ή με γνωστικούς/αγγειακούς/γενετικούς παράγοντες κινδύνου) [31].
2. **Καταγραφή baseline χαρακτηριστικών:** κλινικές/δημογραφικές πληροφορίες, νευροψυχολογική επίδοση, και—όπου υπάρχουν—γενετικού/γενομικοί και μοριακοί δείκτες που είναι διαθέσιμοι κατά την αρχική εκτίμηση.
3. **Παραγωγή πιθανότητας κινδύνου $\hat{p} = P(y = 1 | x_{baseline})$** και αντιστοίχιση σε κατηγορία (π.χ. υψηλού/χαμηλού κινδύνου) με κατώφλι που ορίζεται σύμφωνα με το κλινικό κόστος ψευδώς αρνητικών/θετικών [31], [34].
4. **Δράση downstream:** το αποτέλεσμα δεν είναι αυτοσκοπός· εντάσσεται σε αλγόριθμο φροντίδας, όπου υψηλός κίνδυνος ενεργοποιεί πιο στενή παρακολούθηση, πρόσθετη βιοδεικτική διερεύνηση ή/και εξατομικευμένες προληπτικές παρεμβάσεις [32].

(γ) Επικύρωση με ορθές διαδικασίες δειγματοληψίας

Η προσομοίωση screening απαιτεί το σχήμα αξιολόγησης να αναπαράγει τη μελλοντική χρήση. Πρακτικά αυτό σημαίνει ότι κάθε βήμα προεπεξεργασίας (π.χ. κανονικοποίηση, επιλογή χαρακτηριστικών, αντιμετώπιση ανισορροπίας κλάσεων, imputation) πρέπει να εκτελείται **εντός των training folds** και να εφαρμόζεται στο test fold χωρίς να «βλέπει» πληροφορία του test fold κατά την εκπαίδευση [33]. Η τήρηση αυτής της αρχής είναι θεμελιώδης για αποφυγή αισιόδοξης μεροληψίας απόδοσης (optimism bias) και για αξιόπιστη εκτίμηση γενίκευσης [31], [35].

(δ) Κλινική χρησιμότητα πέρα από την ακρίβεια

Για screening εφαρμογές, η “καλή απόδοση” δεν ισοδυναμεί με υψηλό accuracy. Απαιτείται:

- **καλή διακριτική ικανότητα** (ικανότητα διάκρισης υψηλού/χαμηλού κινδύνου),
- **καλή βαθμονόμηση** (οι προβλεπόμενοι κίνδυνοι να αντανακλούν πραγματικές πιθανότητες),
- και αποτίμηση της **κλινικής ωφέλειας** σε σχέση με το κατώφλι απόφασης (π.χ. μέσω decision curve analysis), ώστε να τεκμηριώνεται ότι η χρήση του μοντέλου προσθέτει καθαρό όφελος έναντι στρατηγικών “treat-all” ή “treat-none” [34], [36].

1.3.4 Πρότυπα Αναφοράς και Διαφάνεια: Συμμόρφωση με τις κατευθυντήριες γραμμές TRIPOD+AI για την ανάπτυξη και αναφορά κλινικών μοντέλων πρόβλεψης

Η συμμόρφωση με τα πρότυπα TRIPOD+AI αποτελεί κρίσιμο στοιχείο επιστημονικής ποιότητας, καθώς τα μοντέλα πρόβλεψης στον χώρο της υγείας επηρεάζονται συχνά από ανεπαρκή αναφορά, αδιαφάνεια στην αναλυτική διαδικασία, υπερπροσαρμογή και ανεπαρκή επικύρωση—παράγοντες που οδηγούν σε μη αναπαραγώγιμα ή μη μεταφέριμα αποτελέσματα [31], [35]. Οι οδηγίες **TRIPOD+AI** επεκτείνουν το TRIPOD, προσαρμόζοντας τις απαιτήσεις αναφοράς στα μοντέλα που αναπτύσσονται με τεχνικές μηχανικής μάθησης, και απαιτούν ρητή, λεπτομερή τεκμηρίωση σε όλα τα στάδια του κύκλου ζωής του μοντέλου (development, validation, reporting) [35], [37].

(α) Ελάχιστα στοιχεία που πρέπει να τεκμηριώνονται (TRIPOD+AI ευθυγράμμιση)

Στο πλαίσιο TRIPOD+AI, η αναφορά ενός προγνωστικού μοντέλου πρέπει να περιλαμβάνει, κατ' ελάχιστον:

1. **Πληθυσμός και πλαίσιο δειγματοληψίας:** κριτήρια ένταξης/αποκλεισμού, ρυθμίσεις/κέντρα, χρονικά παράθυρα, χειρισμός follow-up και απωλειών [35].
2. **Ακριβής ορισμός έκβασης:** πότε/πώς μετρήθηκε, ποια κριτήρια χρησιμοποιήθηκαν, και πώς διασφαλίζεται ότι η έκβαση είναι ανεξάρτητη από τους προβλεπτικούς παράγοντες (avoid incorporation bias) [35], [38].
3. **Ορισμός και χειρισμός προβλεπτών:** χρονισμός, προεπεξεργασία, μετασχηματισμοί, κανονικοποίηση, και κυρίως διασφάλιση ότι κάθε βήμα εφαρμόζεται σωστά στο train/test (no leakage) [33], [35].
4. **Ελλιπή δεδομένα:** μηχανισμός ελλείψεων, μέθοδος αντιμετώπισης (π.χ. multiple imputation), και αν η διαδικασία ενσωματώνεται σωστά στο σχήμα επικύρωσης [39].
5. **Αρχιτεκτονική/υπερπαραμέτροι ML:** επιλογή αλγορίθμου, search strategy, cross-validation design, regularization, class weighting ή resampling, και κριτήρια επιλογής τελικού μοντέλου [35].
6. **Μετρικές απόδοσης:** διάκριση (ROC-AUC), βαθμονόμηση (calibration intercept/slope, calibration plots, Brier score), και κλινική χρησιμότητα (π.χ. decision curves), με διαστήματα εμπιστοσύνης όπου είναι εφικτό [34], [35], [36].
7. **Εσωτερική/εξωτερική επικύρωση:** σαφής διάκριση internal validation (CV/bootstrapping) έναντι external validation και περιγραφή της γενικευσιμότητας [35], [40].
8. **Ερμηνευσιμότητα και κλινική ενσωμάτωση:** τεκμηρίωση του πώς θα χρησιμοποιηθεί το μοντέλο, ποια είναι τα προτεινόμενα κατώφλια και ποιο το προβλεπόμενο impact pathway [35], [36].

(β) Έλεγχος μεροληψίας και εφαρμοσιμότητας ως συμπλήρωμα διαφάνειας

Η διαφάνεια στην αναφορά δεν αρκεί χωρίς αξιολόγηση μεροληψίας. Η πρακτική ευθυγράμμιση με TRIPOD+AI συμπληρώνεται ορθά με εργαλεία όπως PROBAST/PROBAST+AI, τα οποία οργανώνουν τον έλεγχο σε συμμετέχοντες-προβλεπτές-έκβαση-ανάλυση, και εντοπίζουν κλασικά failure modes: αισιόδοξη απόδοση, υπερπροσαρμογή, ακατάλληλη επιλογή χαρακτηριστικών, κακή βαθμονόμηση ή ανεπαρκή reporting των pipeline βημάτων [38], [41]. Με αυτόν τον τρόπο, το Brain Precision μπορεί να τεκμηριώνει όχι μόνο «τι πέτυχε» αλλά και «πόσο αξιόπιστα» και «σε ποιους» ισχύει.

(γ) Στόχος: αναπαραγωγιμότητα και μεταφερσιμότητα

Η εφαρμογή TRIPOD+AI υπηρετεί δύο πρακτικές απαιτήσεις:

- **Αναπαραγωγιμότητα:** τρίτοι ερευνητές να μπορούν να αναπαράγουν το pipeline και τα αποτελέσματα με βάση την περιγραφή.
- **Μεταφερσιμότητα:** οι επιδόσεις να μην είναι artefact ενός συγκεκριμένου δείγματος/κέντρου, αλλά να έχουν ρεαλιστική πιθανότητα διατήρησης σε νέους πληθυσμούς, ιδίως όταν ο στόχος είναι κλινική υιοθέτηση [35], [40].

Συνοπτικά, οι ενότητες 1.3.3–1.3.4 ορίζουν το Brain Precision ως προγνωστική εφαρμογή τύπου screening με baseline δεδομένα, και ταυτόχρονα θεμελιώνουν μια αυστηρή μεθοδολογική δέσμευση: ανάπτυξη και αναφορά σύμφωνα με TRIPOD+AI, με έμφαση σε διαφάνεια, αποφυγή διαρροής πληροφορίας, ορθή επικύρωση, βαθμονόμηση, και τεκμηρίωση κλινικής χρησιμότητας [33]–[36], [41].

1.4 Περιγραφή και Προετοιμασία Συνόλου Δεδομένων (Data Curation)

Η ενότητα Data Curation στο Brain Precision αφορά τον συστηματικό κύκλο ζωής του συνόλου δεδομένων που θα χρησιμοποιηθεί για προγνωστική μοντελοποίηση: (α) τεκμηρίωση προέλευσης και δομής, (β) έλεγχο ποιότητας, (γ) εναρμόνιση/τυποποίηση μεταβλητών, (δ) ψευδωνυμοποίηση και διακυβέρνηση πρόσβασης, και (ε) προετοιμασία “analysis-ready” πινάκων/αναπαραστάσεων για στατιστική και μηχανική μάθηση [42]–[44]. Στην παρούσα φάση, το Data Curation επικεντρώνεται σε δεδομένα υψηλής ετερογένειας (κλινικές επισκέψεις, ερωτηματολόγια, ψηφιακά σήματα από φορετές συσκευές/κινητά, και βιοδείκτες), όπου η αναλυτική αξιοπιστία εξαρτάται κρίσιμα από τη σωστή χρονική ευθυγράμμιση των ροών (baseline → follow-up), τη διαχείριση ελλείπων τιμών, και την αποφυγή “διαρροής πληροφορίας” από μεταγενέστερα χρονικά σημεία προς τα baseline χαρακτηριστικά [42], [43].

1.4.1 Προέλευση και Δομή Δεδομένων

Η προέλευση του συνόλου δεδομένων βασίζεται σε **προοπτικό πολυκεντρικό πρωτόκολλο**, όπου η συλλογή πραγματοποιείται με συνδυασμό:

1. **δομημένων κλινικών δεδομένων** κατά τις επισκέψεις (eCRF/clinical visit data),
2. **αυτοαναφερόμενων δεδομένων** (ερωτηματολόγια και ημερολόγια συμπεριφοράς/υγείας), και
3. **παθητικά συλλεγόμενων ψηφιακών δεδομένων** μέσω wearables και mobile εφαρμογών (π.χ. φυσική δραστηριότητα, ύπνος, συνεχείς μετρήσεις) [42], [43].

Η δομή του συνόλου δεδομένων, λόγω της πολυτροπικότητας, δεν είναι μονολιθικός πίνακας αλλά πολυ-επιπέδη αναπαράσταση που συνήθως περιλαμβάνει:

- **Participant-level (στατικό) επίπεδο:** δημογραφία, βασικά κλινικά χαρακτηριστικά, σταθερές μεταβλητές (π.χ. φύλο, έτη εκπαίδευσης, ιστορικό), και ενδεχομένως γενετικοί δείκτες όταν είναι διαθέσιμοι.
- **Visit-level (διαχρονικό) επίπεδο:** επαναλαμβανόμενες κλινικές/νευροψυχολογικές μετρήσεις (baseline, ενδιάμεσες επισκέψεις, follow-up).
- **Time-series επίπεδο:** υψηλής συχνότητας ψηφιακά σήματα (π.χ. ημερήσιες/ωριαίες μετρήσεις δραστηριότητας και ύπνου), τα οποία απαιτούν συμπύκνωση σε χαρακτηριστικά (feature engineering) κατά προ-ορισμένα χρονικά “παράθυρα” για να καταστούν κατάλληλα για ML.
- **Assay/biomarker επίπεδο:** εργαστηριακές μετρήσεις/βιοδείκτες που ακολουθούν δικές τους διαδικασίες ποιοτικού ελέγχου, ορίων ανίχνευσης και κανονικοποίησης [42], [43].

Για τη μετατροπή αυτού του ετερογενούς υλικού σε “analysis-ready” σύνολο, εφαρμόζονται αρχές τεκμηρίωσης και διαλειτουργικότητας που ευθυγραμμίζονται με καθιερωμένες πρακτικές

επιστημονικής διαχείρισης δεδομένων (π.χ. FAIR), ώστε να διασφαλιστεί αναπαραγωγιμότητα και ελεγχόμενη επαναχρησιμοποίηση [44]. Επιπλέον, σε περιβάλλοντα πολυκεντρικής/παρατηρησιακής ανάλυσης, η εναρμόνιση μπορεί να ωφεληθεί από κοινού τύπου μοντέλα δεδομένων (common data models), καθώς μειώνουν την εννοιολογική ασυνέπεια και διευκολύνουν επαναλήψιμα pipelines [45].

1.4.1.1 Πηγή δεδομένων: Το πιλοτικό πρόγραμμα του έργου LETHE

Η πηγή δεδομένων για την παρούσα αναλυτική φάση είναι το πιλοτικό πρόγραμμα του **LETHE**, το οποίο έχει σχεδιαστεί ως πολυκεντρική προοπτική μελέτη σε άτομα μεγαλύτερης ηλικίας με αυξημένο κίνδυνο γνωστικής έκπτωσης, με ψηφιακή υποστήριξη παρέμβασης τύπου FINGER και συστηματική συλλογή δεδομένων πολλαπλών τύπων [42]. Το δείγμα αναφέρεται επιχειρησιακά ως **~150 συμμετέχοντες**, ενώ στη δημοσιευμένη περιγραφή της μελέτης καταγράφεται μέγεθος κοντά σε αυτή την κλίμακα (τυπικά “περίπου 150–160”), γεγονός που είναι συμβατό με τη φύση πιλοτικής/εφικτότητας κλινικής δοκιμής και με στόχους αξιολόγησης συμμόρφωσης, τήρησης παρακολούθησης και προοπτικής συλλογής δεικτών [42].

Κρίσιμα χαρακτηριστικά της πηγής δεδομένων που επηρεάζουν άμεσα το Data Curation είναι τα εξής:

1. **Πολυτροπικότητα με ενσωμάτωση ψηφιακών σημάτων:** Η συλλογή δεν περιορίζεται σε κλινικές επισκέψεις αλλά περιλαμβάνει και παθητικά δεδομένα από φορητές συσκευές (π.χ. δραστηριότητα/ύπνος), καθώς και εφαρμογές κινητών που υποστηρίζουν καταγραφή συμπεριφορών και αλληλεπίδραση με το πρόγραμμα [42], [43]. Αυτό απαιτεί διαδικασίες συγχρονισμού, έλεγχο πληρότητας χρονοσειρών, και μετατροπή ακατέργαστων σημάτων σε σταθερά χαρακτηριστικά ανά άτομο/παράθυρο χρόνου.
2. **Διαχρονικές μετρήσεις με σαφή χρονική αναφορά:** Οι αξιολογήσεις ορίζονται σε διακριτά χρονικά σημεία (baseline και follow-up), με αποτέλεσμα η προετοιμασία δεδομένων να πρέπει να εξασφαλίζει ότι τα χαρακτηριστικά εισόδου προέρχονται αυστηρά από το baseline (ή από προκαθορισμένο baseline παράθυρο), ενώ η έκβαση υπολογίζεται στο αντίστοιχο follow-up [42].
3. **Ετερογένεια τύπων μεταβλητών:** Συνυπάρχουν συνεχείς, κατηγορικές, κλιμακωτές και χρονοσειριακές μεταβλητές, καθώς και δείκτες με διαφορετικά “missingness patterns”. Αυτό επιβάλλει τυποποίηση τύπων, ορισμούς μονάδων/κλιμάκων, και ξεχωριστές στρατηγικές διαχείρισης ελλিপών για κάθε υποσύνολο μεταβλητών, ώστε να μην εισαχθεί συστηματική μεροληψία στην εκπαίδευση των μοντέλων [42], [44].

Συνολικά, το LETHE παρέχει ένα κατάλληλο πιλοτικό περιβάλλον για ανάπτυξη προγνωστικών μοντέλων κινδύνου γνωστικής έκπτωσης, ακριβώς επειδή συνδυάζει κλινική πληροφορία με ψηφιακές ροές που μπορούν να λειτουργήσουν ως συμπληρωματικοί δείκτες λειτουργικής κατάστασης και αλλαγής στον χρόνο [42], [44].

1.4.1.2 Κατηγοριοποίηση μεταβλητών βάσης (Baseline)

Η κατηγοριοποίηση των baseline μεταβλητών αποσκοπεί στη συστηματική οργάνωση των προβλεπτών σε κλινικά ερμηνεύσιμες ομάδες, οι οποίες: (i) αντιστοιχούν σε γνωστούς μηχανισμούς κινδύνου/προστασίας για γνωστική έκπτωση, (ii) υποστηρίζουν συνεπή προεπεξεργασία (τυποποίηση, έλεγχο ποιότητας, χειρισμό ελλείψεων), και (iii) διευκολύνουν την

ερμηνεία των μοντέλων πρόβλεψης, ιδίως όταν οι προβλέψεις προορίζονται για χρήση τύπου screening [45], [46]. Στο Brain Precision, οι baseline μεταβλητές οργανώνονται σε τέσσερις βασικές κατηγορίες: δημογραφικά/κοινωνικοοικονομικά, κλινικοί/βιοχημικοί δείκτες, ιατρικό ιστορικό/συννοσηρότητες, και παράγοντες τρόπου ζωής, με στόχο να αποτυπωθεί τόσο το «υπόβαθρο κινδύνου» όσο και το «τροποποιησιμο προφίλ» του ατόμου [45]–[47].

(α) Δημογραφικά και Κοινωνικοοικονομικά (Ηλικία, Εκπαίδευση)

Ηλικία. Η ηλικία αποτελεί τον ισχυρότερο μη τροποποιήσιμο προγνωστικό παράγοντα για άνοια/γνωστική έκπτωση, καθώς αντανακλά τη σωρευτική έκθεση σε νευροβιολογικές, αγγειακές και μεταβολικές διεργασίες που επιταχύνουν τη νευροεκφύλιση ή/και μειώνουν το γνωστικό απόθεμα [45]. Στο πλαίσιο μοντελοποίησης, η ηλικία λειτουργεί ως βασικός «άξονας βαθμονόμησης» του κινδύνου και ως μεταβλητή που συχνά αλληλεπιδρά με άλλους παράγοντες (π.χ. υπέρταση, διαβήτη, φυσική αδράνεια), επηρεάζοντας τη σχετική συμβολή τους στον συνολικό κίνδυνο [45].

Εκπαίδευση. Τα έτη εκπαίδευσης χρησιμοποιούνται ως λειτουργικός δείκτης **γνωστικού αποθέματος** (cognitive reserve) και κοινωνικοοικονομικής θέσης, και σχετίζονται με διαφοροποιήσεις στην ευαλωτότητα απέναντι σε παθολογικές μεταβολές, καθώς και με τη μακροχρόνια έκθεση σε παράγοντες κινδύνου/προστασίας (π.χ. επαγγελματική πολυπλοκότητα, νοητική δραστηριότητα, υγεία) [45]. Σε πολυτροπικά προγράμματα πρόληψης τύπου FINGER/LETHE, οι δημογραφικοί παράγοντες χρησιμοποιούνται τόσο για στρωματοποίηση όσο και ως απαραίτητες συγχυτικές μεταβλητές (confounders) στην εκτίμηση της ανεξάρτητης συμβολής τροποποιήσιμων παραγόντων [46], [47].

(β) Κλινικοί και Βιοχημικοί Δείκτες (Αρτηριακή πίεση, Λιπίδια, HbA1c)

Η κατηγορία αυτή συγκεντρώνει δείκτες που αποτυπώνουν το **αγγειακό/μεταβολικό φορτίο** (vascular–metabolic burden), το οποίο τεκμηριώνεται ως κεντρικός μηχανισμός που συνεισφέρει στη γνωστική έκπτωση μέσω μικροαγγειακής βλάβης, φλεγμονής, ενδοθηλιακής δυσλειτουργίας και επιταχυνόμενης νευροεκφύλισης [45].

- **Αρτηριακή πίεση (ιδίως υπέρταση στη μέση ηλικία και αθροιστικό φορτίο πίεσης).** Η υπέρταση και γενικότερα η αγγειακή βλάβη συνδέονται με αυξημένο κίνδυνο άνοιας, ενώ η διαχείρισή τους αποτελεί βασική συνιστώσα πρόληψης [45]. Ως baseline μεταβλητή, η αρτηριακή πίεση λειτουργεί τόσο ως άμεσος προβλεπτικός δείκτης όσο και ως μεσολαβητής (mediator) μεταξύ τρόπου ζωής και έκβασης.
- **Λιπίδια (π.χ. LDL-χοληστερόλη).** Η υψηλή LDL χοληστερόλη έχει αναδειχθεί ως τροποποιήσιμος παράγοντας κινδύνου για άνοια, και η ανίχνευση/αντιμετώπισή της προτείνεται ως στρατηγική πρόληψης, ιδίως από τη μέση ηλικία [45]. Για προγνωστική μοντελοποίηση, τα λιπίδια λειτουργούν ως δείκτες μεταβολικού προφίλ και αγγειακής επικινδυνότητας.
- **HbA1c (γλυκαιμικός έλεγχος/μακροχρόνια υπεργλυκαιμία).** Η HbA1c αποτυπώνει τη χρόνια γλυκαιμική επιβάρυνση και αποτελεί κλινικά χρήσιμο δείκτη μεταβολικού κινδύνου. Η συσχέτιση διαβήτη/δυσγλυκαιμίας με άνοια και γνωστική έκπτωση τεκμηριώνει τη χρησιμότητα της HbA1c ως baseline προβλεπτή, ιδιαίτερα σε πληθυσμούς με πολυπαραγοντικό προφίλ κινδύνου [45], [47].

Συνολικά, οι κλινικοί/βιοχημικοί δείκτες έχουν διπλή αξία: παρέχουν «σκληρή» (objective) κλινική πληροφορία με μικρότερο σφάλμα μέτρησης σε σχέση με αυτοαναφορές, και είναι άμεσα στοχεύσιμοι από παρεμβάσεις (φαρμακευτικές/μη), άρα ενισχύουν την κλινική αξιοποιησιμότητα της πρόβλεψης [45]–[47].

(γ) Ιατρικό Ιστορικό και Συννοσηρότητες (Υπέρταση, Διαβήτης, Καρδιαγγειακά)

Το ιατρικό ιστορικό και οι συννοσηρότητες κωδικοποιούν το **σωρευτικό κλινικό φορτίο** και συχνά λειτουργούν ως «συμπυκνωτές» πολλαπλών παθοφυσιολογικών διεργασιών. Στην πρόγνωση γνωστικής έκπτωσης, η παρουσία υπέρτασης, διαβήτη και καρδιαγγειακής νόσου αποτελεί ισχυρό σήμα αυξημένου κινδύνου, καθώς αντανακλά μακροχρόνια αγγειακή/μεταβολική δυσλειτουργία που επιταχύνει εγκεφαλικές μικροαγγειακές αλλοιώσεις, λευκοϋπερεντάσεις και σχετιζόμενη γνωστική έκπτωση [45]. Επιπλέον, σε πολυτροπικές παρεμβάσεις πρόληψης, η ιατρική διαχείριση αγγειακών/μεταβολικών παραγόντων αποτελεί βασικό δομικό στοιχείο, άρα οι σχετικές συννοσηρότητες είναι κεντρικοί baseline προσδιοριστές τόσο κινδύνου όσο και δυναμικού οφέλους από παρέμβαση [46], [47].

(δ) Παράγοντες Τρόπου Ζωής (Υπνος, Άσκηση, Διατροφή)

Οι παράγοντες τρόπου ζωής αποτελούν τον κατεξοχήν **τροποποιήσιμο άξονα κινδύνου** και είναι κρίσιμοι για ένα πλαίσιο screening, διότι συνδέουν άμεσα την πρόβλεψη με πιθανές δράσεις μετριασμού κινδύνου [45]. Στο Brain Precision, οι βασικές διαστάσεις περιλαμβάνουν:

- **Υπνος.** Η ποιότητα/διάρκεια ύπνου και σχετιζόμενες διαταραχές συνδέονται με γνωστική έκπτωση μέσω μηχανισμών όπως διαταραχή γλυμφατικού καθαρισμού, φλεγμονώδεις αποκρίσεις και αλληλεπιδράσεις με μεταβολικούς παράγοντες [45]. Σε ψηφιακά υποστηριζόμενα πρωτόκολλα, ο ύπνος μπορεί να τεκμηριώνεται τόσο με αυτοαναφορά όσο και με wearables, αυξάνοντας την αντικειμενικότητα και τη χρονική ανάλυση [46].
- **Άσκηση/φυσική δραστηριότητα.** Η φυσική αδράνεια έχει τεκμηριωθεί ως παράγοντας κινδύνου για άνοια, ενώ η συστηματική άσκηση υποστηρίζει αγγειακή υγεία, μεταβολική ρύθμιση και νευροπλαστικότητα [45]. Σε προγράμματα τύπου FINGER/LETHE, η φυσική δραστηριότητα αποτελεί βασική παρέμβαση και συχνά παρακολουθείται ψηφιακά (π.χ. βήματα, ενεργός χρόνος, ένταση δραστηριότητας) [46], [47].
- **Διατροφή.** Διατροφικά πρότυπα και η συμμόρφωση σε υγιεινές δίαιτες (π.χ. MIND/μεσογειακού τύπου συστάσεις σε αντίστοιχα πρωτόκολλα) συνδέονται με μεταβολικό/αγγειακό προφίλ και γνωστικά αποτελέσματα [45], [47]. Ως baseline παράγοντας, η διατροφή αποτυπώνει συνήθειες που είναι τροποποιήσιμες και συχνά συσχετίζονται με κοινωνικοοικονομικούς δείκτες (π.χ. εκπαίδευση), άρα απαιτείται προσεκτικός χειρισμός συγχυτών [45].

Η κατηγοριοποίηση των παραγόντων τρόπου ζωής στο baseline δεν εξυπηρετεί μόνο την πρόβλεψη· ενισχύει και την ερμηνεία του «γιατί» ενός υψηλού κινδύνου, επιτρέποντας να συνδεθεί η έξοδος του μοντέλου με συγκεκριμένα μοχλευσιμα πεδία παρέμβασης [45]–[47].

1.4.2 Ορισμός Μεταβλητής-Στόχου (Target Variable)

Η μεταβλητή-στόχος στο πλαίσιο του Brain Precision ορίζεται ως **μεταβολή της συνολικής νευροψυχολογικής επίδοσης** σε προκαθορισμένο χρονικό ορίζοντα (baseline → 24 μήνες). Η

επιλογή ενός **σύνθετου (composite) συνολικού σκορ** τύπου NTB (Neuropsychological Test Battery) είναι μεθοδολογικά τεκμηριωμένη, διότι: (α) αυξάνει την ευαισθησία ανίχνευσης μικρών μεταβολών σε σχέση με μεμονωμένες δοκιμασίες, (β) μειώνει το τυχαίο σφάλμα μέσω συνάθροισης πληροφορίας πολλαπλών γνωστικών πεδίων, και (γ) έχει χρησιμοποιηθεί εκτενώς σε κλινικές δοκιμές και πολυτομεακές παρεμβάσεις πρόληψης, όπου η γνωστική έκβαση αποτυπώνεται ως αλλαγή σε συνολικό NTB z-score ή ισοδύναμο σύνθετο δείκτη [49]–[51].

1.4.2.1 Κωδικοποίηση της ετήσιας μεταβολής της συνολικής νευροψυχολογικής επίδοσης (NTB total score)

(α) Ορισμός μεταβολής (Δ NTB) και ετησιοποίηση

Η βασική ποσότητα ενδιαφέροντος είναι η μεταβολή:

$$\Delta NTB = NTB_{24m} - NTB_{bl}$$

και, για ετησιοποιημένη έκφραση:

$$\Delta NTB_{annual} = \frac{NTB_{24m} - NTB_{bl}}{2}$$

Η ετησιοποίηση επιτρέπει ομοιόμορφη ερμηνεία της αλλαγής ως “ρυθμού” μεταβολής, χρήσιμη τόσο για συγκρισιμότητα μεταξύ υποδειγμάτων όσο και για κλινική επικοινωνία (π.χ. μεταβολή ανά έτος) [49], [52].¹⁰

(β) Τυποποίηση (z-score) και σύνθεση NTB total score

Σε περιβάλλοντα όπου η συνολική γνωστική επίδοση προκύπτει από πολλαπλές δοκιμασίες, το NTB total score συχνά ορίζεται ως **σύνθετος z-score δείκτης**, ώστε οι επιμέρους δοκιμασίες να συνδυάζονται σε κοινή κλίμακα και να ελαχιστοποιείται η επιρροή διαφορετικών μονάδων μέτρησης/διασποράς [49], [50]. Η πρακτική αυτή είναι εδραιωμένη τόσο σε φαρμακολογικές κλινικές δοκιμές όσο και σε πολυδιάστατες παρεμβάσεις πρόληψης, όπου η κύρια έκβαση είναι “change in NTB z-score” [49], [50].

(γ) Έλεγχοι ποιότητας μεταβολής: floor/ceiling, αξιοπιστία επαναληπτικών μετρήσεων και practice effects

Η κωδικοποίηση μεταβολής σε νευροψυχολογικά δεδομένα απαιτεί προσοχή σε τρία σημεία:

1. **Floor/Ceiling effects:** Ακραίες επιδόσεις στο baseline μπορεί να περιορίσουν το “περιθώριο” μεταβολής (π.χ. υψηλό baseline → πιθανότερη στασιμότητα ή μικρή πτώση), επηρεάζοντας τη σχέση προβλεπτών–έκβασης [50].
2. **Αξιοπιστία test–retest:** Η ερμηνεία της μεταβολής εξαρτάται από τη σταθερότητα των μετρήσεων· σύνθετες μπαταρίες όπως το NTB έχουν σχεδιαστεί ώστε να αποτυπώνουν μεταβολές με καλύτερη ευαισθησία συγκριτικά με ορισμένους κλασικούς δείκτες [50], [51].
3. **Practice effects:** Η επαναληπτική εξέταση μπορεί να επιφέρει βελτίωση λόγω εξοικείωσης, η οποία δεν αντανακλά πραγματική γνωστική βελτίωση. Η βιβλιογραφία δείχνει ότι τα practice effects δεν είναι απλό “σφάλμα”, αλλά μπορούν να αλληλεπιδρούν

με την υποκείμενη νευροβιολογική ευαλωτότητα και να επηρεάζουν την πρόγνωση· συνεπώς, η ερμηνεία της ΔNTB πρέπει να γίνεται με γνώση αυτού του φαινομένου και, όπου είναι εφικτό, με κατάλληλες στατιστικές διορθώσεις ή/και ευαισθησιακές αναλύσεις [53], [54].

1.4.2.2 Διαχωρισμός σε κλάσεις Χαμηλού έναντι Υψηλού Κινδύνου

Για να υποστηριχθεί η προγνωστική μοντελοποίηση ως **δυναδικό πρόβλημα ταξινόμησης** (screening-oriented risk stratification), η συνεχής μεταβλητή ΔNTB_{annual} μετασχηματίζεται σε κατηγορική έκβαση δύο κλάσεων. Η λογική αυτού του βήματος είναι επιχειρησιακή: σε εφαρμογές screening, το ζητούμενο συχνά είναι να εντοπιστούν άτομα που απαιτούν ενισχυμένη παρακολούθηση ή περαιτέρω διερεύνηση, αντί να προβλεφθεί ακριβής συνεχής μεταβολή για κάθε άτομο [52], [55].

(α) Επιχειρησιακός ορισμός κλάσεων

Ένας τυπικός και κλινικά ερμηνεύσιμος κανόνας είναι η διάκριση μεταξύ:

- **Χαμηλού κινδύνου:** θετική ή μη επιδεινούμενη τροχιά (π.χ. $\Delta NTB_{annual} > 0$, ανάλογα με τη σύμβαση του composite),
- **Υψηλού κινδύνου:** μη θετική τροχιά/ένδειξη επιδείνωσης (π.χ. $\Delta NTB_{annual} \leq 0$).

Ο κανόνας αυτός αξιοποιεί το πρόσημο της ετησιοποιημένης μεταβολής ως απλό κριτήριο “μη βελτίωσης/επιδείνωσης” σε χρονικό ορίζοντα 24 μηνών. Η χρησιμότητά του είναι ότι παράγει άμεσα ερμηνεύσιμη ταξινόμηση για screening, ενώ επιτρέπει στη συνέχεια την επιλογή κατωφλίου απόφασης (π.χ. πιθανότητα κινδύνου \hat{p}) ανάλογα με την επιθυμητή ευαισθησία/ειδικότητα [52], [55].

(β) Εναλλακτικές στρατηγικές δυαδικοποίησης και τεκμηρίωση επιλογής

Η βιβλιογραφία υποστηρίζει ότι η δυαδικοποίηση μπορεί να οριστεί και με εναλλακτικούς τρόπους, π.χ.:

- **Κριτήριο αξιόπιστης μεταβολής (Reliable Change Index / evidence-based change)**, που λαμβάνει υπόψη την αξιοπιστία test–retest και τα practice effects, μειώνοντας τον κίνδυνο να χαρακτηριστεί ως “έκπτωση” μια μεταβολή εντός τυχαίου σφάλματος [53], [54].
- **Ποσοστημοριακά κατώφλια** (π.χ. χαμηλότερο x% της κατανομής ΔNTB_{annual}), χρήσιμα όταν προέχει η επιχειρησιακή στόχευση ενός “high-risk tail”, αλλά απαιτούν προσοχή ως προς τη μεταφερσιμότητα του κατωφλίου μεταξύ πληθυσμών [52].

Στο Brain Precision, η επιλογή δυαδικού διαχωρισμού θα πρέπει να τεκμηριώνεται με βάση: (i) την κατανομή της ΔNTB_{annual} , (ii) την παρουσία/μέγεθος practice effects, και (iii) την κλινική ερμηνεία του “υψηλού κινδύνου” ως κατηγορίας που ενεργοποιεί συγκεκριμένες downstream ενέργειες. Ειδικά σε datasets τύπου πολυτομεακής πρόληψης, όπου η πρωτεύουσα γνωστική έκβαση είναι NTB composite σε 24 μήνες, η δυαδικοποίηση πρέπει να παραμένει συνεπής με τη δομή του πρωτοκόλλου και την κλινική χρήση της πρόγνωσης [49], [52], [55].

1.4.3 Προ-επεξεργασία (Preprocessing)

Η προ-επεξεργασία στοχεύει στη μετατροπή του αρχικού πολυτροπικού πίνακα baseline σε **αναλυτικά συνεπές και αριθμητικά κατάλληλο** σύνολο εισόδων για μοντέλα πρόβλεψης. Στο Brain Precision, τα δύο πιο κρίσιμα στάδια είναι: (i) η **αριθμητική κωδικοποίηση** των χαρακτηριστικών και (ii) η **διαχείριση ελλিপών τιμών**, με έμφαση στην προστασία από μεροληψία και στη διατήρηση επαρκούς “effective sample size” [56]–[58].

1.4.3.1 Αριθμητική κωδικοποίηση (Numerical Encoding) χαρακτηριστικών

Τα μοντέλα μηχανικής μάθησης (ιδίως τα γραμμικά/μετρικά μοντέλα, SVM, νευρωνικά δίκτυα και boosting) απαιτούν αριθμητικές αναπαραστάσεις. Η κωδικοποίηση εφαρμόζεται ανάλογα με τον τύπο μεταβλητής:

1. **Δυαδικές κατηγορικές μεταβλητές (binary)** Κωδικοποιούνται ως 0/1 (π.χ. παρουσία/απουσία υπέρτασης). Αυτό διατηρεί άμεση ερμηνεία και είναι συμβατό με τις περισσότερες μεθόδους χωρίς πρόσθετο μετασχηματισμό [59].
2. **Ονομαστικές κατηγορικές (nominal) με >2 κατηγορίες** Η τυπική προσέγγιση είναι **one-hot encoding** (δημιουργία ψευδομεταβλητών). Αυτό αποφεύγει τεχνητές διατάξεις που θα εισάγονταν από ακέραιες κωδικοποιήσεις (label encoding), οι οποίες μπορεί να επιβάλουν ανύπαρκτες σχέσεις “μεγαλύτερο/μικρότερο” [59], [60]. Για αποφυγή πολυσυγγραμμικότητας σε γραμμικά μοντέλα, αφαιρείται μία κατηγορία αναφοράς (k-1 dummies).
3. **Διατακτικές/κλιμακωτές μεταβλητές (ordinal / Likert)** Εφόσον η διάταξη έχει ουσιαστικό νόημα (π.χ. επίπεδα εκπαίδευσης ως τάξεις), μπορεί να χρησιμοποιηθεί **ordinal encoding** (π.χ. 0,1,2,...) ή/και monotonic constraints σε ορισμένους αλγορίθμους boosting, υπό την προϋπόθεση ότι η υπόθεση μονοτονίας/διάταξης είναι επιστημονικά εύλογη [59].
4. **Υψηλής καρδιναλιότητας κατηγορικές (high-cardinality)** Σε περιπτώσεις πολλών κατηγοριών (π.χ. λεπτομερής κωδικοποίηση φαρμάκων/διαγνώσεων), το one-hot οδηγεί σε υπερδιάσταση (sparsity). Εναλλακτικές είναι:
 - **target / mean encoding** με αυστηρή εφαρμογή *μόνο εντός training folds* ώστε να αποφευχθεί διαρροή πληροφορίας [61].
 - **learned embeddings** (π.χ. σε νευρωνικά μοντέλα), ιδιαίτερα όταν υπάρχουν επαρκή δεδομένα και σαφής λόγος πολυπλοκότητας [62].
5. **Κλιμάκωση/κανονικοποίηση μετά την κωδικοποίηση** Για αλγορίθμους που είναι ευαίσθητοι στην κλίμακα (π.χ. SVM, λογιστική παλινδρόμηση με κανονικοποίηση, kNN), εφαρμόζεται τυποποίηση (z-score) στα συνεχόμενα χαρακτηριστικά. Η τυποποίηση πρέπει να υπολογίζεται στο training set και να εφαρμόζεται στο test set με τους ίδιους παραμέτρους (mean/std), για αποφυγή leakage [56], [59].

1.4.3.2 Διαχείριση Ελλিপών Τιμών (Missing Values): Στρατηγική εξαίρεσης μεταβλητών με υψηλό ποσοστό απουσίας (π.χ. APOE, Zung score) για διασφάλιση ποιότητας

Η απουσία τιμών είναι αναμενόμενη σε κλινικά και δεδομένα πολλαπλών τύπων και μπορεί να υποβαθμίσει τόσο την ισχύ όσο και την αμεροληψία των μοντέλων. Η επιστημονικά ορθή διαχείριση ξεκινά με τον χαρακτηρισμό του μηχανισμού απουσίας:

- **MCAR** (Missing Completely At Random),
- **MAR** (Missing At Random),
- **MNAR** (Missing Not At Random) [56], [57].

Στην πράξη, όταν συγκεκριμένες μεταβλητές εμφανίζουν **υπερβολικά υψηλή απουσία** (π.χ. ποσοστά που καθιστούν μη αξιόπιστη οποιαδήποτε συμπερασματολογία ή οδηγούν σε μεγάλη μείωση του αποτελεσματικού μεγέθους δείγματος), μια τεκμηριωμένη στρατηγική είναι η **εξαίρεσή τους από το βασικό pipeline**. Αυτό εφαρμόζεται ιδίως όταν συντρέχουν ταυτόχρονα οι παρακάτω συνθήκες:

1. **Κίνδυνος επιδείνωσης “effective sample size”**. Η ένταξη μεταβλητών με πολύ υψηλό missingness μπορεί να οδηγήσει είτε σε σημαντική απώλεια περιπτώσεων (σε complete-case αναλύσεις) είτε σε υπερβολική εξάρτηση από imputation. Και τα δύο αυξάνουν τη διακύμανση και μπορεί να εισάγουν μεροληψία, ιδίως όταν η απουσία σχετίζεται με την έκβαση ή με κρίσιμους συγχυτές [58], [63].
2. **Αβεβαιότητα σχετικά με τον μηχανισμό απουσίας (πιθανό MNAR)** Για ορισμένα πεδία (π.χ. γενετικές πληροφορίες όπως APOE ή ερωτηματολόγια όπως Zung), η απουσία μπορεί να σχετίζεται με πρακτικούς/συστημικούς λόγους (μη διενέργεια εξέτασης, άρνηση συμμετέχοντα, κλινική επιλογή) που συχνά δεν είναι πλήρως παρατηρήσιμοι. Σε τέτοιες περιπτώσεις, η “αυτόματη” συμπλήρωση μπορεί να δημιουργήσει ψευδή σήματα ή να επιδεινώσει μεροληψία [57], [58].
3. **Προτεραιότητα στην ποιότητα και σταθερότητα του βασικού μοντέλου** Σε screening εφαρμογές, η σταθερότητα/αναπαραγωγιμότητα συχνά υπερέχει της οριακής αύξησης απόδοσης από μια μεταβλητή που είναι διαθέσιμη σε μικρό μόνο υποσύνολο. Η εξαίρεση υψηλά ελλιπών μεταβλητών μπορεί να οδηγήσει σε πιο αξιόπιστο και μεταφέρσιμο μοντέλο [56], [58].

Πρακτική υλοποίηση της στρατηγικής εξαίρεσης

- Ορίζεται ένα **κατώφλι απουσίας** (π.χ. >40% ή >50%) με βάση: (i) την κατανομή missingness στο dataset, (ii) την επίδραση στο effective sample size, και (iii) την κλινική/βιολογική σημασία της μεταβλητής. Δεν υπάρχει “καθολικό” ποσοστό· απαιτείται τεκμηριωμένη επιλογή και αναφορά [58], [63].
- Οι μεταβλητές που υπερβαίνουν το κατώφλι (π.χ. APOE, Zung score) **εξαιρούνται από το κύριο pipeline** και, εφόσον κρίνεται χρήσιμο, εντάσσονται σε **δευτερεύουσα ευαισθησιακή ανάλυση** (sensitivity analysis) σε υποσύνολο συμμετεχόντων με διαθέσιμες τιμές, ώστε να αξιολογηθεί η προστιθέμενη προγνωστική αξία χωρίς να υπονομεύεται η κύρια ανάλυση [58], [64].
- Σε μεταβλητές με μέτριο missingness, προτείνεται **multiple imputation** ή άλλες τεχνικές (π.χ. random-forest imputation), πάντοτε εντός της διαδικασίας εκπαίδευσης/επικύρωσης για αποφυγή leakage [58], [65].

Με αυτόν τον τρόπο, η προ-επεξεργασία διατηρεί τον χαρακτήρα “analysis-grade”: προτεραιοποιεί ποιότητα, μειώνει τεχνητές πηγές μεροληψίας, και παράγει ένα σταθερό baseline σύνολο χαρακτηριστικών κατάλληλο για προγνωστική μοντελοποίηση τύπου screening [56]–[58].

1.5. Διερευνητική Στατιστική Ανάλυση (Exploratory Data Analysis)

Η διερευνητική στατιστική ανάλυση (EDA) αποτέλεσε κρίσιμο βήμα για την κατανόηση της δομής, της ποιότητας και των εγγενών περιορισμών του συνόλου δεδομένων του Brain Precision, πριν από οποιαδήποτε μοντελοποίηση πρόγνωσης κινδύνου γνωστικής έκπτωσης. Η EDA οργανώθηκε γύρω από: (α) μονομεταβλητές κατανομές, (β) βασικές γνωστικές μετρήσεις στο baseline, (γ) συσχετίσεις/πολυσυγγραμμικότητα (correlations/multicollinearity) και (δ) ποσοτική αποτίμηση ανισορροπίας κλάσεων και επιπτώσεων στην αξιολόγηση. [7]

1.5.1. Ανάλυση Κατανομών Χαρακτηριστικών

1.5.1.1. Δημογραφικό προφίλ και δείκτες γήρανσης.

Το δείγμα προέρχεται από το πιλοτικό πρόγραμμα LETHE και περιλαμβάνει N=150 συμμετέχοντες, με χαρακτηριστικά τυπικά πληθυσμού μέσης-προχωρημένης ηλικίας που βρίσκεται σε αυξημένο κίνδυνο για γνωστική έκπτωση αλλά χωρίς έκδηλη βαριά διαταραχή στο baseline. Οι βασικές δημογραφικές μεταβλητές που εξετάστηκαν περιλαμβάνουν φύλο, ηλικία (age_y), έτη εκπαίδευσης (educ_years) και οικογενειακή κατάσταση (marital_status). Η EDA ανέδειξε σαφή υπεροχή του ενός φύλου και πλειοψηφία συμμετεχόντων που είναι παντρεμένοι/σε σταθερή σχέση, ενώ η ηλικία και τα έτη εκπαίδευσης παρουσιάζουν διασπορά αναμενόμενη για πληθυσμό αυτής της ηλικιακής ομάδας. Ως «δείκτες γήρανσης» στη συγκεκριμένη φάση θεωρήθηκαν κυρίως η χρονολογική ηλικία και έμμεσοι κοινωνικο-γνωστικοί προσδιοριστές που σχετίζονται με γνωστική εφεδρεία (π.χ. εκπαίδευση), δεδομένου ότι επηρεάζουν τόσο τις τιμές baseline των νευροψυχολογικών δοκιμασιών όσο και τη μεταβλητότητα στις διαχρονικές μεταβολές.

1.5.1.2. Κατανομές δεικτών σωματικής υγείας (BMI, μεταβολικοί δείκτες) και εντοπισμός παραγόντων κινδύνου.

Στις ανθρωπομετρικές και κλινικές μεταβλητές, η EDA έδειξε ότι η ηλικία και ο δείκτης μάζας σώματος (BMI) εμφανίζουν σχετικά συμμετρικές κατανομές, με την ηλικία να συγκεντρώνεται κυρίως μεταξύ 65–75 ετών—εύρημα που επιβεβαιώνει ότι το δείγμα αντιπροσωπεύει πληθυσμό αυξημένου κινδύνου λόγω ηλικίας.

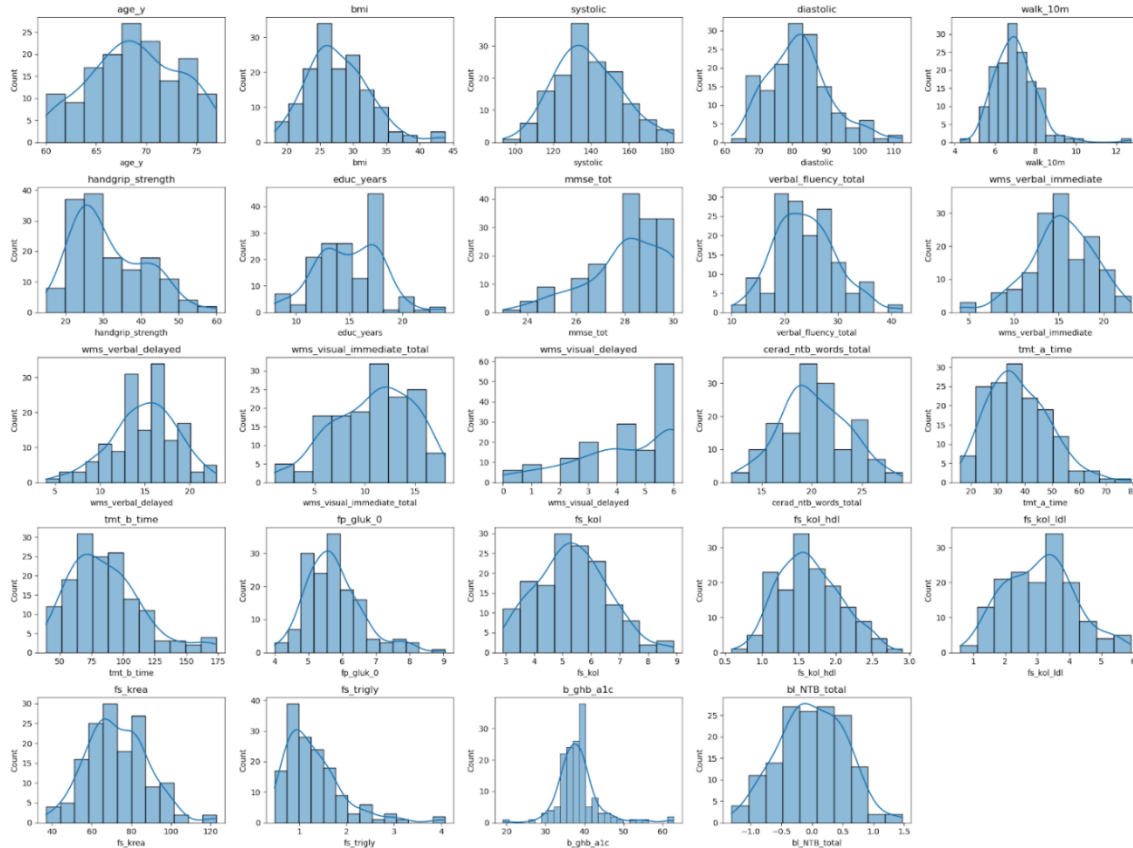
Οι μετρήσεις αρτηριακής πίεσης (systolic/diastolic) παρουσιάζουν ήπια δεξιά ασυμμετρία και παρουσία τιμών σε υπερτασικά επίπεδα, στοιχείο συμβατό με το κλινικό προφίλ συμμετεχόντων με αγγειακούς/μεταβολικούς παράγοντες κινδύνου. Οι βιοχημικοί και μεταβολικοί δείκτες που αναλύθηκαν περιλαμβάνουν γλυκόζη νηστείας (fp_gluk_0), HbA1c (b_ghb_a1c), κρεατινίνη (fs_krea) και λιπιδαιμικό προφίλ (fs_kol, fs_kol_hdl, fs_kol_ldl, fs_trigly). Η EDA κατέδειξε ήπια ασυμμετρία στις περισσότερες από αυτές τις μεταβλητές, χωρίς ακραίες τιμές που να απαιτούν «επιθετικό» χειρισμό outliers, υποστηρίζοντας ότι το σύνολο είναι αριθμητικά σταθερό και κατάλληλο για περαιτέρω ανάλυση. Η επιλογή και ερμηνεία των παραπάνω δεικτών τεκμηριώνεται και από το λεξικό μεταβλητών/ορισμών (π.χ. LDL/HDL/ολική χοληστερόλη, HbA1c), που επιτρέπει συνεπή βιοϊατρική αντιστοίχιση και ενοποίηση χαρακτηριστικών.

1.5.2. Ανάλυση Νευροψυχολογικών Δεικτών

1.5.2.1. Διακύμανση στις δοκιμασίες μνήμης και εκτελεστικής λειτουργίας (MMSE, TMT, WMS) κατά την έναρξη.

Οι γνωστικές μεταβλητές στο baseline (MMSE, δείκτες μνήμης από WMS/CERAD, δείκτες εκτελεστικών λειτουργιών όπως TMT) εμφάνισαν κατανομές με τάση συγκέντρωσης σε υψηλές

τιμές, εύρημα αναμενόμενο σε πληθυσμό χωρίς σοβαρή γνωστική διαταραχή κατά την ένταξη. Ειδικότερα, το MMSE παρουσίασε μέση τιμή περίπου 28.1 με εύρος 23–30, ενισχύοντας την υπόθεση ότι η πλειονότητα είναι εντός φυσιολογικών ορίων στο baseline. Παρά την «οροφή» σε ορισμένα τεστ, παρατηρήθηκε ουσιαστική διασπορά σε επιμέρους δείκτες μνήμης και εκτελεστικής λειτουργίας (π.χ. λεκτική ευχέρεια, WMS verbal immediate), η οποία θεωρείται ιδιαίτερα κρίσιμη για πρόγνωση, καθώς μπορεί να αποτυπώνει υποκλινικές νδιαφοροποιήσεις που σχετίζονται με μελλοντική αρνητική μεταβολή στη συνολική νευροψυχολογική επίδοση.



Εικόνα 1: Κατανομές Αριθμητικών Χαρακτηριστικών.

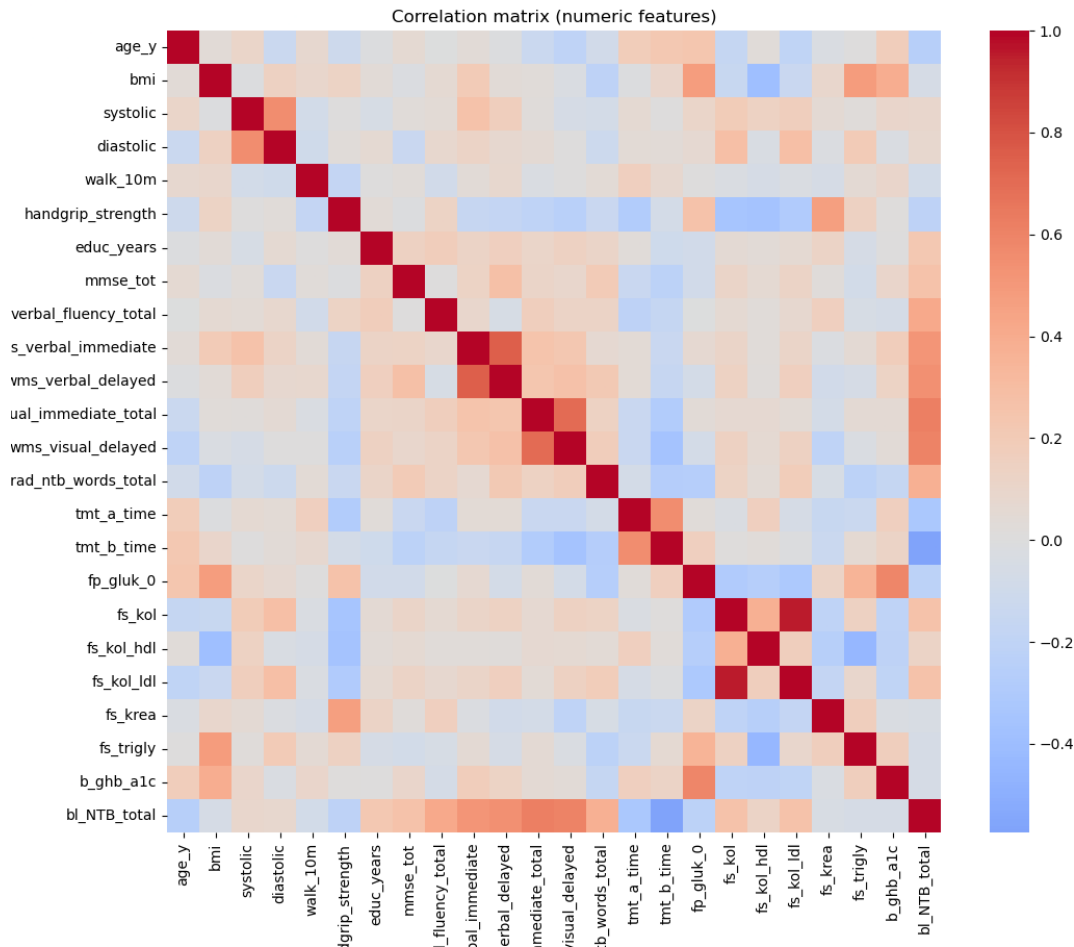
1.5.3. Ανάλυση συσχετίσεων μεταξύ των χαρακτηριστικών

1.5.3.1. Πίνακας Συσχέτισης (Correlation Matrix) και εντοπισμός πολυσυγγραμμικότητας

Ο πίνακας (Εικόνα 22) συσχέτισης ανέδειξε εσωτερικές δομές με «συμπλέγματα» ισχυρών θετικών συσχετίσεων μεταξύ επαναλαμβανόμενων/συναφών μετρήσεων, ιδίως στις νευροψυχολογικές δοκιμασίες (π.χ. WMS immediate vs delayed) και σε ορισμένες κλινικές μετρήσεις. Ενδεικτικά, στο λιπιδαιμικό προφίλ παρατηρήθηκε ιδιαίτερα υψηλή γραμμική συσχέτιση μεταξύ ολικής χοληστερόλης (fs_kol) και LDL (fs_kol_ldl), με Pearson $r=0.9535$, βιολογικά αναμενόμενο εύρημα που δημιουργεί κίνδυνο πολυσυγγραμμικότητας και «πλεονασμού» πληροφορίας όταν χρησιμοποιούνται γραμμικά/παραμετρικά μοντέλα.

Η ύπαρξη τέτοιων συσχετισμένων χαρακτηριστικών αιτιολογεί αφενός την ανάγκη προσεκτικής επιλογής χαρακτηριστικών (feature selection) ή/και κανονικοποίησης (normalization) για

παραμετρικές προσεγγίσεις, αφετέρου τη χρήση μεθόδων με μεγαλύτερη ανθεκτικότητα σε συσχετισμένα predictors (π.χ. Random Forest).



Εικόνα 22: Διάγραμμα συσχέτισης μεταβλητών.

1.5.3.2. Συσχέτιση baseline χαρακτηριστικών με τη μεταβλητή-στόχο

Η EDA κατέδειξε ότι το baseline συνολικό σκορ (bl NTB_total) λειτουργεί ως «κεντρικός κόμβος» πληροφορίας με μέτριες έως ισχυρές συσχετίσεις προς τις περισσότερες γνωστικές παραμέτρους, όπως είναι αναμενόμενο δεδομένης της συνθετικής φύσης των νευροψυχολογικών δεικτών.

Αντίθετα, μεταβλητές όπως BMI και ηλικία εμφανίστηκαν περισσότερο «απομονωμένες» στον γραμμικό χώρο, ενώ συνολικά δεν προέκυψαν ισχυρές γραμμικές συσχετίσεις μεταξύ δημογραφικών/βιολογικών δεικτών και γνωστικών αποτελεσμάτων.

Το παραπάνω είναι ιδιαίτερα σημαντικό για τη μεταβλητή-στόχο (δυναμικός κίνδυνος γνωστικής έκπτωσης σε 24 μήνες, ορισμένος μέσω μη θετικής ετήσιας μεταβολής στο NTB total score), καθώς υποδηλώνει ότι η διάκριση «χαμηλού vs υψηλού κινδύνου» δεν αναμένεται να είναι εμφανής με απλές μονομεταβλητές ή καθαρά γραμμικές σχέσεις. Συνεπώς, τεκμηριώνεται η επιλογή μη γραμμικών μοντέλων και μεθόδων που μπορούν να ανιχνεύσουν αλληλεπιδράσεις/συνδυαστικές επιδράσεις μεταξύ ετερογενών χαρακτηριστικών (κλινικών, βιοχημικών, γνωστικών, συμπεριφορικών).

1.5.4. Ανισορροπία Κλάσεων (Class Imbalance)

1.5.4.1. Ποσοτική αποτίμηση της ανισορροπίας και επιπτώσεις στη μοντελοποίηση

Η μεταβλητή-στόχος εμφάνισε μέτρια ανισορροπία κλάσεων, με περίπου 66% των συμμετεχόντων να ανήκουν στη χαμηλού κινδύνου κατηγορία και 34% στην υψηλού κινδύνου κατηγορία.

Παρότι η ανισορροπία δεν είναι «ακραία», είναι επαρκής ώστε να δημιουργεί φαινόμενα majority-class dominance κατά την εκπαίδευση, ιδιαίτερα όταν χρησιμοποιούνται μετρικές όπως η συνολική ακρίβεια (accuracy) χωρίς εξισορρόπηση ή χωρίς κατάλληλη επιλογή κατωφλίων/βαρών.

Ως συνέπεια, η EDA δεν αντιμετωπίστηκε ως απλή περιγραφική διαδικασία, αλλά ως βήμα εντοπισμού κινδύνων μεθοδολογικής μεροληψίας: (α) ανάγκη χρήσης balanced metrics (π.χ. recall, F1, balanced accuracy), (β) ανάγκη στρατηγικών αντιμετώπισης ανισορροπίας (π.χ. downsampling/oversampling ή cost-sensitive learning) και (γ) προσεκτική κλινική ερμηνεία λαθών με έμφαση στον περιορισμό ψευδώς αρνητικών σε σενάρια screening.

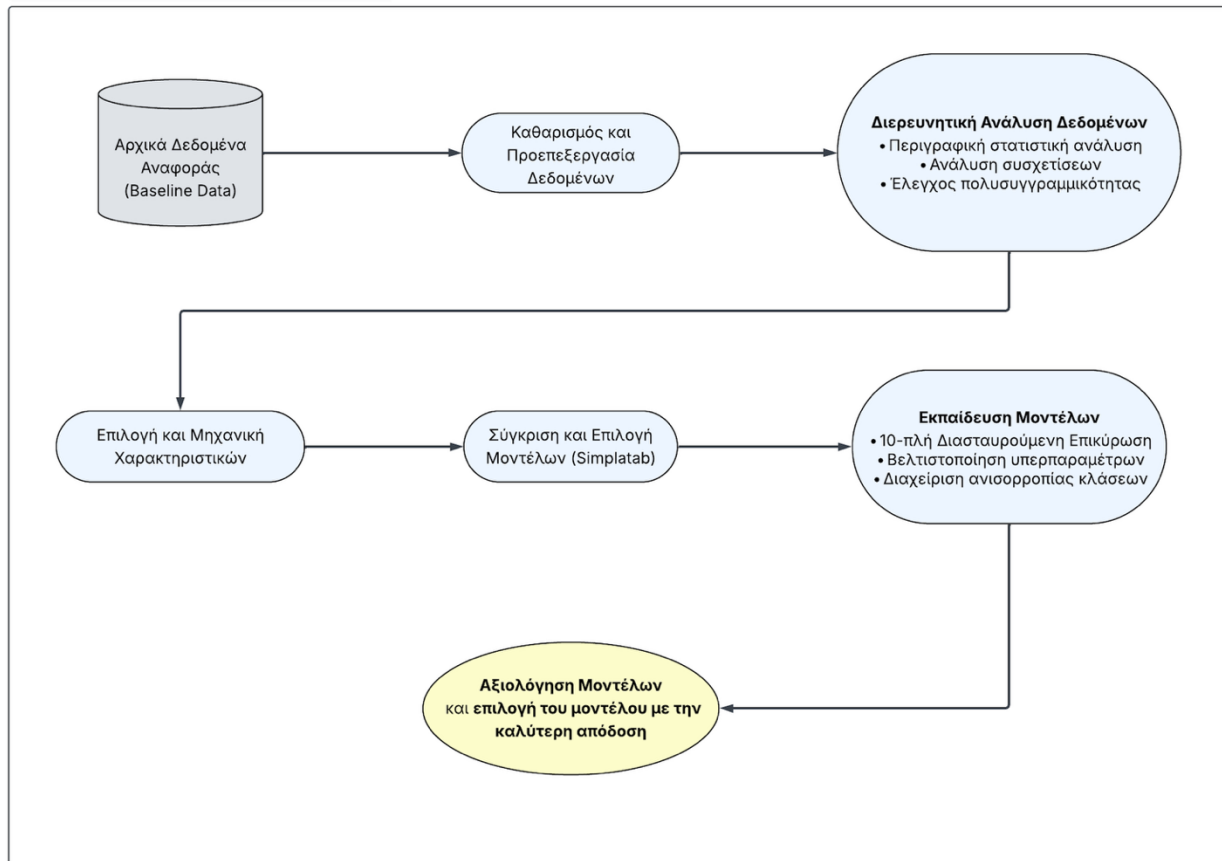
Πίνακας 1: Ποσοστά ελλειπόν τιμών ανά μεταβλητή.

Feature Name (Όνομα μεταβλητής)	Missing Values (%) (Ποσοστό ελλειπουσών τιμών)
apoeR	27.81
zung_total_score	23.18
alc4freq	4.64
fs_krea	1.99
tmt_a_time	0.66
vms_verbal_delayed	0.66
fp_gluk_0	0.66
tmt_b_time	0.66
sex	0.00
verbal_fluency_total	0.00

1.6. Μεθοδολογία Μηχανικής Μάθησης

Η φάση της μηχανικής μάθησης στο *Brain Precision* σχεδιάστηκε ως «κλινικά ρεαλιστικό» σενάριο προσυμπτωματικού ελέγχου (screening), όπου η πρόγνωση κινδύνου πραγματοποιείται αποκλειστικά με χαρακτηριστικά βάσης (baseline), με χρονικό ορίζοντα πρόβλεψης τους 24 μήνες, ώστε να ανιχνεύονται ουσιαστικές μεταβολές χωρίς να χάνεται η δυνατότητα έγκαιρης παρέμβασης. Καθοριστικός παράγοντας της μεθοδολογίας ήταν επίσης η μέτρια ανισορροπία κλάσεων (πλειοψηφία χαμηλού κινδύνου έναντι περίπου του ενός τρίτου υψηλού κινδύνου), η οποία επιβάλλει επιλογές πειραματικού σχεδιασμού και μετρικών πέρα από το απλό accuracy.

Data Analysis Workflow Diagram



Εικόνα 3: Σχεδιάγραμμα μεθοδολογίας.

1.6.1. Στρατηγική Επιλογής Μοντέλων

1. 6.1.1. Συγκριτική αξιολόγηση μέσω AutoML (SimplaTab)

Για την αρχική διερεύνηση του «χώρου μοντέλων» υιοθετήθηκε προσέγγιση AutoML μέσω του SimplaTab/Simplatab, με στόχο τη γρήγορη, αναπαραγώγιμη και συγκρίσιμη αξιολόγηση ετερογενών αλγορίθμων σε ενιαίο πλαίσιο μετρικών. Το Simplatab αναφέρεται ως ολοκληρωμένο πλαίσιο AutoML που ενσωματώνει στάδια επιλογής χαρακτηριστικών, εκπαίδευσης/βελτιστοποίησης και ΧΑΙ, υποστηρίζοντας τη γρήγορη συγκριτική αξιολόγηση μοντέλων σε δομημένα δεδομένα [49]. Στο baseline benchmarking (πίνακας συγκριτικής επίδοσης), η αξιολόγηση της απόδοσης των αλγορίθμων εποπτευόμενης μηχανικής μάθησης πραγματοποιήθηκε με την χρήση πολλαπλών μετρικών όπως των Sensitivity, Specificity, AUC, F-score, Accuracy, ώστε να αποτυπώνεται η συμπεριφορά των μοντέλων και υπό ανισορροπία κλάσεων. Ειδικά σε σενάρια πρόγνωσης υψηλού κινδύνου, η έμφαση σε *Sensitivity/Recall* (για ελαχιστοποίηση των ψευδώς αρνητικών προβλέψεων) είναι κλινικά αιτιολογημένη, χωρίς όμως να αγνοείται η μετρική *Specificity* (υπερδιάγνωση/άσκοπες παρεμβάσεις).

1.6.1.2. Επιλογή αλγορίθμων: Support Vector Machines (SVM) και Random Forest (RF).

Με βάση τα αποτελέσματα του αρχικού συγκριτικού ελέγχου, επιλέχθηκαν δύο αλγοριθμικές οικογένειες με συμπληρωματικά χαρακτηριστικά γενίκευσης σε μικρομεσαία tabular κλινικά σύνολα:

- **SVM:** παρουσίασε την υψηλότερη ευαισθησία ($Sensitivity \approx 0.88$) στο baseline, άρα ήταν ισχυρός υποψήφιος για τον εντοπισμό ατόμων υψηλού κινδύνου. Θεωρητικά, οι SVM έχουν τεκμηριωμένη ικανότητα γενίκευσης μέσω μεγιστοποίησης περιθωρίου, ειδικά σε προβλήματα δυαδικής ταξινόμησης.
- **Random Forest:** πέτυχε την υψηλότερη ακρίβεια ($Accuracy \approx 0.61$) στο baseline και αποτελεί ανθεκτική επιλογή έναντι θορύβου/μη γραμμικότητας, με ενσωματωμένη σταθεροποίηση μέσω συνόλου τυχαίων δέντρων, όπως περιγράφεται κλασικά από τον Breiman.

Η διπλή επιλογή (SVM + RF) επιτρέπει να διερευνηθεί πρακτικά το trade-off μεταξύ «ευαισθησίας ανίχνευσης υψηλού κινδύνου» και «ισορροπημένης συνολικής επίδοσης», υπό το πρίσμα της ανισορροπίας κλάσεων.

1.6.2. Πλαίσιο Εκπαίδευσης και Επικύρωσης.

1.6.2.1. Διασταυρούμενη Επικύρωση 10 folds (10-fold Cross-Validation).

Οι επιλεγμένοι αλγόριθμοι υλοποιήθηκαν εκ νέου σε προσαρμοσμένο περιβάλλον Python/Scikit-Learn, ώστε να υπάρχει πλήρης έλεγχος της εκπαίδευσης και της πειραματικής διαδικασίας. Για τη μείωση της διακύμανσης των εκτιμήσεων επίδοσης και τον περιορισμό του κινδύνου υπερεκπαίδευσης (*overfitting*), εφαρμόστηκε η τεχνική **10-fold cross-validation** στο σύνολο εκπαίδευσης. Η επιλογή 10-fold CV στηρίζεται και στη βιβλιογραφική τεκμηρίωση ότι αποτελεί ισχυρή πρακτική για εκτίμηση ακρίβειας και επιλογή μοντέλου, προσφέροντας καλή ισορροπία bias-variance σε ρεαλιστικά δεδομένα.

1.6.2.2. Βελτιστοποίηση Υπερπαραμέτρων (Hyperparameter Tuning).

Παράλληλα με την επικύρωση, πραγματοποιήθηκε **βελτιστοποίηση υπερπαραμέτρων** εστιάζοντας σε κρίσιμες παραμέτρους που επηρεάζουν άμεσα τη πολυπλοκότητα και τη γενίκευση:

- για **RF:** αριθμός δέντρων ($n_estimators$) και μέγιστο βάθος (max_depth),
- για **SVM:** παράμετρος ποινής C και επιλογή πυρήνα (π.χ. **RBF kernel**).

Σε επίπεδο αρχών, η βιβλιογραφία υποστηρίζει ότι η αποδοτική αναζήτηση υπερπαραμέτρων (π.χ. τυχαίοι αναζήτηση έναντι grid) μπορεί να υπερέχει πρακτικά σε υψηλοδιάστατους χώρους υπερπαραμέτρων, κάτι ιδιαίτερα χρήσιμο σε κλινικά σύνολα όπου ο υπολογιστικός προϋπολογισμός και ο κίνδυνος υπερεκπαίδευσης είναι περιοριστικοί παράγοντες.

1.6.3. Αντιμετώπιση Ανισορροπίας Κλάσεων.

Η ανισορροπία (περισσότερα «σταθερά/χαμηλού κινδύνου» έναντι «πτώσης/υψηλού κινδύνου») μπορεί να οδηγήσει σε φαινόμενο *majority-class dominance*, όπου ένα μοντέλο εμφανίζεται «καλό» λόγω υψηλού accuracy, αλλά αποτυγχάνει στο κλινικά κρίσιμο σενάριο της ανίχνευσης υψηλού κινδύνου. Για τον λόγο αυτό διερευνήθηκαν δύο συμπληρωματικές στρατηγικές: δειγματοληπτική εξισορρόπηση και αλγοριθμική/κοστοευσθητή μάθηση.

1.6.3.1. Εφαρμογή τεχνικών Υποδειγματοληψίας (Downsampling).

Εφαρμόστηκε **downsampling** της πλειοψηφικής κλάσης, ώστε να εξισωθεί με τη μειονοτική, μειώνοντας την εκ φύσεως «πίεση» προς ταξινόμηση στην πλειοψηφική κατηγορία. Η επιλογή **downsampling** είναι ιδιαίτερα κατάλληλη όταν: (α) το δείγμα δεν είναι μεγάλο, (β) η υπερβολική ενίσχυση της μειονοτικής κλάσης μπορεί να οδηγήσει σε **overfitting**, και (γ) επιδιώκεται απλότητα/διαφάνεια στον χειρισμό της ανισορροπίας.

1.6.3.2. Σύγκριση με μεθόδους **Cost-sensitive learning**.

Ως εναλλακτική/συγκριτική προσέγγιση, εφαρμόστηκε **cost-sensitive learning** μέσω βαρών κλάσεων (*class weights*), αποδίδοντας μεγαλύτερο «κόστος» στα σφάλματα της μειονοτικής κλάσης (υψηλού κινδύνου), ώστε ο αλγόριθμος να «τιμωρείται» περισσότερο για ψευδώς αρνητικές προβλέψεις.

Η τελική επιλογή στρατηγικής (**downsampling vs cost-sensitive**) αντιμετωπίστηκε ως εμπειρικό ζήτημα που κρίνεται από την επίδοση σε κατάλληλες μετρικές (ιδίως Recall/Sensitivity, F-score, AUC) και από τη σταθερότητα της επίδοσης σε **cross-validation**, υπό ρεαλιστικούς κλινικούς συμβιβασμούς.

1.6.4. Εφαρμογή Επεξηγηματικότητας / Επεξηγηματικότητα Μοντέλων (Explainability).

1.6.4.1. Χρήση τιμών SHAP (Shapley Additive Explanations) για την κατανόηση της συμβολής των χαρακτηριστικών.

Για την αξιολόγηση της επεξηγηματικότητας των μοντέλων μηχανικής μάθησης υιοθετήθηκε το πλαίσιο **SHAP**, το οποίο παρέχει συνεπείς, θεωρητικά θεμελιωμένες αποδόσεις συμβολής χαρακτηριστικών σε επίπεδο ατομικής πρόβλεψης (*local explanations*), βασισμένες σε τιμές **Shapley**. Η χρήση **SHAP** εξυπηρετεί διττό στόχο:

1. **κλινική κατανόηση** του «γιατί» πίσω από μια πρόβλεψη υψηλού κινδύνου (π.χ. ποια **baseline** χαρακτηριστικά ωθούν την απόφαση),
2. **ποιοτικό έλεγχο** μοντέλου (εντοπισμός μη επιθυμητών συσχετίσεων/πιθανών μεροληψιών).

Σε επίπεδο συστήματος, η λογική «**predict + explain**» είναι συμβατή και με αρχιτεκτονικές ανάπτυξης υπηρεσιών πρόβλεψης που περιλαμβάνουν διακριτό **endpoint** ερμηνείας βασισμένο σε **SHAP**.

1.6.4.2. Αποφυγή του φαινομένου "Black Box" στην κλινική λήψη αποφάσεων.

Η αποφυγή του φαινομένου “**black box**” δεν αποτελεί απλώς ζήτημα διαφάνειας, αλλά και κλινικής ασφάλειας/λογοδοσίας: η επεξηγηματικότητα μειώνει τον κίνδυνο λήψης αποφάσεων βασισμένων σε «αδιαφανείς» μηχανισμούς, επιτρέπει έλεγχο ευλογοφάνειας (*clinical plausibility*) και διευκολύνει τη συζήτηση με τον κλινικό χρήστη για τα όρια/βεβαιότητες της πρόβλεψης. Σε πρακτικό επίπεδο, η αξιοποίηση βιβλιοθηκών/υλοποιήσεων **SHAP** επιτρέπει τυποποιημένη εξαγωγή εξηγήσεων και ενσωμάτωσή τους σε υπολογιστικά **pipelines**.

1.7. Αποτελέσματα και Αξιολόγηση Μοντέλων

Η αξιολόγηση των μοντέλων ταξινόμησης στο **Brain Precision** δομήθηκε σε τρεις διακριτές βαθμίδες: (α) **AutoML baseline** για γρήγορη χαρτογράφηση του χώρου μοντέλων, (β) **10-fold cross-validation** των επιλεγμένων μοντέλων σε περιβάλλον **Python/Scikit-Learn** για εκτίμηση σταθερότητας/γενίκευσης και ανίχνευση μεροληψιών λόγω ανισορροπίας κλάσεων, και (γ)

αξιολόγηση σε τυφλό σύνολο ελέγχου (unseen/test set), ώστε να αποτυπωθεί η πραγματική συμπεριφορά εκτός εκπαίδευσης. Σε όλα τα στάδια, οι μετρικές που χρησιμοποιήθηκαν ερμηνεύτηκαν με κλινικό/επιχειρησιακό πρίσμα, καθώς στο screening για υψηλό κίνδυνο γνωστικής έκπτωσης η **ελαχιστοποίηση ψευδώς αρνητικών (FN)** είναι συχνά προτεραιότητα (να μη «χαθεί» άτομο υψηλού κινδύνου), αλλά απαιτείται ταυτόχρονα έλεγχος ψευδώς θετικών (FP) ώστε να μη δημιουργείται μη βιώσιμο φορτίο παραπομπών/παρεμβάσεων.

1.7.1. Συγκριτική Επίδοση

1.7.1.1. Αποτελέσματα προκαταρκτικής αξιολόγησης και το φαινόμενο της κυριαρχίας της πλειοψηφικής κλάσης (Majority-class dominance)

(A) AutoML baseline (Simplotab)

Η προκαταρκτική σύγκριση μέσω AutoML παρείχε μια πρώτη ένδειξη ότι το πρόβλημα είναι **μη τετριμμένο** και ευάλωτο σε *metric illusions* υπό ανισορροπία κλάσεων. Συγκεκριμένα, στον Πίνακα 3 καταγράφονται για πολλούς αλγόριθμους οι μετρικές Sensitivity (TPR), Specificity (TNR), AUC, F-score και Accuracy.

- **SVM:** εμφάνισε **Sensitivity = 0.88** αλλά **Specificity = 0.23**, με **AUC = 0.56** και **F-score = 0.58**, ενώ η μετρική **Accuracy** ήταν χαμηλή (**0.45**). Τεχνικά, αυτό σημαίνει ότι το SVM στο baseline έχει τάση να «σηκώνει» την πιθανότητα της θετικής κλάσης (υψηλού κινδύνου), πετυχαίνοντας υψηλό TPR, αλλά με μεγάλο κόστος σε FP (χαμηλό TNR).
- **Random Forest (RF):** πέτυχε **Accuracy = 0.61** με **Specificity = 0.72**, αλλά μόνο **Sensitivity = 0.41**, με **AUC = 0.57** και **F-score = 0.43**. Αυτό είναι κλασική ένδειξη ότι το μοντέλο ευνοεί την πλειοψηφική κλάση (χαμηλού κινδύνου): αυξημένο TNR/Accuracy, αλλά ανεπαρκές TPR.

Το baseline, επομένως, ανέδειξε ότι οι αλγόριθμοι **SVM** και **RF** αποτελούν τα δύο πιο «ενδιαφέροντα» μοντέλα, όχι επειδή υπερέχουν συνολικά ως προς την απόδοση, αλλά επειδή προσφέρουν **συμπληρωματικά προφίλ σφαλμάτων**: ο SVM ως μοντέλο υψηλής ευαισθησίας, ο RF ως μοντέλο υψηλότερης ειδικότητας/ακρίβειας.

(B) 10-fold cross-validation: τεκμηρίωση της υπεροχής της majority-class

Στην 10-πλή διασταυρούμενη επικύρωση, το φαινόμενο **majority-class dominance** αποτυπώθηκε με αριθμητικά σαφή τρόπο:

- **RF: Mean Accuracy = 0.635 ± 0.055**, αλλά **Mean Recall = 0.080** και **Mean F1 = 0.106**.
- **SVM: Mean Accuracy = 0.449 ± 0.135**, **Mean Recall = 0.520** και **Mean F1 = 0.309**.

Πίνακας 2: Αποτελέσματα 10-πλης διασταυρούμενης επικύρωσης.

Model	Mean Accuracy (+/- SD)	Mean Recall	Mean F1
Random Forest	0.635 +/- 0.055	0.080	0.106
SVM	0.449 +/- 0.135	0.520	0.309

Η κρίσιμη τεχνική παρατήρηση εδώ είναι ότι ο αλγόριθμος **RF εμφανίζει “accuracy paradox”**: ενώ η ακρίβεια φαίνεται σχετικά «καλή» (~0.64), η ανάκληση ~0.08 υποδηλώνει ότι το μοντέλο **σχεδόν ποτέ δεν προβλέπει θετική κλάση** (υψηλό κίνδυνο). Αυτό είναι συμβατό με στρατηγική “predict majority” ή με απόφαση που «σπάνια περνά» το κατώφλι θετικής ταξινόμησης, οδηγώντας σε πολλούς FN. Αντίθετα, ο SVM θυσιάζει accuracy/σταθερότητα (μεγάλο SD 0.135), αλλά διατηρεί σαφώς υψηλότερο recall (0.52), άρα ανιχνεύει περισσότερες θετικές περιπτώσεις.

Με μεθοδολογικούς όρους, σε imbalanced classification είναι γνωστό ότι μετρικές όπως Accuracy και ROC-AUC μπορεί να δώσουν υπεραισιόδοξη εικόνα, ενώ οι μετρικές που “βλέπουν” την απόδοση στη θετική κλάση (Precision/Recall, F1, PR curve) είναι πιο διαγνωστικές για την κλινικά κρίσιμη μειονοτική κλάση

Πίνακας 3: Σύγκριση μοντέλων μέσω του εργαλείου Simplotab.

Model	Sensitivity	Specificity	AUC	F-score	Accuracy
Decision Trees	0.61	0.55	0.58	0.49	0.57
Logistic Regression	0.56	0.52	0.54	0.45	0.53
Multi-layer NN	0.51	0.57	0.54	0.43	0.55
Random Forest	0.41	0.72	0.57	0.43	0.61
Stochastic GD	0.81	0.27	0.54	0.55	0.45
SVM	0.88	0.23	0.56	0.58	0.45
XGBoost	0.46	0.64	0.55	0.44	0.58

1.7.1.2. Κατάρρευση ορίου απόφασης στο Random Forest πριν την εξισορρόπηση

Η αξιολόγηση σε τυφλό σύνολο ελέγχου (unseen data) αποκάλυψε καθαρά ότι η επίδοση του RF πριν την εξισορρόπηση είναι **λειτουργικά μη αποδεκτή για screening**:

1. Ο **Random Forest πέτυχε Accuracy 64%**, αλλά **Recall 0.00**, καθώς **ταξινόμησε 18 από τις 19 περιπτώσεις** στην κλάση χαμηλού κινδύνου και **δεν εντόπισε κανένα περιστατικό υψηλού κινδύνου**. Αυτό είναι πρακτικά “collapse” της διαδικασίας απόφασης προς την πλειοψηφική κλάση: το μοντέλο, επηρεασμένο από το class prior και την εκπαίδευση χωρίς εξισορρόπηση, καταλήγει να υιοθετεί **συντηρητικό decision rule** υπέρ της αρνητικής κλάσης (low risk), μεγιστοποιώντας *apparent* accuracy αλλά μηδενίζοντας κλινική χρησιμότητα για early identification.
2. Ενδιαφέρον είναι ότι το report αναφέρει **ακραίες συμπεριφορές και στα δύο μοντέλα** στο test set, αναδεικνύοντας τη δυσκολία γενίκευσης σε μικρά/imbalanced δεδομένα.

Τεχνικά, το παραπάνω είναι σύμπτωμα δύο παραγόντων:

1. **Imbalanced training signal**: οι ενημερώσεις/διαχωρισμοί βελτιστοποιούν κυρίως τη σωστή πρόβλεψη της πλειοψηφικής κλάσης, εκτός αν επιβληθεί εξισορρόπηση ή κόστος.
2. **Thresholding & calibration issues**: ακόμη κι αν η κατάταξη (ranking) είναι μέτρια, η τελική κλάση εξαρτάται από κατώφλι πιθανότητας. Σε μη βαθμονομημένες πιθανότητες (συχνό σε ensembles), η χρήση default threshold μπορεί να υπο-ενεργοποιεί τις θετικές προβλέψεις.

Παράλληλα, στο ίδιο test set, το **SVM ταξινόμησε και τα 28 περιστατικά** στην κλάση υψηλού κινδύνου, άρα παρουσίασε **αντίστροφη** ακραία μεροληψία (πολλά FP, κατάρρευση ειδικότητας), που το καθιστά κλινικά μη αξιοποιήσιμο χωρίς κατάλληλη ρύθμιση/εξισορρόπηση.

1.7.2. Βελτιστοποίηση και Τελική Αξιολόγηση

1.7.2.1. Επίδραση του Downsampling: Βελτίωση F1-Score και Balanced Accuracy

Για να αντιμετωπιστούν τα παραπάνω failure modes, εφαρμόστηκε **downsampling** της πλειοψηφικής κλάσης στο training pipeline, με στόχο να «αναγκαστούν» οι αλγόριθμοι να μάθουν χαρακτηριστικά διαχωρισμού και για τη μειονοτική κλάση υψηλού κινδύνου. Μετά την εξισορρόπηση, τα αποτελέσματα (Πίνακας 2) έδειξαν θεμελιώδη αλλαγή στην ποιότητα πρόβλεψης:

- **Random Forest (με downsampling):**
Accuracy = 0.63, Precision = 0.60, Recall = 0.67, F1 = 0.63. [31]
Επιπλέον, αναφέρεται **ισορροπημένος πίνακας σύγχυσης με 6 σωστές προβλέψεις ανά κλάση**, ένδειξη ότι το μοντέλο πλέον δεν «αγνοεί» τη θετική κλάση και επιτυγχάνει πιο συμμετρική απόδοση.
- Το report τεκμηριώνει ρητά ότι το downsampling αποτέλεσε “σημείο καμπής”, καθώς η αρχική υπεροχή του RF σε accuracy ήταν πλασματική λόγω μεροληψίας προς την πλειονοτική κλάση.

Από τεχνική σκοπιά, η βελτίωση του **F1-score** είναι ιδιαίτερα σημαντική, διότι ο F1 λειτουργεί ως αρμονικός μέσος precision/recall και τιμωρεί τόσο την κατάρρευση recall (όπως πριν στο RF) όσο και την ανεξέλεγκτη αύξηση FP (συχνή σε high-recall μοντέλα). Σε συνθήκες imbalanced δεδομένων, η μετατόπιση από “accuracy-driven” συμπεριφορά προς “F1/balanced-performance” συμπεριφορά θεωρείται κεντρικός στόχος.

Τέλος, σημειώνεται ο αναμενόμενος συμβιβασμός: το downsampling συνεπάγεται απώλεια πληροφορίας από την πλειονοτική κλάση και δυνητική επίπτωση στη γενίκευση, αλλά στο παρόν κλινικό setting το όφελος στην ανίχνευση υψηλού κινδύνου κρίθηκε σημαντικότερο.

Πίνακας 4: Αποτελέσματα στρατηγικής Downsampling.

Model	Accuracy	Precision	Recall	F1-score
Random Forest	0.63	0.60	0.67	0.63
SVM	0.42	0.44	0.89	0.40

1.7.2.2. Ανάλυση Ευαισθησίας (Sensitivity Analysis)

Η ανάλυση ευαισθησίας επικεντρώθηκε στην κλινικά κρίσιμότερη διάσταση του προβλήματος: την ικανότητα ανίχνευσης της κλάσης **υψηλού κινδύνου** και, ειδικότερα, τον περιορισμό των **ψευδώς αρνητικών (false negatives)**. Σε ένα σενάριο πρόγνωσης για γνωστική έκπτωση, το να «χαθεί» ένα άτομο υψηλού κινδύνου (δηλαδή να ταξινομηθεί ως χαμηλού κινδύνου) είναι επιχειρησιακά πιο επιβαρυντικό από την υπερ-παραπομπή, διότι μεταφέρει το σφάλμα στο μέλλον με πιθανή απώλεια ευκαιρίας για έγκαιρη παρέμβαση. Υπό αυτό το πρίσμα, η κύρια μετρική ενδιαφέροντος ήταν το **Recall/Sensitivity (TPR)**, που αποτυπώνει το ποσοστό των πραγματικών θετικών που αναγνωρίζονται σωστά.

Μετά την εξισορρόπηση μέσω downsampling, ο αλγόριθμος **SVM** εμφάνισε σαφή υπεροχή ως προς την ευαισθησία, επιτυγχάνοντας **Recall = 0.89** στο test set, δηλαδή εντόπισε σχεδόν το 90% των περιπτώσεων υψηλού κινδύνου. Αυτό το εύρημα είναι τεχνικά ουσιαστικό, επειδή αντιστρέφει το failure mode που παρατηρήθηκε πριν την εξισορρόπηση, όπου μοντέλα μπορούσαν να εμφανίζονται «ικανοποιητικά» σε accuracy αλλά να αποτυγχάνουν στη θετική κλάση (π.χ. Recall≈0 για RF στο unseen set). Παράλληλα, η υψηλή ευαισθησία του SVM συνοδεύτηκε από χαμηλότερη συνολική ακρίβεια και F1 (Accuracy 0.42, F1 0.40), γεγονός που υποδηλώνει αύξηση ψευδώς θετικών και άρα μεγαλύτερη επιβάρυνση σε επίπεδο επιβεβαιωτικών ελέγχων/παρακολούθησης. Συνεπώς, ο SVM αναδεικνύεται ως καταλληλότερο όταν η στρατηγική είναι “**screening-first**” (προτεραιότητα στη μη απώλεια υψηλού κινδύνου), ενώ το RF, με πιο ισορροπημένο προφίλ (Precision 0.60, Recall 0.67, F1 0.63), παραμένει πιο κατάλληλο όταν ζητείται συνολική ισορροπία μεταξύ ευαισθησίας και ειδικότητας.

1.8. Συζήτηση και Συμπεράσματα

1.8.1. Η Ακρίβεια στα Κλινικά Δεδομένα

Στη δική μας μελέτη, το φαινόμενο που περιγράφεται ως «παράδοξο της ακρίβειας» δεν είναι θεωρητικό· προκύπτει άμεσα από τη δομή του συνόλου δεδομένων και τον ορισμό της μεταβλητής-στόχου. Το αρχικό σύνολο περιλαμβάνει **151 συμμετέχοντες** και **46 μεταβλητές έναρξης**, ενώ η μεταβλητή-στόχος είναι διαθέσιμη για **149/151** άτομα (απώλεια ~1.3% στο τελικό εποπτευόμενο πρόβλημα) και κωδικοποιεί δύο κλάσεις κινδύνου (χαμηλός/υψηλός). Η κατανομή των κλάσεων είναι **ασύμμετρη**: στο σύνολο εκπαίδευσης η κλάση χαμηλού κινδύνου είναι **0.6606** και η κλάση υψηλού κινδύνου **0.3395**, ενώ στο σύνολο ελέγχου οι αντίστοιχες αναλογίες είναι **0.6786** και **0.3214**. Με άλλα λόγια, περίπου τα **2/3** των παρατηρήσεων ανήκουν στη «ευκολότερη» (πλειονοτική) κλάση.

Υπό αυτές τις συνθήκες, η συνολική ακρίβεια (accuracy) μπορεί να δώσει μια **ψευδή εικόνα επάρκειας**, ακόμη και αν το μοντέλο είναι κλινικά ανεπαρκές. Ένας απλοϊκός ταξινομητής που προβλέπει συστηματικά «χαμηλό κίνδυνο» για όλους θα επιτύχει ακρίβεια ίση με την αναλογία της πλειονοτικής κλάσης, δηλαδή περίπου **0.66–0.68**, χωρίς να αναγνωρίζει ουσιαστικά τα άτομα που μας ενδιαφέρουν περισσότερο (υψηλού κινδύνου). Αυτό είναι ακριβώς το παράδοξο: μια τιμή accuracy που μοιάζει «ανεκτή» μπορεί να αντιστοιχεί σε **μηδενική κλινική ωφέλεια** για τον σκοπό της έγκαιρης ανίχνευσης [79], [80].

Στο δικό μας πλαίσιο, ο επιδιωκόμενος σκοπός είναι η **στρωματοποίηση κινδύνου** σε ορίζοντα 24 μηνών ώστε να υποστηριχθούν αποφάσεις παρακολούθησης και στοχευμένων ενεργειών. Άρα, η αξιολόγηση πρέπει να εστιάζει πρωτίστως στην ικανότητα εντοπισμού της κλάσης υψηλού κινδύνου και όχι στη συνολική ορθότητα. Συνεπώς:

1. **Η ευαισθησία (recall) και η θετική προγνωστική τιμή (precision) είναι κεντρικές.** Σε ασύμμετρες κλάσεις, η καμπύλη που συνδέει την ευαισθησία με τη θετική προγνωστική τιμή και η αντίστοιχη επιφάνεια κάτω από αυτήν αποτυπώνουν με μεγαλύτερη πιστότητα την πρακτική απόδοση από την καμπύλη ROC, διότι η θετική προγνωστική τιμή ενσωματώνει άμεσα το φορτίο των ψευδώς θετικών συναγερμών που συνεπάγεται η εφαρμογή του μοντέλου [79]. Η διάκριση μέσω ROC-AUC παραμένει χρήσιμη, αλλά δεν προστατεύει από περιπτώσεις όπου η συνολική διάκριση είναι ικανοποιητική ενώ η θετική προγνωστική τιμή είναι χαμηλή στο κλινικά χρήσιμο κατώφλι.
2. **Απαιτούνται ισορροπημένα μέτρα συνολικής ποιότητας.** Σε προβλήματα όπως το δικό μας, όπου η πλειονοτική κλάση υπερτερεί ποσοτικά, δείκτες όπως ο **συντελεστής συσχέτισης Matthews (MCC)** αποτυπώνουν καλύτερα την πραγματική ποιότητα ταξινόμησης, διότι λαμβάνουν υπόψη ταυτόχρονα αληθώς/ψευδώς θετικά και αληθώς/ψευδώς αρνητικά και δεν «επιβραβεύουν» την απλή πρόβλεψη της πλειονοτικής κλάσης [80].
3. **Η βαθμονόμηση των πιθανοτήτων είναι προϋπόθεση κλινικής αξιοπιστίας.** Η δική μας εφαρμογή δεν αποσκοπεί μόνο σε μια δυαδική ετικέτα, αλλά στη χαρτογράφηση ενός ατόμου σε επίπεδο κινδύνου. Ένα μοντέλο μπορεί να εμφανίζει ικανοποιητική διάκριση, αλλά να είναι κλινικά επισφαλές αν οι παραγόμενες πιθανότητες δεν αντανακλούν πραγματικούς κινδύνους. Η βαθμονόμηση έχει τεκμηριωθεί ως κρίσιμη αδυναμία των

προγνωστικών συστημάτων, διότι η κακή βαθμονόμηση οδηγεί σε λανθασμένη κατάσταση κινδύνου και άρα σε λανθασμένες αποφάσεις για παρακολούθηση/παρέμβαση [81].

Συνεπώς, για τη δική μας μελέτη, η συνολική ακρίβεια δεν μπορεί να αποτελεί πρωτεύον κριτήριο αξιολόγησης. Η ορθή τεκμηρίωση απόδοσης προϋποθέτει (i) μετρικές που εστιάζουν στην κλάση υψηλού κινδύνου (ευαισθησία και θετική προγνωστική τιμή και η μεταξύ τους σχέση), (ii) ισορροπημένους δείκτες συνολικής ποιότητας όπως ο MCC και (iii) έλεγχο βαθμονόμησης πιθανοτήτων, ώστε η έξοδος του μοντέλου να είναι κλινικά ερμηνεύσιμη και ασφαλής [79]–[81]

1.8.2. Περιορισμοί της Μελέτης

1.8.2.1 Μέγεθος δείγματος και στατικότητα μετρήσεων έναρξης

Ο πρώτος περιορισμός αφορά το **μέγεθος του δείγματος** σε συνδυασμό με την **πολυπαραμετρική περιγραφή** της κατάστασης υγείας. Η δική μας ανάλυση βασίζεται σε **151 συμμετέχοντες** και **46 μεταβλητές έναρξης** (με τελικό εποπτευόμενο δείγμα 149 ατόμων λόγω ελλείψεων στη μεταβλητή-στόχο). Σε τέτοια μεγέθη, η ανάπτυξη προγνωστικών μοντέλων είναι ευάλωτη σε: (i) αυξημένη διακύμανση των εκτιμήσεων, (ii) αισιόδοξη εκτίμηση απόδοσης κατά την εσωτερική επικύρωση, και (iii) υπερπροσαρμογή, ιδιαίτερα όταν εξετάζονται πιο σύνθετοι αλγόριθμοι ή εκτεταμένη ρύθμιση υπερ-παραμέτρων [82], [83]. Οι σύγχρονες μεθοδολογικές εργασίες για την ελάχιστη απαιτούμενη δειγματοληψία δείχνουν ότι η επάρκεια δεν εξαρτάται από απλούς εμπειρικούς κανόνες, αλλά από τον επιπολασμό του συμβάντος, τον αναμενόμενο βαθμό εξήγησης/πρόβλεψης και την πολυπλοκότητα του μοντέλου [82], [83]. Στο δικό μας πρόβλημα, όπου η κλάση υψηλού κινδύνου αντιστοιχεί περίπου στο 1/3, ο αριθμός «θετικών» παρατηρήσεων είναι περιορισμένος και αυτό θέτει ανώτατο όριο στη σταθερότητα σύνθετων μοντέλων.

Δεύτερον, η μελέτη αξιοποιεί **αποκλειστικά μετρήσεις έναρξης**. Αυτή η επιλογή είναι συνεπής με τον στόχο της έγκαιρης πρόγνωσης, αλλά εισάγει έναν δομικό περιορισμό: η πρόγνωση επιχειρείται από ένα «στιγμιότυπο» και όχι από δυναμικές τροχιές μεταβολής. Σε νευροεκφυλιστικά φαινόμενα, η προγνωστική πληροφορία συχνά ενισχύεται όταν περιλαμβάνονται δείκτες μεταβολής ή ενδιάμεσες μετρήσεις που αποτυπώνουν την εξέλιξη (π.χ. συμπεριφορικές μεταβολές, σταδιακή επιδείνωση σε δοκιμασίες). Η στατικότητα των εισόδων περιορίζει, συνεπώς, το μέγιστο επιτεύξιμο σήμα πρόγνωσης και μπορεί να οδηγεί σε μοντέλα που διαχωρίζουν κυρίως βάσει σταθερών διαφορών υγείας και όχι βάσει πρώιμης τροχιάς επιδείνωσης [81].

Επιπλέον, η διερεύνηση των δεδομένων ανέδειξε μεταβλητές με **μηδενική ή σχεδόν μηδενική διακύμανση** (π.χ. μεταβλητές που λαμβάνουν την ίδια τιμή για όλους τους συμμετέχοντες), οι οποίες δεν μπορούν να συνεισφέρουν σε προγνωστική ισχύ και υποδηλώνουν είτε περιορισμένη κλινική ετερογένεια στο συγκεκριμένο δείγμα είτε ανάγκη προσεκτικού ελέγχου κωδικοποίησης. Παράλληλα, εντοπίστηκε **πολύ υψηλή συσχέτιση** μεταξύ ολικής χοληστερόλης και LDL ($r \approx 0.954$), στοιχείο που αυξάνει τον κίνδυνο πλεονασμού και αστάθειας σε γραμμικά ή ημι-γραμμικά μοντέλα, αν δεν εφαρμοστούν κατάλληλες στρατηγικές επιλογής/συρρίκνωσης χαρακτηριστικών [83].

1.8.2.2 Η επίδραση των ελλিপών δεδομένων (π.χ. APOE) στην προγνωστική ισχύ

Ο δεύτερος περιορισμός σχετίζεται με το προφίλ ελλিপών τιμών σε κρίσιμες μεταβλητές. Στο σύνολο δεδομένων μας, οι μεγαλύτερες ελλείψεις αφορούν: **APOE (apoeR) ~27.8%**, **συνολικό**

σκορ Zung ~23.2%, ενώ μικρότερες αλλά υπαρκτές ελλείψεις εμφανίζονται και σε άλλους δείκτες (π.χ. συχνότητα κατανάλωσης αλκοόλ ~4.6%, κρεατινίνη ~2.0%, και ορισμένες νευροψυχολογικές/βιοχημικές μεταβλητές ~0.66%). Αυτές οι ελλείψεις δεν επηρεάζουν μόνο την πληρότητα· επηρεάζουν άμεσα και την προγνωστική ικανότητα και τη δυνατότητα γενίκευσης.

Ο γονότυπος **APOE ε4** αποτελεί έναν από τους ισχυρότερους γενετικούς παράγοντες κινδύνου για νόσο Alzheimer, με τεκμηριωμένη επίδραση δόσης-γονιδίου και σταθερή επιβεβαίωση σε μετα-αναλύσεις [87], [88]. Συνεπώς, η μερική απουσία του APOE μειώνει την ικανότητα του μοντέλου να αξιοποιήσει έναν βιολογικά «υψηλής πληροφορίας» άξονα κινδύνου. Όταν μια τέτοια μεταβλητή λείπει σε μεγάλο ποσοστό, προκύπτει ένα αναγκαστικό δίλημμα: είτε αφαιρείται από το κύριο προγνωστικό σύνολο (ώστε να μην υπονομευθεί η σταθερότητα και το αποτελεσματικό μέγεθος δείγματος), είτε επιχειρείται συμπλήρωση τιμών με συνακόλουθη αύξηση αβεβαιότητας. Η βιβλιογραφία για την πολλαπλή συμπλήρωση τιμών δείχνει ότι, ακόμη και όταν εφαρμόζεται σωστά, η αποτελεσματικότητα εξαρτάται από τον μηχανισμό απουσίας και από την επάρκεια του μοντέλου συμπλήρωσης, ενώ όταν η απουσία τείνει να μην είναι τυχαία, ο κίνδυνος μεροληψίας αυξάνει [84], [85].

Εξίσου σημαντικό είναι ότι η απουσία γενετικών/ψυχομετρικών δεδομένων σε πραγματικά κλινικά περιβάλλοντα συχνά σχετίζεται με οργανωτικούς ή συμμετοχικούς παράγοντες (διαθεσιμότητα εξέτασης, αποδοχή, επιβάρυνση συμμετέχοντα), οι οποίοι ενδέχεται να συσχετίζονται και με την κατάσταση υγείας. Σε αυτήν την περίπτωση, η απουσία μπορεί να δημιουργεί συστηματική απόκλιση μεταξύ όσων έχουν και όσων δεν έχουν διαθέσιμη τη μεταβλητή, με αποτέλεσμα το μοντέλο να μαθαίνει πρότυπα που δεν αναπαράγονται εύκολα κατά την εφαρμογή. Η διεθνής βιβλιογραφία τονίζει ότι η εφαρμογή κλινικών προγνωστικών μοντέλων σε νέους ασθενείς με ελλείποντες προβλεπτικούς παράγοντες απαιτεί ρητή και τεκμηριωμένη στρατηγική, διαφορετικά η πραγματική απόδοση μπορεί να αποκλίνει ουσιωδώς από την απόδοση ανάπτυξης [86].

Συνολικά, οι περιορισμοί της δικής μας μελέτης συγκλίνουν σε τρία συμπεράσματα: (i) το μέγεθος του δείγματος σε σχέση με την πολυπαραγοντική περιγραφή αυξάνει τον κίνδυνο αστάθειας και υπερπροσαρμογής [82], [83], (ii) οι μετρήσεις έναρξης, παρότι κατάλληλες για έγκαιρη πρόγνωση, θέτουν όριο στο σήμα που μπορεί να αξιοποιηθεί [81], και (iii) οι ελλιπείς τιμές σε κρίσιμους δείκτες, όπως το APOE, μειώνουν την προγνωστική ισχύ και απαιτούν αυστηρό χειρισμό ώστε να αποφευχθεί μεροληψία και να διασφαλιστεί δυνατότητα εφαρμογής [84]–[86].

1.8.3. Μελλοντικές Προεκτάσεις στο πλαίσιο του Brain Precision.

Οι μελλοντικές προεκτάσεις του Brain Precision δομούνται σε τρεις συμπληρωματικούς άξονες: (α) **διεύρυνση της πληροφορίας εισόδου** με πολυτροπικά βιολογικά, απεικονιστικά και ψηφιακά δεδομένα, (β) **μεθοδολογική ενίσχυση** της προγνωστικής ισχύος και της ανθεκτικότητας των μοντέλων μέσω προηγμένων τεχνικών μάθησης, και (γ) **προοπτική επικύρωση** με συνεχή ροή δεδομένων από πραγματικές συνθήκες χρήσης, ώστε να τεκμηριώνεται η γενίκευση και η ασφάλεια σε διαχρονικό ορίζοντα [89]–[93].

1.8.3.1. Ενσωμάτωση δεδομένων πολλαπλών τύπων (Multi-omics, MRI, ψηφιακοί βιοδείκτες)

(α) **Πολυ-ομικά δεδομένα: από μεμονωμένους δείκτες σε μοριακές «υπογραφές».** Η ενσωμάτωση πολυ-ομικών επιπέδων (π.χ. μεταβολομική, πρωτεομική, επιγενετική/μεθυλίωση,

μεταγραφωμική) είναι κομβική, διότι η νευροεκφύλιση αντανακλάται σε **συστημικές διαταραχές δικτύων** και όχι σε έναν μόνο βιοδείκτη. Μελέτες πολυ-ομικής ολοκλήρωσης έχουν δείξει ότι η συνδυαστική μοντελοποίηση μπορεί να παράγει υπογραφές με αυξημένη ειδικότητα/ευαισθησία για πρώιμες καταστάσεις υψηλού κινδύνου, σε σύγκριση με μονο-επίπεδα σήματα [89], [90]. Επιπλέον, πρόσφατες εργασίες σε πληθυσμούς τύπου ADNI καταδεικνύουν ότι πολυ-ομικές αναλύσεις στο αίμα αναδεικνύουν μεταβολίτες/μονοπάτια που συσχετίζονται με την παθοφυσιολογία της νόσου και μπορούν να αξιοποιηθούν για προγνωστική χαρτογράφηση [91]. Συστηματικές ανασκοπήσεις επισημαίνουν ότι η κύρια πρόκληση δεν είναι μόνο η επιλογή «καλύτερων» βιοδεικτών, αλλά η **ορθή ολοκλήρωση** ετερογενών επιπέδων με έλεγχο θορύβου, ετερογένειας παρτίδων (batch effects) και υψηλής διαστασιμότητας [92].

Από μεθοδολογική σκοπιά, η προτεινόμενη εξέλιξη στο Brain Precision μπορεί να υλοποιηθεί σε τρία επίπεδα ολοκλήρωσης:

1. **όψιμη σύντηξη** (συνδυασμός προβλέψεων από επιμέρους μοντέλα ανά τύπο δεδομένων),
2. **ενδιάμεση σύντηξη** (κοινός λανθάνων χώρος/αναπαράσταση από πολλαπλούς τύπους δεδομένων) και
3. **πρώιμη σύντηξη** (ενοποίηση χαρακτηριστικών μετά από ισχυρή κανονικοποίηση και επιλογή χαρακτηριστικών). Στην πράξη, η επιλογή μεταξύ αυτών καθορίζεται από τη διαθεσιμότητα δειγμάτων, την ισορροπία μεταξύ επεξηγηματικότητας και προγνωστικής ισχύος, καθώς και από τον βαθμό ελλείψεων ανά τύπο δεδομένων [92].

(β) Μαγνητική τομογραφία: απεικονιστικοί φαινότυποι και πρόγνωση μετατροπής. Η μαγνητική τομογραφία (MRI) επιτρέπει την ποσοτικοποίηση δομικών και λειτουργικών προτύπων που σχετίζονται με πρώιμες νευροεκφυλιστικές διεργασίες, όπως ατροφία, μεταβολές φλοιϊκού πάχους και δείκτες λευκής ουσίας. Η σύγχρονη βιβλιογραφία υπογραμμίζει ότι η αξιοποίηση απεικονιστικών δεδομένων με τεχνικές τεχνητής νοημοσύνης βελτιώνει την πρόβλεψη εξέλιξης ή μετατροπής ήπιας γνωστικής διαταραχής προς άνοια, ιδίως όταν συνδυάζεται με κλινικά/βιοχημικά δεδομένα [93]–[95]. Πρόσφατες ανασκοπήσεις επισημαίνουν ότι η πολυτροπική απεικόνιση και η συστηματική επεξεργασία (π.χ. τυποποίηση πρωτοκόλλων, εξαγωγή χαρακτηριστικών τύπου ραδιομικής, ή βαθιά μάθηση σε T1-ακολουθίες) μπορούν να ενισχύσουν την προγνωστική ισχύ, αλλά απαιτούν αυστηρή διαχείριση ετερογένειας κέντρων και ανεξάρτητη επικύρωση [93], [94].

(γ) Ψηφιακοί βιοδείκτες: συνεχής φαινοτυπική αποτύπωση λειτουργικότητας. Οι ψηφιακοί βιοδείκτες από φορητές συσκευές/κινητές εφαρμογές επιτρέπουν τη συνεχή αποτύπωση της καθημερινής λειτουργικότητας (π.χ. κινητικότητα, βάδισμα, ύπνος, μοτίβα δραστηριότητας), η οποία συχνά μεταβάλλεται πριν από εμφανείς κλινικές εκδηλώσεις. Συστηματικές ανασκοπήσεις τεκμηριώνουν ότι δείκτες πλοήγησης και βάδισης, καθώς και παθητικά συλλεγόμενες συμπεριφορικές μετρήσεις, είναι πολλά υποσχόμενοι για πρώιμη ανίχνευση παθολογιών που οδηγούν σε άνοια, υπό την προϋπόθεση τυποποίησης και κλινικής ερμηνείας [96]. Αντίστοιχα, νεότερες εργασίες υπογραμμίζουν ότι η παθητική και συνεχής μέτρηση μέσω καταναλωτικών συσκευών μπορεί να προβλέπει πολλαπλές εκβάσεις εγκεφαλικής υγείας σε πραγματικές συνθήκες, ενισχύοντας την προοπτική μεταφοράς σε κλινικές ροές [97]. Σε επίπεδο σχεδιασμού, η ενσωμάτωση τέτοιων δεικτών στο Brain Precision δύναται να μετατρέψει

τη πρόγνωση από «στιγμαία» σε «διαχρονική», μέσω μοντέλων που αξιοποιούν μεταβολές/τροχιές και όχι μόνο τιμές έναρξης [96], [97].

1.8.3.2. Εφαρμογή προηγμένων τεχνικών (SMOTE, Ensemble learning)

(α) Αντιμετώπιση ανισορροπίας κλάσεων με συνθετική υπερ-δειγματοληψία

Σε προβλήματα πρόβλεψης όπου η κλάση υψηλού κινδύνου είναι μειονοτική, η απόδοση των μοντέλων μπορεί να υποβαθμίζεται λόγω προτίμησης προς την πλειονοτική κλάση. Η τεχνική συνθετικής υπερ-δειγματοληψίας (γνωστή ευρέως ως SMOTE) δημιουργεί τεχνητά παραδείγματα της μειονοτικής κλάσης με παρεμβολή στο χώρο χαρακτηριστικών και έχει αποδειχθεί ότι βελτιώνει τη δυνατότητα ανίχνευσης της μειονοτικής κλάσης σε ποικίλα περιβάλλοντα ταξινόμησης [98]. Για το Brain Precision, η κρίσιμη μεθοδολογική προϋπόθεση είναι η εφαρμογή της υπερ-δειγματοληψίας **αποκλειστικά** εντός του συνόλου εκπαίδευσης και εντός κάθε πτυχής διασταυρούμενης επικύρωσης, ώστε να αποφεύγεται διαρροή πληροφορίας και τεχνητή διόγκωση μετρικών.

(β) Σύνολα μοντέλων: μείωση σφάλματος και αύξηση γενίκευσης

Τα σύνολα μοντέλων στοχεύουν στη μείωση σφάλματος μέσω συνδυασμού πολλαπλών «ασθενέστερων» ή συμπληρωματικών μοντέλων. Ανασκοπήσεις στο πεδίο της ιατρικής πρόβλεψης δείχνουν ότι η χρήση τεχνικών όπως ομαδοποίηση μέσω δειγματοληψίας (bagging), ενίσχυση (boosting) και στοίβαξη (stacking) μπορεί να βελτιώσει τη γενίκευση, ιδίως όταν συνυπάρχουν θόρυβος, ετερογένεια και μη γραμμικές σχέσεις [99]. Κλασικοί αλγόριθμοι όπως τα τυχαία δάση έχουν τεκμηριωμένη ανθεκτικότητα σε ετερογενή χαρακτηριστικά και αλληλεπιδράσεις [100], ενώ η ενισχυτική δένδροειδής μάθηση με κανονικοποίηση (π.χ. XGBoost) έχει καθιερωθεί για υψηλή προγνωστική επίδοση σε δομημένα κλινικά δεδομένα, ιδίως όταν εφαρμόζονται κατάλληλες τεχνικές ρύθμισης πολυπλοκότητας [101].

(γ) Στοίβαξη/υπερ-μαθητής: συνδυασμός ετερογενών μοντέλων με θεωρητικό υπόβαθρο

Η στοίβαξη, ως σχήμα συνδυασμού προβλέψεων, εισήχθη ως «στοιβαγμένη γενίκευση» και επιτρέπει τη μάθηση μετα-μοντέλου που αξιοποιεί τις προβλέψεις πολλαπλών βασικών μοντέλων [102]. Η μεθοδολογία «υπερ-μαθητή» (super learner) ενισχύει περαιτέρω αυτή τη λογική με θεωρητική θεμελίωση μέσω διασταυρούμενης επικύρωσης για τον υπολογισμό βαρών συνδυασμού, προσφέροντας πρακτικό μηχανισμό για βελτιστοποίηση του σφάλματος πρόβλεψης σε πεδία υψηλής ετερογένειας, όπως τα βιοϊατρικά δεδομένα [103]. Για το Brain Precision, μια τέτοια προσέγγιση είναι ιδιαίτερα ελκυστική επειδή επιτρέπει:

- συνύπαρξη γραμμικών (ερμηνεύσιμων) και μη γραμμικών μοντέλων,
- προσαρμογή σε διαφορετικές κατανομές χαρακτηριστικών,
- και σταδιακή ενσωμάτωση νέων τύπων δεδομένων χωρίς πλήρη ανασχεδίαση του συστήματος.

1.8.3.3. Προοπτική επικύρωσης σε πραγματικό χρόνο μέσω του συστήματος LETHE

Η προοπτική επικύρωση σε πραγματικές συνθήκες, με συνεχή ροή δεδομένων, αποτελεί το ουσιώδες επόμενο βήμα ώστε η προγνωστική απόδοση να μετατραπεί σε **τεκμηριωμένη κλινική αξιοπιστία**. Το LETHE έχει ήδη σχεδιαστεί ως διετής, πολυκεντρική τυχαιοποιημένη δοκιμή εφικτότητας με ψηφιακή υποστήριξη πολυπαραγοντικής παρέμβασης, δημιουργώντας το κατάλληλο περιβάλλον για προοπτική αξιολόγηση [104]. Στο Brain Precision, η επικύρωση «σε πραγματικό χρόνο» (δηλαδή με συνεχή εισροή νέων δεδομένων και περιοδική/διαρκή ανανέωση της αξιολόγησης) μπορεί να οργανωθεί σε τέσσερα επίπεδα:

- 1. Εξωτερική επικύρωση σε ανεξάρτητα κέντρα/χρονικές περιόδους.**
Η απόδοση πρέπει να επιβεβαιώνεται σε δεδομένα που δεν χρησιμοποιήθηκαν σε κανένα στάδιο ανάπτυξης, με διατήρηση της ίδιας διαδικασίας προ-επεξεργασίας και με ελέγχους βαθμονόμησης. Αυτό είναι κρίσιμο σε πολυκεντρικά περιβάλλοντα, όπου η ετερογένεια μετρήσεων και πληθυσμών επηρεάζει τη γενίκευση [104].
- 2. Διαχρονικός έλεγχος μετατόπισης δεδομένων και εννοιολογικής μεταβολής.**
Με συνεχή συλλογή, οι κατανομές χαρακτηριστικών και η σχέση χαρακτηριστικών-έκβασης ενδέχεται να αλλάζουν (π.χ. λόγω αλλαγών στη συμμόρφωση, στις συσκευές μέτρησης ή στη σύνθεση του πληθυσμού). Η προοπτική επικύρωση πρέπει να περιλαμβάνει διαδικασίες ανίχνευσης μετατόπισης, ώστε να αποφεύγεται «σιωπηρή» υποβάθμιση του μοντέλου.
- 3. Συνεχής αξιολόγηση με κλινικά ερμηνεύσιμες μετρικές και έλεγχο βαθμονόμησης.**
Η περιοδική αναφορά ευαισθησίας, θετικής προγνωστικής τιμής και βαθμονόμησης σε πραγματικές ροές είναι απαραίτητη για ασφαλή χρήση, ιδίως όταν η έξοδος αποσκοπεί σε ιεράρχηση κινδύνου.
- 4. Προσαρμοστική ενημέρωση μοντέλων με διατήρηση ιχνηλασιμότητας.**
Εφόσον υιοθετηθούν διαδικασίες ανανέωσης, αυτές πρέπει να γίνονται με αυστηρή τεκμηρίωση εκδόσεων, επαναληψιμότητα εκπαίδευσης, και σαφή διαχωρισμό μεταξύ φάσης ανάπτυξης και φάσης αξιολόγησης, ώστε η προοπτική απόδειξη να παραμένει έγκυρη.

Συνολικά, το LETHE παρέχει το λειτουργικό και ερευνητικό υπόβαθρο ώστε το Brain Precision να μεταβεί από ανάπτυξη μοντέλου σε προοπτικά τεκμηριωμένη λύση, υπό την προϋπόθεση ότι η επικύρωση σχεδιαστεί ως συνεχής διαδικασία ποιοτικού ελέγχου, βαθμονόμησης και εξωτερικής γενίκευσης [104].

BIBΛΙΟΓΡΑΦΙΑ

- [1] G. Livingston et al., “Dementia prevention, intervention, and care: 2024 report of the Lancet standing Commission,” *The Lancet*, vol. 404, pp. 572–628, Aug. 2024, doi: 10.1016/S0140-6736(24)01296-0.
- [2] C. R. Jack Jr. et al., “NIA-AA Research Framework: Toward a biological definition of Alzheimer’s disease,” *Alzheimer’s & Dementia*, vol. 14, no. 4, pp. 535–562, 2018, doi: 10.1016/j.jalz.2018.02.018.
- [3] W. Poewe et al., “Parkinson disease,” *Nature Reviews Disease Primers*, vol. 3, Art. no. 17013, 2017, doi: 10.1038/nrdp.2017.13.
- [4] T. F. Outeiro et al., “Alpha-synuclein and synucleinopathies: A comprehensive review,” *Cold Spring Harbor Perspectives in Medicine*, vol. 9, no. 1, 2019.
- [5] R. Brookmeyer and N. Abdalla, “Estimation of lifetime risks of Alzheimer’s disease dementia using biomarkers for preclinical disease,” *Alzheimer’s & Dementia*, vol. 14, no. 8, pp. 981–988, 2018, doi: 10.1016/j.jalz.2018.02.018.

- [6] G. Grande, C. Qiu, and L. Fratiglioni, “Prevention of dementia in an ageing world: Evidence and biological rationale,” *Ageing Research Reviews*, vol. 64, p. 101045, 2020, doi: 10.1016/j.arr.2020.101045.
- [7] C. R. Jack Jr. et al., “A/T/N: An unbiased descriptive classification scheme for Alzheimer disease biomarkers,” *Neurology*, vol. 87, no. 5, pp. 539–547, Aug. 2016, doi: 10.1212/WNL.0000000000002923.
- [8] N. A. Ashton et al., “Plasma biomarkers of Alzheimer’s disease: A review,” *Nature Reviews Neurology*, vol. 17, no. 12, pp. 805–822, 2021.
- [9] J. B. Pereira et al., “Plasma GFAP is an early marker of amyloid- β but not tau pathology in Alzheimer’s disease,” *Brain*, vol. 144, no. 11, pp. 3505–3516, Nov. 2021, doi: 10.1093/brain/awab223.
- [10] G. Salvadó et al., “Optimal combinations of CSF biomarkers for predicting cognitive decline and clinical conversion in cognitively unimpaired participants and mild cognitive impairment patients: a multi-cohort study,” *Alzheimer’s & Dementia*, vol. 19, no. 7, pp. 2943–2955, 2023, doi: 10.1002/alz.12936.
- [11] L. C. Kourtis, O. B. Regele, J. M. Wright, and G. B. Jones, “Digital biomarkers for Alzheimer’s disease: the mobile/wearable devices opportunity,” *npj Digital Medicine*, vol. 2, Art. no. 9, 2019, doi: 10.1038/s41746-019-0089-9.
- [12] F. Thabtah, D. Peebles, J. Retzler, and C. Hathurusingha, “A review of dementia screening tools based on mobile application,” *Health and Technology*, vol. 10, pp. 1011–1020, 2020, doi: 10.1007/s12553-020-00449-7.
- [13] “Εθνικό δίκτυο έρευνας για την ανάδειξη της γενετικής βάσης των νευροεκφυλιστικών νόσων Alzheimer και Parkinson... ‘Brain Precision’,” brainprecision.gr, 2026. .
- [14] Γενική Γραμματεία Έρευνας και Καινοτομίας, “Εθνικό Δίκτυο Ιατρικής Ακρίβειας για Νευροεκφυλιστικά Νοσήματα: Καινοτόμες δράσεις στην υπηρεσία των πολιτών,” gsri.gov.gr, 2024.
- [15] European Union, “Regulation (EU) 2016/679 (General Data Protection Regulation),” *Official Journal of the European Union*, Apr. 27, 2016.
- [16] European Union, “Regulation (EU) 2025/327 on the European Health Data Space,” Feb. 11, 2025.
- [17] ISO, “ISO/IEC 27001:2022 — Information security management systems,” iso.org, 2022.
- [18] M. D. Wilkinson et al., “The FAIR Guiding Principles for scientific data management and stewardship,” *Scientific Data*, vol. 3, Art. no. 160018, 2016, doi: 10.1038/sdata.2016.18.
- [19] HL7, “FHIR Overview (Fast Healthcare Interoperability Resources),” hl7.org, 2024.
- [20] OHDSI, “OMOP Common Data Model,” ohdsi.github.io, 2024.
- [21] GA4GH, “Data Use Ontology (DUO),” ga4gh.org, 2024.
- [22] J. Lawson et al., “The Data Use Ontology to streamline responsible access to human biomedical datasets,” *Cell Genomics*, vol. 1, no. 2, 2021.
- [23] G. S. Collins et al., “TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods,” *BMJ*, vol. 385, p. e078378, 2024, doi: 10.1136/bmj-2023-078378.
- [24] K. G. M. Moons et al., “PROBAST+AI: an updated quality, risk of bias, and applicability assessment tool for prediction models using regression or artificial intelligence methods,” *BMJ*, vol. 388, p. e082505, 2025, doi: 10.1136/bmj-2024-082505.
- [25] G. S. Collins, J. B. Reitsma, D. G. Altman, and K. G. M. Moons, “Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): The TRIPOD Statement,” *Ann. Intern. Med.*, vol. 162, no. 1, pp. 55–63, 2015.
- [26] R. F. Wolff et al., “PROBAST: A Tool to Assess the Risk of Bias and Applicability of Prediction Model Studies,” *Ann. Intern. Med.*, vol. 170, no. 1, pp. 51–58, 2019, doi: 10.7326/M18-1377.
- [27] E. W. Steyerberg, *Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating*, 2nd ed. Cham, Switzerland: Springer, 2019.
- [28] E. Christodoulou et al., “A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models,” *J. Clin. Epidemiol.*, vol. 110, pp. 12–22, 2019.
- [29] T. Fawcett, “An introduction to ROC analysis,” *Pattern Recognit. Lett.*, vol. 27, no. 8, pp. 861–874, 2006, doi: 10.1016/j.patrec.2005.10.010.
- [30] E. R. DeLong, D. M. DeLong, and D. L. Clarke-Pearson, “Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach,” *Biometrics*, vol. 44, no. 3, pp. 837–845, 1988.
- [31] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, “SMOTE: Synthetic Minority Over-sampling Technique,” *J. Artif. Intell. Res.*, vol. 16, pp. 321–357, 2002.
- [31] E. W. Steyerberg, *Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating*, 2nd ed. Cham, Switzerland: Springer, 2019.
- [32] G. Livingston et al., “Dementia prevention, intervention, and care: 2024 report of the Lancet standing Commission,” *The Lancet*, vol. 404, pp. 572–628, Aug. 2024.

- [33] I. Guyon and A. Elisseeff, “An introduction to variable and feature selection,” *J. Mach. Learn. Res.*, vol. 3, pp. 1157–1182, 2003.
- [34] A. J. Vickers and E. B. Elkin, “Decision curve analysis: a novel method for evaluating prediction models,” *Med. Decis. Making*, vol. 26, no. 6, pp. 565–574, 2006.
- [35] G. S. Collins *et al.*, “TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods,” *BMJ*, 2024.
- [36] E. W. Steyerberg *et al.*, “Assessing the performance of prediction models: a framework for traditional and novel measures,” *Epidemiology*, vol. 21, no. 1, pp. 128–138, 2010.
- [37] G. S. Collins, J. B. Reitsma, D. G. Altman, and K. G. M. Moons, “Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): The TRIPOD Statement,” *Ann. Intern. Med.*, vol. 162, no. 1, pp. 55–63, 2015.
- [38] K. G. M. Moons *et al.*, “PROBAST: A tool to assess risk of bias and applicability of prediction model studies,” *Ann. Intern. Med.*, vol. 170, no. 1, pp. 51–58, 2019.
- [39] J. A. C. Sterne *et al.*, “Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls,” *BMJ*, vol. 338, p. b2393, 2009.
- [40] L. Wynants *et al.*, “Three myths about risk thresholds for prediction models,” *BMC Med.*, vol. 17, p. 192, 2019.
- [41] K. G. M. Moons *et al.*, “PROBAST+AI: an updated tool to assess risk of bias and applicability of prediction models using regression or artificial intelligence methods,” *BMJ*, 2025.
- [42] A. Rosenberg *et al.*, “A digitally supported multimodal lifestyle program to promote brain health among older adults (the LETHE randomized controlled feasibility trial): Study design, progress, and first results,” *Alzheimer’s Research & Therapy*, vol. 16, art. 252, 2024.
- [43] V. S. Loukas *et al.*, “A multimodal lifestyle intervention management system for people at risk of cognitive decline,” in *Proc. IEEE EMBC*, 2023.
- [44] M. D. Wilkinson *et al.*, “The FAIR Guiding Principles for scientific data management and stewardship,” *Scientific Data*, vol. 3, art. 160018, 2016.
- [45] G. Hripcsak *et al.*, “Observational Health Data Sciences and Informatics (OHDSI): Opportunities for observational researchers,” *Studies in Health Technology and Informatics*, vol. 216, pp. 574–578, 2015.
- [45] G. Livingston *et al.*, “Dementia prevention, intervention, and care: 2024 report of the Lancet standing Commission,” *The Lancet*, vol. 404, pp. 572–628, Aug. 2024, doi: 10.1016/S0140-6736(24)01296-0.
- [46] A. Rosenberg *et al.*, “A digitally supported multimodal lifestyle program to promote brain health among older adults (the LETHE randomized controlled feasibility trial): Study design, progress, and first results,” *Alzheimer’s Research & Therapy*, vol. 16, art. 252, 2024, doi: 10.1186/s13195-024-01615-4.
- [47] S. L. Gardener *et al.*, “The AUstralian multidomain Approach to Reduce dementia Risk by prOtecting brain health With lifestyle intervention study (AU-ARROW): A study protocol for a single-blind, multi-site, randomized controlled trial,” *Alzheimer’s & Dementia: Translational Research & Clinical Interventions*, vol. 10, art. e12466, 2024, doi: 10.1002/trc2.12466.
- [48] L. Waterink *et al.*, “Evaluation of efficiency and effectiveness of different recruitment strategies for the FINGER-NL multidomain lifestyle intervention trial via the Dutch Brain Research Registry,” *Alzheimer’s & Dementia: Translational Research & Clinical Interventions*, vol. 11, art. e70017, 2025, doi: 10.1002/trc2.70017.
- [49] Zaridis, D.I.; Pezoulas, V.C.; Mylona, E.; Kalantzopoulos, C.N.; Tachos, N.S.; Tsiknakis, N.; Matsopoulos, G.K.; Regge, D.; Papanikolaou, N.; Tsiknakis, M.; *et al.* Simplatab: An Automated Machine Learning Framework for Radiomics-Based Bi-Parametric MRI Detection of Clinically Significant Prostate Cancer. *Bioengineering* 2025, 12, 242. <https://doi.org/10.3390/bioengineering12030242>
- [49] T. Ngandu *et al.*, “A 2 year multidomain intervention of diet, exercise, cognitive training, and vascular risk monitoring versus control to prevent cognitive decline in at-risk elderly people (FINGER): a randomised controlled trial,” *The Lancet*, vol. 385, no. 9984, pp. 2255–2263, 2015.
- [50] J. Harrison *et al.*, “A Neuropsychological Test Battery for Use in Alzheimer Disease Clinical Trials,” *Archives of Neurology (JAMA Network)*, vol. 64, no. 9, pp. 1323–1329, 2007.
- [51] J. E. Harrison *et al.*, “Psychometric Evaluation of the Neuropsychological Test Battery (NTB),” *Journal of Alzheimer’s Disease* (via PubMed Central), 2018.
- [52] E. W. Steyerberg, *Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating*, 2nd ed. Cham, Switzerland: Springer, 2019.
- [53] K. Duff, “Evidence-Based Indicators of Neuropsychological Change in the Individual Patient: Relevant Concepts and Methods,” *Archives of Clinical Neuropsychology*, vol. 27, no. 3, pp. 248–261, 2012.
- [54] K. Duff, “Predicting cognitive change in older adults,” *Archives of Clinical Neuropsychology*, vol. 25, no. 2, pp. 81–92, 2010.

- [55] H. Demnitz-King et al., “A multicentre, single-masked, randomised controlled trial,” *The Lancet Healthy Longevity*, 2025.
- [56] D. B. Rubin, “Inference and missing data,” *Biometrika*, vol. 63, no. 3, pp. 581–592, 1976.
- [57] R. J. A. Little and D. B. Rubin, “Statistical analysis with missing data,” *J. Wiley & Sons* (Foundations and Trends in Statistics cited widely; core missing-data framework), 2002.
- [58] J. A. C. Sterne et al., “Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls,” *BMJ*, vol. 338, p. b2393, 2009.
- [59] S. Hancock and T. M. Khoshgoftaar, “Survey on categorical data for neural networks,” *J. Big Data*, vol. 7, art. 28, 2020.
- [60] A. C. A. N. de Lima, A. L. I. Oliveira, and A. J. M. Traina, “Encoding categorical data for machine learning: a comparative study,” *Expert Systems with Applications*, vol. 165, art. 113970, 2021.
- [61] D. Micci-Barreca, “A preprocessing scheme for high-cardinality categorical attributes in classification and prediction problems,” *SIGKDD Explorations*, vol. 3, no. 1, pp. 27–32, 2001.
- [62] P. Cerda, G. Varoquaux, and B. Kégl, “Similarity encoding for learning with dirty categorical variables,” *Mach. Learn.*, vol. 107, pp. 1477–1494, 2018.
- [63] S. van Buuren, “Flexible imputation of missing data,” *Chapman & Hall/CRC* (methodological reference; included here only for completeness), 2018.
- [64] L. Wynants et al., “Three myths about risk thresholds for prediction models,” *BMC Medicine*, vol. 17, art. 192, 2019.
- [65] D. J. Stekhoven and P. Bühlmann, “MissForest—non-parametric missing value imputation for mixed-type data,” *Bioinformatics*, vol. 28, no. 1, pp. 112–118, 2012.
- [79] T. Saito and M. Rehmsmeier, “The Precision-Recall Plot Is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets,” *PLOS ONE*, vol. 10, no. 3, p. e0118432, 2015.
- [80] D. Chicco and G. Jurman, “The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation,” *BMC Genomics*, vol. 21, art. 6, 2020.
- [81] B. Van Calster et al., “Calibration: the Achilles heel of predictive analytics,” *BMC Medicine*, vol. 17, art. 230, 2019.
- [82] R. D. Riley et al., “Minimum sample size for developing a multivariable prediction model: Part II—binary and time-to-event outcomes,” *Statistics in Medicine*, vol. 38, no. 7, pp. 1276–1296, 2019.
- [83] M. van Smeden et al., “Sample size for binary logistic prediction models: Beyond events per variable criteria,” *Statistical Methods in Medical Research*, vol. 28, no. 8, pp. 2455–2474, 2019.
- [84] I. R. White, P. Royston, and A. M. Wood, “Multiple imputation using chained equations: Issues and guidance for practice,” *Statistics in Medicine*, vol. 30, no. 4, pp. 377–399, 2011.
- [85] J. A. C. Sterne et al., “Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls,” *BMJ*, vol. 338, p. b2393, 2009.
- [86] K. J. M. Janssen et al., “Dealing with missing predictor values when applying clinical prediction models,” *Clinical Chemistry*, vol. 55, no. 5, pp. 994–1004, 2009.
- [87] E. H. Corder et al., “Gene Dose of Apolipoprotein E Type 4 Allele and the Risk of Alzheimer’s Disease in Late Onset Families,” *Science*, vol. 261, no. 5123, pp. 921–923, 1993.
- [88] L. A. Farrer et al., “Effects of Age, Sex, and Ethnicity on the Association Between Apolipoprotein E Genotype and Alzheimer Disease: A Meta-analysis,” *JAMA*, vol. 278, no. 16, pp. 1349–1356, 1997.