



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

**Ανάλυση βιοχημικών φαινοτύπων της καρωτιδικής αθηρωμάτωσης:  
Συσχετίσεις με απεικονιστικούς και κλινικούς δείκτες με τη βοήθεια  
τεχνικών συσταδοποίησης**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ  
ΤΟΥ  
**ΣΕΡΑΦΕΙΜ Λ. ΛΟΥΚΑ**

**Επιβλέπουσα** : Κωνσταντίνα Σπ. Νικήτα

Καθηγήτρια Ε.Μ.Π.

Αθήνα, Ιούλιος 2015





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

**Ανάλυση βιοχημικών φαινοτύπων της καρωτιδικής αθηρωμάτωσης:  
Συσχετίσεις με απεικονιστικούς και κλινικούς δείκτες με τη βοήθεια  
τεχνικών συσταδοποίησης**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ  
ΤΟΥ  
**ΣΕΡΑΦΕΙΜ Λ. ΛΟΥΚΑ**

**Επιβλέπουσα :** Κωνσταντίνα Σπ. Νικήτα

Καθηγήτρια Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 24<sup>η</sup> Ιουλίου, 2015.

.....  
Κωνσταντίνα Νικήτα  
Καθηγήτρια Ε.Μ.Π.

.....  
Γεώργιος Στάμου  
Επίκουρος Καθηγητής Ε.Μ.Π.

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

Αθήνα, Ιούλιος 2015



.....

Σεραφείμ Λ. Λούκας

Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π

Copyright © Σεραφείμ Λούκας, 2015.

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

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

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



## Περίληψη

Η αθηρωμάτωση αποτελεί μια πάθηση εκφυλιστικού χαρακτήρα, η οποία προσβάλλει τις αρτηρίες αλλοιώνοντας το εσωτερικό τους και αποτελεί την κύρια αιτία πρόκλησης ισχαιμικών επεισοδίων. Κύριο διαγνωστικό μέσο αποτελεί το υπερηχογράφημα, το οποίο παρέχει λεπτομερείς απεικονίσεις της καρωτίδας. Σκοπός της παρούσας διπλωματικής εργασίας αποτελεί η αναγνώριση φαινοτύπων με βάση τους συμβατικούς και εξειδικευμένους βιοχημικούς δείκτες. Για την ανίχνευση αυτή, συνδυάστηκαν τεχνικές ιεραρχικής συσταδοποίησης και παραγωγής δενδρογράμματος. Η ορθότητα των εξαγόμενων αποτελεσμάτων εκτιμήθηκε μέσω συσχέτισης με απεικονιστικούς και κλινικούς δείκτες. Υπολογίστηκε ο αριθμός των φαινοτύπων με βάση τους συμβατικούς, τους εξειδικευμένους και το συνδυασμό τους. Στη συνέχεια αναλύοντας τις γραφικές απεικονίσεις αυτών των συστάδων σε κάθε περίπτωση εντοπίστηκαν οι δείκτες που παρουσιάζουν μεγάλες διακυμάνσεις μέσα στο δείγμα και συνεπώς αποτελούν τα αίτια διαχωρισμού σε συστάδες. Τέλος, αναλύθηκαν οι μεταβολές των απεικονιστικών δεικτών εντός των φαινοτύπων σε κάθε περίπτωση και συσχετίστηκαν με τους βιοχημικούς.

**Λέξεις κλειδιά:** *Αθηρωματική πλάκα, Αθηροσκλήρωση, Καρωτίδα, Υπέρηχος, Ιεραρχική συσταδοποίηση, Βιοχημικοί δείκτες, Απεικονιστικοί δείκτες.*





## **Abstract**

Atherosclerosis is a degenerative disease that affects the arteries by altering their interior and it constitutes the major cause of ischemic strokes. The main imaging method used for clinical diagnosis is carotid ultrasound that provides detailed illustrations of the carotid artery. The purpose of this diploma thesis is the identification of imaging phenotypes based on conventional and specialized biochemical indicators. For the detection of the phenotypes, hierarchical clustering was combined with the production of the dendrogram and heatmaps. The accuracy of the exported results was estimated through the comparison with image-based and clinical indicators. The number of phenotypes was estimated based on either conventional or specialized biochemical indicators and then based on their combination. Then by analyzing the graphic illustrations in each case the indicators that show large variations in the sample were calculated and thus constitute the main cause of separation in clusters. Finally, the variations of image-based indicators in each cluster and each case were correlated with the biochemical indicators.

**Keywords:** *Atheromatous plaque, Atherosclerosis, Carotid, Ultrasound, Hierarchical Clustering, Biochemical markers, Imaging Biomarkers.*

*Ευχαριστώ θερμά την οικογένειά μου και τη Σοφία Κώστογλου για την υπομονή και τη στήριξη.*

*Ευχαριστώ την καθηγήτρια μου Κ. Σπ. Νικήτα που μου έδωσε το κίνητρο μέσω των μαθημάτων  
κα των διαλέξεών της, να ασχοληθώ εμπράκτως με τον τομέα της βιοιατρικής.*

*Ευχαριστώ ακόμη τη διδάκτορα Αιμιλία Γαστουνιώτη για την άμεση βοήθεια σε κάθε βήμα της  
παρούσας διπλωματικής εργασίας.*

# Περιεχόμενα

Κεφάλαιο 1: Εισαγωγή.....	1
1.1 Στοιχεία φυσιολογίας της καρωτίδας.....	2
1.2 Η νόσος της καρωτιδικής αθηρωμάτωσης.....	4
1.3 Συσχέτιση βιοχημικών δεικτών με την παρουσία και εξέλιξη της νόσου.....	9
1.4 Ο ρόλος και η συνεισφορά του υπερήχου.....	16
1.4.1 Απεικονιστικές μέθοδοι.....	16
1.4.2 Απεικόνιση καρωτίδας με χρήση υπερήχων.....	17
1.4.3 Ανάλυση εικόνων υπερήχου.....	20
1.4.4 Καθιερωμένοι και ευρέως χρησιμοποιούμενοι απεικονιστικοί δείκτες.....	25
1.5 Σκοπός της παρούσας εργασίας.....	26
Κεφάλαιο 2 : Μεθοδολογία.....	28
2.1 Θεωρητικό υπόβαθρο του αλγορίθμου.....	29
2.1.1 Ανάλυση κατά συστάδες.....	29
2.1.2 Μέτρα Απόστασης.....	31
2.1.3 Μέτρα Ομοιότητας.....	35
2.1.4 Κατηγορίες ιεραρχικών μεθόδων συσταδοποίησης.....	36
2.1.5 Συσσωρευτικές ιεραρχικές μέθοδοι ομαδοποίησης.....	37
2.2 Σχεδιασμός της μελέτης.....	42
2.2.1 Περιγραφή των δεδομένων.....	42
2.2.2 Περιγραφή της μεθόδου.....	47
Κεφάλαιο 3 : Αποτελέσματα.....	52
3.1 Χρήση συμβατικών βιοχημικών δεικτών.....	53
3.2 Χρήση εξειδικευμένων βιοχημικών δεικτών.....	64
3.3 Συνδυασμός συμβατικών και εξειδικευμένων δεικτών.....	74
Κεφάλαιο 4 : Συζήτηση.....	88
4.1 Συμπεράσματα.....	89
4.2 Επίλογος και μελλοντικές προεκτάσεις.....	92
Βιβλιογραφία.....	95

# Κατάλογος Εικόνων

Εικόνα 1.1 : Σχηματική αναπαράσταση της φυσιολογίας της καρωτίδας.....	2
Εικόνα 1.2 : Αναπαράσταση δομής και σύστασης του αρτηριακού τοιχώματος.....	3
Εικόνα 1.3 : Απεικόνιση υγιούς αρτηρίας και αρτηρίας με αθηρωματική πλάκα.....	4
Εικόνα 1.4 : Η εξέλιξη της αθηρωματικής πλάκας. ....	7
Εικόνα 1.5 : Ισχαμικό επεισόδιο λόγω θρόμβωσης.....	8
Εικόνα 1.6 : Το υπερηχογράφημα της καρωτίδας.....	16
Εικόνα 1.7 : Ο μετατροπέας υπερήχων.....	17
Εικόνα 1.8 : Τα διαφορετικά είδη συστοιχιών του μετατροπέα υπερήχων.....	18
Εικόνα 1.9 : Στιγμιότυπο κοινής καρωτίδας από β-σάρωση υπερήχου.....	19
Εικόνα 1.10 : Εντοπισμός της αθηρωματικής πλάκας σε εικόνα υπερήχου β-σάρωσης.....	22
Εικόνα 1.11 : Ανάλυση του πάχους του ενδιάμεσου στρώματος (IMT) από τρεις παρατηρητές A, B και C.....	22
Εικόνα 2.1 : Συσσωρευτική και διαιρετική ομαδοποίηση.....	37
Εικόνα 2.2 : Διάγραμμα ροής συσσωρευτικού αλγορίθμου συσταδοποίησης.....	38
Εικόνα 2.3 : Απόσταση ομάδων βάσει των τριών μεθόδων.....	40
Εικόνα 2.4 : Μέθοδος κεντροειδούς σύνδεσης.....	41
Εικόνα 2.5 : Ιστογράμματα τεσσάρων συμβατικών βιοχημικών δεικτών.....	46
Εικόνα 2.6 : Ιστογράμματα τεσσάρων εξειδικευμένων βιοχημικών δεικτών.....	47
Εικόνα 2.7 : Διάγραμμα ροής της μεθοδολογίας.....	48
Εικόνα 3.1 : Γραφική αναπαράσταση του δενδρογράμματος των συμβατικών βιοχημικών δεικτών.....	53
Εικόνα 3.2 : Γραφική αναπαράσταση μεγεθυμένου τμήματος του δενδρογράμματος των συμβατικών βιοχημικών δεικτών.....	54
Εικόνα 3.3 : Γραφική παράσταση μέσης απόστασης των στοιχείων μεταξύ των ομάδων συναρτήσει του πλήθους των ομάδων για την εύρεση του βέλτιστου αριθμού συστάδων.....	54

Εικόνα 3.4 : Το heatmap που αντιστοιχεί στην συσταδοποίηση βάσει των συμβατικών βιοχημικών δεικτών.....	55
Εικόνα 3.5 : Απεικονιστικά χαρακτηριστικά των συστάδων.....	56
Εικόνα 3.6 : Βoxplot για το συστολικό GSM κατά τη συσταδοποίηση με βάση τους συμβατικούς βιοχημικούς δείκτες.....	56
Εικόνα 3.7 : Βoxplot για το διαστολικό GSM κατά τη συσταδοποίηση με βάση τους συμβατικούς βιοχημικούς δείκτες. ....	57
Εικόνα 3.8 : Βoxplot για το εμβαδόν της καρωτίδας κατά τη συσταδοποίηση με βάση τους συμβατικούς βιοχημικούς δείκτες. ....	57
Εικόνα 3.9 : Το boxplot του συμβατικού δείκτη της Ουρίας.....	60
Εικόνα 3.10 : Το boxplot του συμβατικού δείκτη των λεμφοκυττάρων. ....	61
Εικόνα 3.11 : Βoxplot των τριγλυκεριδίων. ....	62
Εικόνα 3.12 : Βoxplots των τριγλυκεριδίων και της γλυκόζης. ....	63
Εικόνα 3.13 : Γραφική αναπαράσταση του δενδρογράμματος των εξειδικευμένων βιοχημικών δεικτών.....	64
Εικόνα 3.14 : Μεγεθυμένο τμήμα του δενδρογράμματος των εξειδικευμένων βιοχημικών δεικτών.....	65
Εικόνα 3.15 : Γραφική παράσταση μέσης απόστασης των στοιχείων μεταξύ των ομάδων συναρτήσει του πλήθους των ομάδων για την εύρεση του βέλτιστου αριθμού συστάδων. ....	65
Εικόνα 3.16 : Το heatmap που αντιστοιχεί στην συσταδοποίηση βάσει των εξειδικευμένων βιοχημικών δεικτών.....	66
Εικόνα 3.17 : Απεικονιστικά χαρακτηριστικά των συστάδων.....	67
Εικόνα 3.18 : Βoxplot για το συστολικό GSM κατά τη συσταδοποίηση με βάση τους εξειδικευμένους βιοχημικούς δείκτες.....	67
Εικόνα 3.19 : Βoxplot για το διαστολικό GSM κατά τη συσταδοποίηση με βάση τους εξειδικευμένους βιοχημικούς δείκτες.....	68
Εικόνα 3.20 : Βoxplot για το διαστολικό GSM κατά τη συσταδοποίηση με βάση τους εξειδικευμένους βιοχημικούς δείκτες.....	68

Εικόνα 3.21 : Το boxplot του εξειδικευμένου δείκτη της ινσουλίνης.....	71
Εικόνα 3.22 : Το boxplot του εξειδικευμένου δείκτη RBP4. ....	71
Εικόνα 3.23 : Boxplot της Γαλεκτίνης-3. ....	72
Εικόνα 3.24 : Γραφική αναπαράσταση του δενδρογράμματος των συμβατικών και εξειδικευμένων βιοχημικών δεικτών. ....	74
Εικόνα 3.25 : Μεγεθυμένο τμήμα του δενδρογράμματος των εξειδικευμένων και συμβατικών βιοχημικών δεικτών.....	75
Εικόνα 3.26 : Γραφική παράσταση μέσης απόστασης των στοιχείων μεταξύ των ομάδων συναρτήσει του πλήθους των ομάδων για την εύρεση του βέλτιστου αριθμού συστάδων. ....	75
Εικόνα 3.27 : Το heatmap που αντιστοιχεί στην συσταδοποίηση βάσει των συμβατικών και των εξειδικευμένων βιοχημικών δεικτών. ....	76
Εικόνα 3.28 : Απεικονιστικά χαρακτηριστικά των συστάδων .....	77
Εικόνα 3.29 : Boxplot για το συστολικό GSM κατά τη συσταδοποίηση με βάση τους εξειδικευμένους και συμβατικούς βιοχημικούς δείκτες. ....	77
Εικόνα 3.30 : Boxplot για το διαστολικό GSM κατά τη συσταδοποίηση με βάση τους εξειδικευμένους και συμβατικούς βιοχημικούς δείκτες. ....	78
Εικόνα 3.31 : Boxplot για το εμβαδόν της καρωτίδας κατά τη συσταδοποίηση με βάση τους εξειδικευμένους και συμβατικούς βιοχημικούς δείκτες. ....	78
Εικόνα 3.32 : Boxplot της Αιμοσφαιρίνης. ....	82
Εικόνα 3.33 : Boxplot των Λεμφοκυττάρων. ....	82
Εικόνα 3.34 : Boxplot της Γαλεκτίνης-3. ....	83
Εικόνα 3.35 : Boxplot των Τριγλυκεριδίων.....	84
Εικόνα 3.36 : Boxplot του TMP1. ....	85
Εικόνα 3.37 : Boxplot του RBP4. ....	86

# Κατάλογος Πινάκων

Πίνακας 1.1: Κατηγοριοποίηση των τύπων της αθηρωματικής πλάκας και η πορεία εξέλιξης της νόσου. ....	6
Πίνακας 1.2 : Συνοπτική παρουσίαση εργασιών που συσχέτισαν βιοχημικούς δείκτες με τη νόσο της αθηρωμάτωσης. ....	13
Πίνακας 2.1 : Συνοπτική παρουσίαση του προφίλ των ασθενών. ....	43
Πίνακας 2.2 : Πλήρης ονομασία και συντομογραφία συμβατικών βιοχημικών δεικτών. ....	44
Πίνακας 2.3 : Πλήρης ονομασία και συντομογραφία εξειδικευμένων βιοχημικών δεικτών. ....	45
Πίνακας 2.4 : Οι απεικονιστικοί δείκτες. ....	46
Πίνακας 3.1 : Πίνακας των ποσοστών των ασθενών κάθε φαινοτύπου στις κατηγορίες υψηλή, μεσαία και χαμηλή. ....	58
Πίνακας 3.2 : Πίνακας των ποσοστών των ασθενών κάθε φαινοτύπου στις κατηγορίες υψηλή, μεσαία και χαμηλή. ....	58
Πίνακας 3.3 : Πίνακας των ποσοστών των ασθενών κάθε φαινοτύπου στις κατηγορίες υψηλή, μεσαία και χαμηλή. ....	59
Πίνακας 3.4 : Οι συμβατικοί βιοχημικοί δείκτες που παρουσιάζουν τη μεγαλύτερη συσχέτιση με το GSM διαστολής. ....	62
Πίνακας 3.5 : Οι συμβατικοί βιοχημικοί δείκτες που παρουσιάζουν τη μεγαλύτερη συσχέτιση με το GSM διαστολής. ....	63
Πίνακας 3.6 : Πίνακας των ποσοστών των ασθενών κάθε φαινοτύπου στις κατηγορίες υψηλή, μεσαία και χαμηλή. ....	69
Πίνακας 3.7 : Πίνακας των ποσοστών των ασθενών κάθε φαινοτύπου στις κατηγορίες υψηλή, μεσαία και χαμηλή. ....	69
Πίνακας 3.8 : Πίνακας των ποσοστών των ασθενών κάθε φαινοτύπου στις κατηγορίες υψηλή, μεσαία και χαμηλή. ....	70

Πίνακας 3.9 : Οι εξειδικευμένοι βιοχημικοί δείκτες που παρουσιάζουν τη μεγαλύτερη συσχέτιση με το GSM διαστολής.....	72
Πίνακας 3.10 : Οι εξειδικευμένοι βιοχημικοί δείκτες που παρουσιάζουν τη μεγαλύτερη συσχέτιση με το GSM συστολής.....	73
Πίνακας 3.11 : Οι εξειδικευμένοι βιοχημικοί δείκτες που παρουσιάζουν τη μεγαλύτερη συσχέτιση με το εμβαδόν της καρωτίδας.....	73
Πίνακας 3.12 : Πίνακας των ποσοστών των ασθενών κάθε φαινοτύπου στις κατηγορίες υψηλή, μεσαία και χαμηλή.....	79
Πίνακας 3.13 : Πίνακας των ποσοστών των ασθενών κάθε φαινοτύπου στις κατηγορίες υψηλή, μεσαία και χαμηλή.....	79
Πίνακας 3.14 : Πίνακας των ποσοστών των ασθενών κάθε φαινοτύπου στις κατηγορίες υψηλή, μεσαία, χαμηλή και του είδους συμπτωμάτων.....	80
Πίνακας 3.15 : Οι συμβατικοί και εξειδικευμένοι βιοχημικοί δείκτες που παρουσιάζουν τη μεγαλύτερη συσχέτιση με το GSM διαστολής.....	83
Πίνακας 3.16 : Οι συμβατικοί και εξειδικευμένοι βιοχημικοί δείκτες που παρουσιάζουν τη μεγαλύτερη θετική και αρνητική συσχέτιση με το GSM συστολής.....	84
Πίνακας 3.17 : Οι συμβατικοί και εξειδικευμένοι βιοχημικοί δείκτες που παρουσιάζουν τη μεγαλύτερη θετική και αρνητική συσχέτιση με το εμβαδόν της καρωτίδας.....	85



# Κεφάλαιο 1: Εισαγωγή

## Περίληψη

Στο παρόν κεφάλαιο παρουσιάζονται αρχικά τα στοιχεία φυσιολογίας της καρωτίδας καθώς και η νόσος της αθηρωματικής πλάκας, μια από τις πιο σοβαρές καρδιαγγειακές παθήσεις. Στη συνέχεια, γίνεται εκτενής αναφορά στους συμβατικούς βιοχημικούς δείκτες και στη συσχέτισή τους με τη νόσο. Γίνεται επισήμανση του ρόλου του υπερήχου για την καταγραφή και ανάλυση τους αρτηριακού τοιχώματος και αναλύονται εκτενώς οι μέθοδοι επεξεργασίας εικόνων υπερήχων για την ανίχνευση της αθηρωματικής πλάκας. Επίσης γίνεται αναφορά τους απεικονιστικούς δείκτες της νόσου. Τέλος, παρατίθεται ο σκοπός της παρούσας διπλωματικής εργασίας.

## 1.1 Στοιχεία φυσιολογίας της καρωτίδας

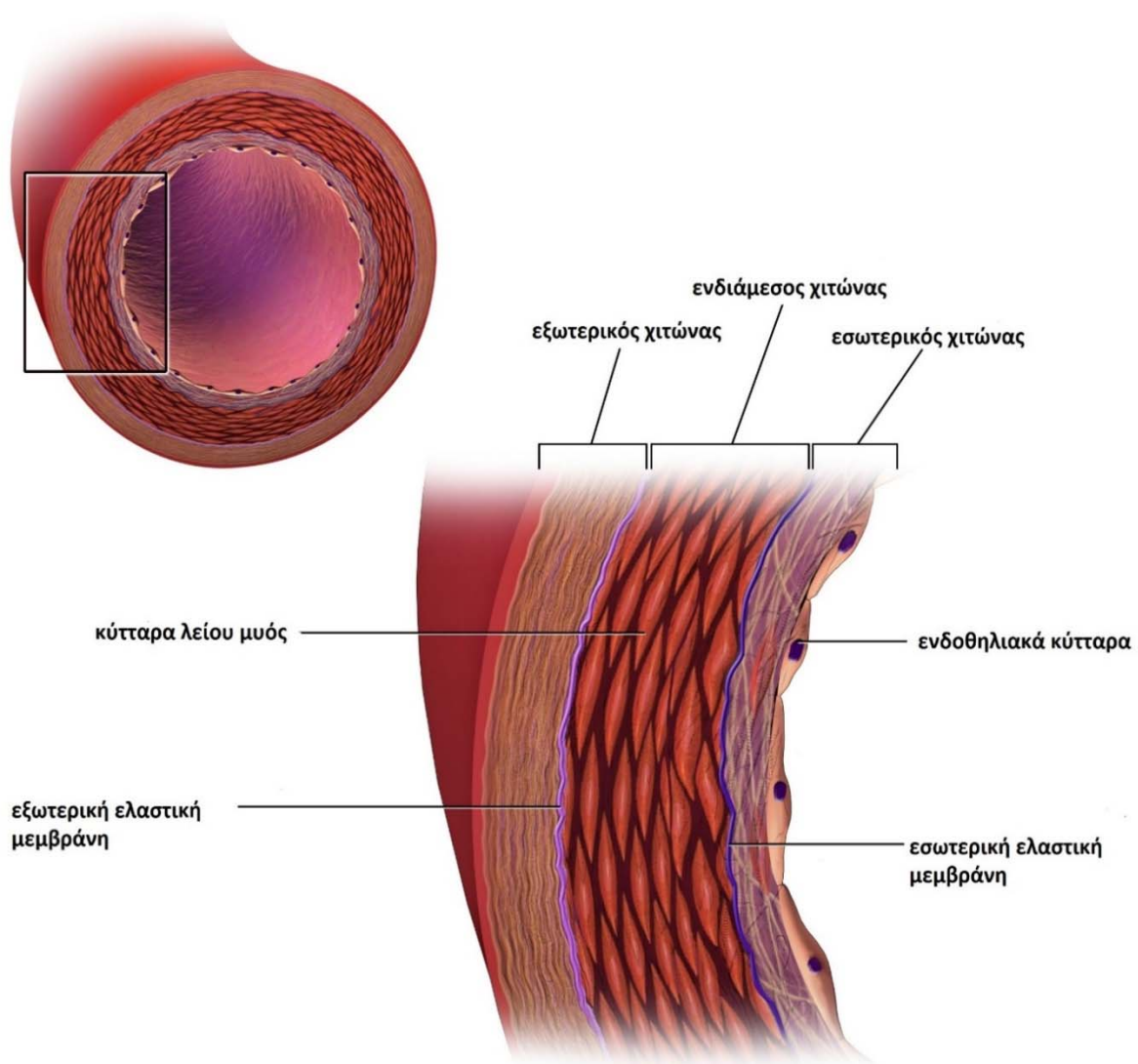
Οι καρωτίδες είναι δύο αρτηρίες, οι οποίες τροφοδοτούν με αίμα το λαιμό και την κεφαλή. Διαχωρίζονται σε δεξιά και αριστερή καρωτίδα και κάθε μία από αυτές χωρίζεται με τη σειρά τους στην εσωτερική (έσω) και εξωτερική (έξω) καρωτίδα όπως φαίνεται στην Εικόνα 1.1. Οι εξωτερικές καρωτίδες παρουσιάζουν πολλαπλές διακλαδώσεις, τροφοδοτώντας περιοχές εξωτερικά του κρανίου όπως το πρόσωπο και το τριχωτό της κεφαλής. Από την άλλη πλευρά, οι εσωτερικές αρτηρίες τροφοδοτούν με αίμα τον εγκέφαλο, τους οφθαλμούς, τη μύτη και τέλος το μέτωπο [1].



Εικόνα 1.1 : Σχηματική αναπαράσταση της φυσιολογίας της καρωτίδας [1]

Το καρωτιδικό τοίχωμα αποτελείται από τρεις χιτώνες, τον εξωτερικό, τον ενδιάμεσο και τον εσωτερικό χιτώνα (Εικόνα 1.2). Ο εξωτερικός χιτώνας (tunica adventitia) αποτελείται κυρίως από συνδετικό ιστό και κολλαγόνο. Ο ρόλος του είναι η προστασία του αγγείου και η ευστάθεια που παρέχει στις πιθανές διευθύνσεις κίνησης λόγω της ακανόνιστης μορφής των ινών κολλαγόνου. Ο ενδιάμεσος χιτώνας (tunica media), αποτελείται από ελαστικές ίνες και από λεία

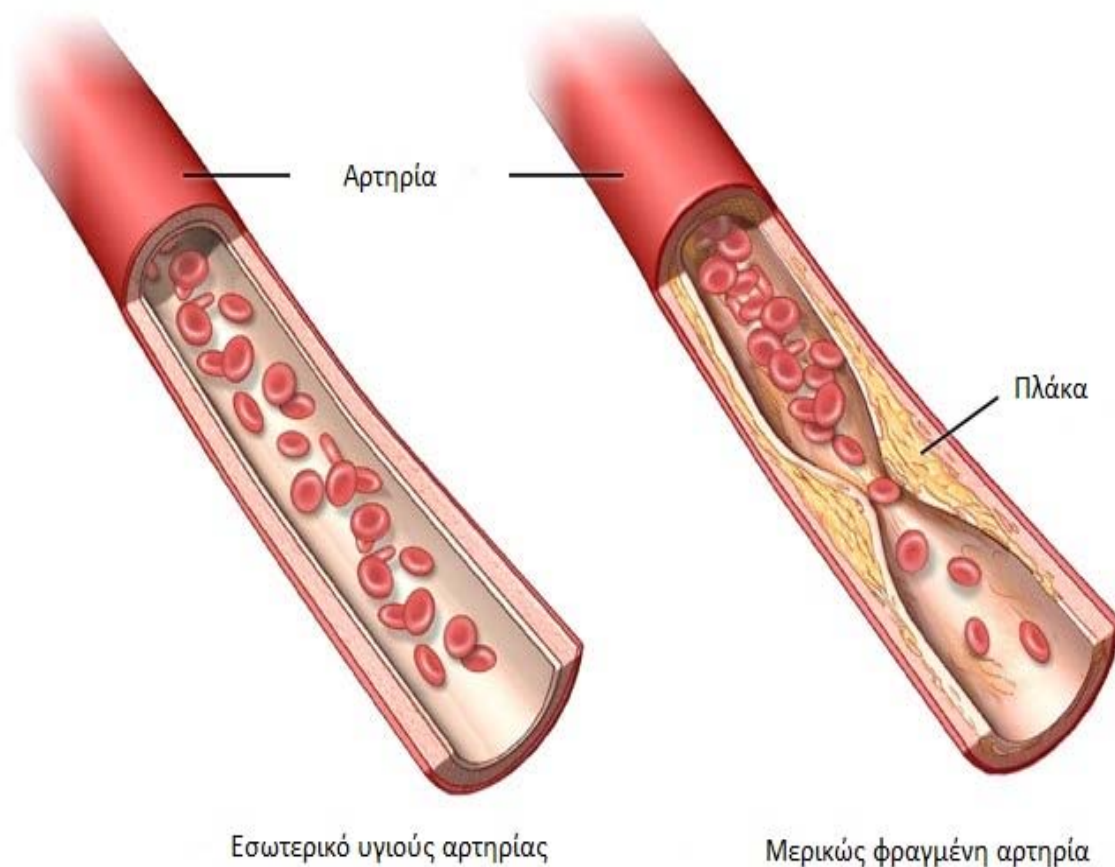
μυϊκά κύτταρα. Καθορίζει κατά κύριο λόγο το πάχος της αρτηρίας και τον τρόπο κίνησής της. Τέλος, ο εσωτερικός χιτώνας (tunica intima) απαρτίζεται από μια στοιβάδα επιθηλιακών κυττάρων και από μία ελαστική μεμβράνη (internal elastic lamina). Το πάχος του εσωτερικού χιτώνα είναι σημαντικά μικρότερο σε σχέση με τους δυο παραπάνω χιτώνες. Επιπλέον, η διάταξη των επιθηλιακών κυττάρων διευκολύνει την ροή του αίματος [2].



Εικόνα 1.2 : Αναπαράσταση δομής και σύστασης του αρτηριακού τοιχώματος [3].

## 1.2 Η νόσος της καρωτιδικής αθηρωμάτωσης

Η αθηρωμάτωση ή αθηροσκλήρωση αποτελεί μια από τις κύριες αιτίες καρδιακών παθήσεων και εμφραγμάτων παγκοσμίως. Είναι μια εκφυλιστική πάθηση που προσβάλλει τις αρτηρίες και χαρακτηρίζεται από τη συσσώρευση λιπιδίων, κυτταρικών αποβλήτων, ασβεστίου και ινωδών στοιχείων στα εσωτερικά τοιχώματα των αρτηριών. Το γεγονός αυτό έχει σαν αποτέλεσμα τη σταδιακή αλλοίωση του ενδοθηλίου δηλαδή του εσωτερικού τοιχώματος της αρτηρίας και το σχηματισμό της λεγόμενης αθηρωματικής πλάκας (Εικόνα 1.3), η οποία αυξάνεται συνεχώς με απώτερο αποτέλεσμα τη στένωση της αρτηρίας και τη μείωση της ελαστικότητας του αρτηριακού τοιχώματος [2] [4].



Εικόνα 1.3 : Απεικόνιση υγιούς αρτηρίας και αρτηρίας με αθηρωματική πλάκα [5].

Αναμφίβολα, τα αίτια και οι παράγοντες που σχετίζονται με τη νόσο συνδέονται με τον τρόπο ζωής και με διάφορες άλλες παθήσεις. Από επιδημιολογικές μελέτες που έχουν γίνει τα τελευταία χρόνια, έχουν χωριστεί οι υπαίτιοι παράγοντες της νόσου σε δύο κατηγορίες, πρώτον σε παράγοντες με γενετικό υπόβαθρο και σε περιβαλλοντικούς παράγοντες [6].

Όσον αφορά το γενετικό υπόβαθρο, οι κυριότεροι παράγοντες είναι το αυξημένο επίπεδο λιποπρωτεϊνών, η αυξημένη αιματική πίεση (υπέρταση), η κληρονομικότητα, το φύλο και η ηλικία, τα αυξημένα επίπεδα ομοκυστεΐνης (μη πρωτεϊνικό αμινοξύ), η πάθηση του διαβήτη και της παχυσαρκίας και το μεταβολισμό σύνδρομο.

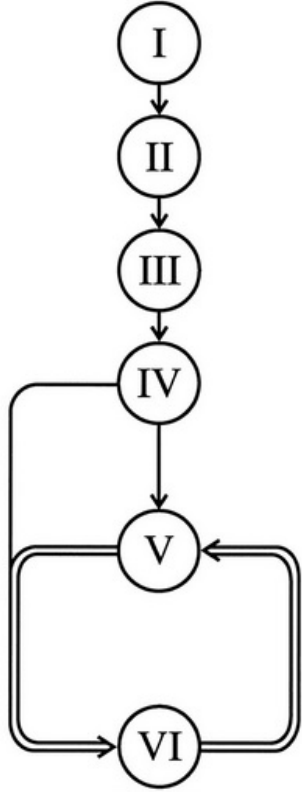
Από την άλλη πλευρά, οι παράγοντες με περιβαλλοντικό χαρακτήρα είναι η διατροφή με υψηλή περιεκτικότητα λιπαρών, το κάπνισμα, η έλλειψη σωματικής άσκησης και τα χαμηλά επίπεδα αντιοξειδωτικών.

## Εξέλιξη της νόσου

Βάσει της σοβαρότητας, οι αθηρωματικές πλάκες χωρίζονται σε έξι τύπους από I έως VI. Οι πλάκες τύπου IV, V, VI αποτελούν προχωρημένες αλλοιώσεις, οι τύπου I και II αποτελούν τις πρώιμες αλλοιώσεις ενώ ο τύπος III αποτελεί το μεταβατικό στάδιο από τις πρώιμες αλλοιώσεις στις προχωρημένες (Πίνακας 1.1).

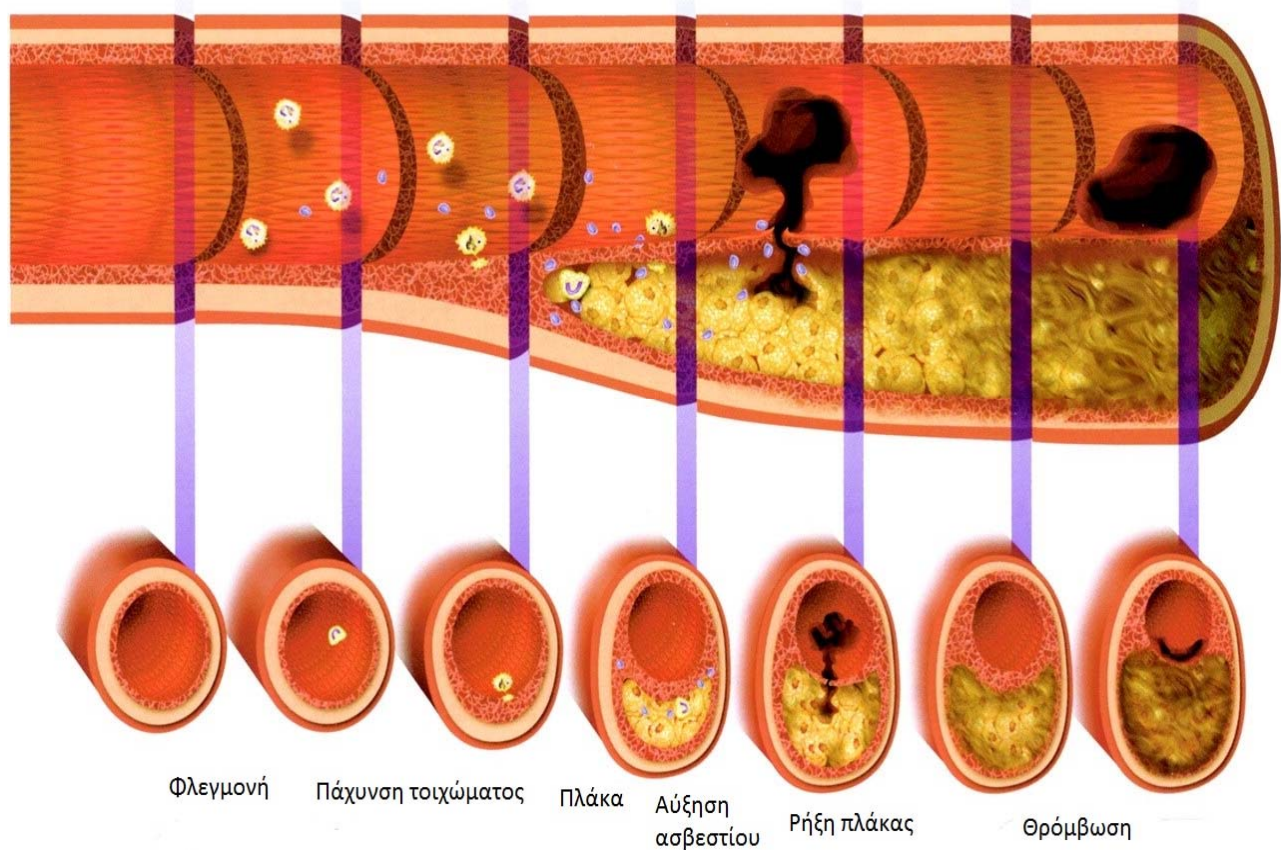
Οι αθηρωματικές αλλοιώσεις τύπου I χαρακτηρίζονται από την πρώτη συσσώρευση-εναπόθεση αφρωδών κυττάρων (μακροφάγων), στο εσωτερικό τοίχωμα των αρτηριών. Οι αλλοιώσεις τύπου II χαρακτηρίζονται από τις λιπώδεις γραμμώσεις, οι οποίες δημιουργούνται από την εναπόθεση αφρωδών κυττάρων σε μορφή στοιβάδων (συμπαγές δομές). Οι αλλοιώσεις τύπου III προκύπτουν από την εναπόθεση και συσσώρευση επιπλέον λιπιδίων στο εσωτερικό του λείου μυ. Η Εικόνα 1.4 προσομοιώνει την εξέλιξη των παραπάνω τύπων. Όσον αφορά την πλάκα τύπου IV, βασικό χαρακτηριστικό γνώρισμα αποτελεί η ύπαρξη λιπιδικού πυρήνα, ο οποίος αυξάνεται ταχεία. Τέλος, αθηρωματικές αλλοιώσεις τύπου V χαρακτηρίζονται από την παρουσία ινώδους ιστού. Αν παρουσιαστεί ρήξη ή θρόμβος στις πλάκες τύπου IV και V τότε χαρακτηρίζονται ως τύπου VI δηλαδή περίπλοκη αθηρωματική αλλοίωση όπως φαίνεται και από τη σχηματική παράσταση που υπάρχει στον (Πίνακας 1.1) [7].

Πίνακας 1.1: Κατηγοριοποίηση των τύπων της αθηρωματικής πλάκας και η πορεία εξέλιξης της νόσου.

Ορολογία	Πορεία Εξέλιξης
<b>Τύπου I</b> Αρχική αλλοίωση (Initial lesion)	
<b>Τύπου II</b> Λιπώδης Γράμμωση (Fatty streak lesion)	
<b>Τύπου III</b> Προαθήρωμα (Intermediate lesion)	
<b>Τύπου IV</b> Αθήρωμα (Atheroma lesion)	
<b>Τύπου V</b> Ινωδοαθήρωμα (Fibroatheroma lesion)	
<b>Τύπου VI</b> Περίπλοκη αθηρωματική αλλοίωση (Complicated lesion)	

Ο βρόχος που συνδέει τον τύπο V με τον τύπο VI στο διάγραμμα της πορείας της εξέλιξης στον Πίνακα 1.1, αναπαριστά το πώς η πλάκα αυξάνεται συνεχώς σε πάχος λόγω της συσσώρευσης θρομβωτικών εναποθέσεων, οι οποίες εναποθέσεις μπορούν να σχηματίζονται στο ίδιο σημείο γεγονός που αποτελεί το κύριο μηχανισμό για τη προοδευτική στένωση των αρτηριών [7]. Αυτή η στένωση εμποδίζει την ομαλή ροή του αίματος με απώτερο αποτέλεσμα τη μη σωστή παροχή οξυγόνου στον εγκέφαλο και σε άλλα μέρη του σώματος, γεγονός που μπορεί να οδηγήσει μέχρι και στον θάνατο (ισχαιμικό επεισόδιο) (Εικόνα 1.5) [8].



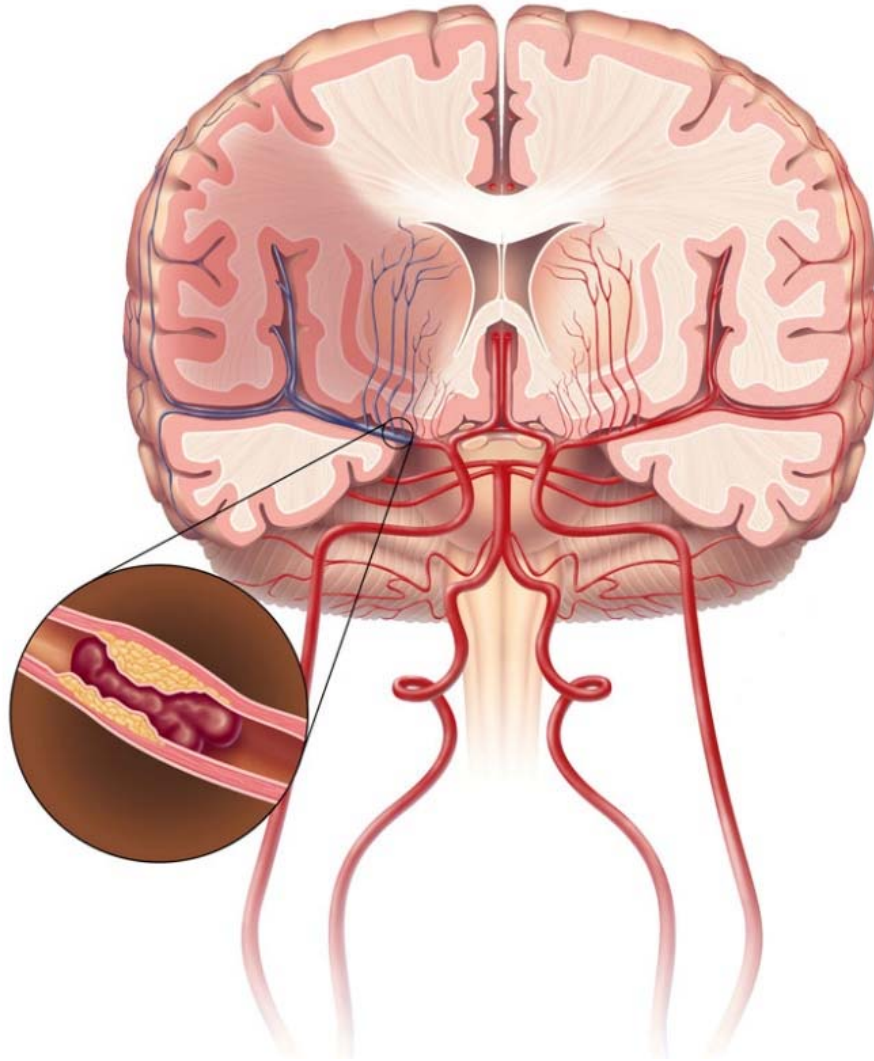


Εικόνα 1.4 : Η εξέλιξη της αθηρωματικής πλάκας [9].

## Συμπτωματολογία της νόσου

Εν γένει τα συμπτώματα της καρωτιδικής αθηρωμάτωσης εμφανίζονται μόνο αφού η στένωση της καρωτιδικής αρτηρίας έχει προχωρήσει σε μεγάλο βαθμό και εμποδίζει έτσι την αιματική ροή προς τον εγκέφαλο. Τις περισσότερες φορές, οι ασθενείς που πάσχουν από τη νόσο δεν το γνωρίζουν μέχρι να εμφανιστεί κάποια καρδιαγγειακή διαταραχή όπως οξύ ισχαιμικό επεισόδιο (Ischemic stroke) ή προσωρινό ισχαιμικό επεισόδιο (TIA-transient ischemic attack) του οποίου τα συμπτώματα είναι προσωρινά και μπορεί να διαρκέσουν από μερικά λεπτά μέχρι μερικές ώρες και δεν προκαλεί συνήθως μόνιμες βλάβες στον εγκέφαλο. Ωστόσο

θα πρέπει να δοθεί μεγάλη σημασία στα συμπτώματά του προσωρινού ισχαιμικού περιστατικού διότι θεωρείται σαν προειδοποίηση η οποία πρέπει να διερευνηθεί. Τέλος, μερικά από τα συμπτώματα αυτά είναι το μούδιασμα ή το αίσθημα αδυναμίας στο πρόσωπο και στα άκρα, δυσκολία στην ομιλία, ξαφνική αδυναμία όρασης από το ένα ή και τα δύο μάτια, ζαλάδα, παραμόρφωση των χειλών και της μίας πλευράς του προσώπου[10].



Εικόνα 1.5 : Ισχαιμικό επεισόδιο λόγω θρόμβωσης [11].



### 1.3 Συσχέτιση βιοχημικών δεικτών με την παρουσία και εξέλιξη της νόσου

Οι βιοχημικοί δείκτες παρέχουν μια αξιόπιστη εικόνα για τη γενικότερη κατάσταση του οργανισμού. Υπάρχουν πολλών ειδών εξετάσεις βιοχημικών δεικτών όπως, η γενική αίματος, η εξέταση για την ταχύτητα καθίζησης ερυθρών, για μέτρηση της γαλακτικής αφυδρογονάσης (LDH), για χοληστερίνη, για κινάση κρεατίνης (CPK), για χαμηλά επίπεδα υψηλής πυκνότητας χοληστερόλης (HDL- γνωστή και ως καλή χοληστερόλη), για υψηλά επίπεδα χαμηλής πυκνότητας χοληστερόλης (LDL) και για τριγλυκερίδια [12]. Από την γενική εξέταση αίματος προκύπτουν τιμές-δείκτες όπως η αιμοσφαιρίνη, ο αιματοκρίτης, αιμοπετάλια κτλ, από τις οποίες τιμές, μπορεί να αξιολογηθεί η γενική κατάσταση του οργανισμού και γενικότερα να φανεί αν υπάρχει κάποια λοίμωξη. Ομοίως και με την εξέταση της ταχύτητας καθίζησης των ερυθρών. Από την εξέταση για τον προσδιορισμό των επιπέδων της γαλακτικής αφυδρογονάσης (LDH), που αποτελεί ένζυμο του οργανισμού, είναι δυνατός ο εντοπισμός παθολογικών καταστάσεων που συνδέονται με αυξημένη τιμή του ενζύμου, ενώ η εξέταση για την κινάση της κρεατίνης (CPK) βοηθά στη διάγνωση μυϊκών παθήσεων. Οι δείκτες χοληστερίνης, HDL, LDL και τριγλυκεριδίων συνθέτουν το λιπιδαιμικό προφίλ του οργανισμού. Όταν οι μετρήσεις αυτές ξεπεράσουν κάποια όρια τότε αυξάνεται η πιθανότητα σχηματισμού αθηρωματικής πλάκας και εκδήλωσης καρδιακών δυσλειτουργιών [14].

Ο ρόλος των βιοχημικών δεικτών με τη νόσο της αθηρωμάτωσης έχει διερευνηθεί εκτενώς σε συναφείς μελέτες. Αρχικά, οι M. Griffin et al. [15] μελέτησαν τη σύνδεση της υπερβολικής ινσουλίνης με τη σύνθεση της χοληστερόλης. Μελετήθηκαν τρεις κατηγορίες ασθενών και η απόκρισή τους στην ινσουλίνη. Υπήρξε θετική συσχέτιση μεταξύ της σύνθεσης της χοληστερόλης και της αντίστασης της ινσουλίνης. Έπειτα από απώλεια βάρους των ασθενών, χωρίστηκαν σε δύο ομάδες, σε αυτούς που μείωναν την αντίσταση ινσουλίνης και σε αυτούς που δεν μείωναν την αντίσταση της ινσουλίνης. Η σημαντική μετα-γευματική αυξημένη σύνθεση της χοληστερόλης δεν παρατηρήθηκε στην πρώτη ομάδα ασθενών ενώ στη δεύτερη ομάδα ασθενών παρέμεινε ανεβασμένη. Αυτά τα αποτελέσματα δείχνουν ότι η υπερβολική ποσότητα ινσουλίνης συσχετίζεται με αυξημένη σύνθεση χοληστερόλης. Αυτή η αύξηση παύει να

παρατηρείται σε ασθενείς στους οποίους η αντίσταση ινσουλίνης βελτιώθηκε με την απώλεια βάρους.

Άλλος συμβατικός δείκτης που έχει συσχετιστεί με τη νόσο της αθηρωμάτωσης είναι ο δείκτης της ομοκυστεΐνης. Σύμφωνα με μελέτες, η υπερομοκυστεϊναιμία αναγνωρίστηκε σαν ανεξάρτητος παράγοντας κινδύνου καρδιαγγειακής νόσου. Η ομοκυστεΐνη είναι ένα αμινοξύ που προέρχεται από την μεθειονίνη της τροφής. Η αύξηση της ομοκυστεΐνης στο πλάσμα πάνω από τα φυσιολογικά όρια των 5-15  $\mu\text{mol/L}$  συνεπάγεται υπερομοκυστεϊναιμία. Το παθογενετικό υπόστρωμα της νόσου είναι η δυσλειτουργία του αγγειακού ενδοθηλίου και ο θρομβωγόνος ρόλος του, αποτέλεσμα της τοξικής δράσης της αυξημένης ομοκυστεΐνης και κυρίως των προϊόντων οξείδωσής της. Ο προσδιορισμός της ουσίας στο πλάσμα μετά από φόρτιση με μεθειονίνη συνίσταται στη μέτρηση της ομοκυστεΐνης πριν και μετά τη χορήγηση μεθειονίνης ανά 100 mg/kg σωματικού βάρους σε νησικά άτομα αν τακτά χρονικά διαστήματα. Στα φυσιολογικά άτομα, υπό δοκιμασία φόρτισής με μεθειονίνη, οι τιμές της ομοκυστεΐνης θα πρέπει να έχουν επανέλθει σε φυσιολογικά επίπεδα μετά από δυο ώρες. Κατά το μεταβολισμό της ουσίας, προϊόντα της οξείδωσής της συνδέονται με τις LDL- λιποπρωτεΐνες με αποτέλεσμα να ενισχύεται το παθογενετικό υπόστρωμα της αθηρωματικής πλάκας με το σχηματισμό αφρωδών λιποκυττάρων. Η μελέτη πραγματοποιήθηκε 750 περιπτώσεις αθηροσκλήρωσης με ασθενείς και των δυο φύλων και με ηλικία κάτω των 60 ετών. [16].

Άλλη μια προσπάθεια συσχέτισης βιοχημικών δεικτών με τη νόσο αποτέλεσε η μελέτη 3258 ενήλικων ασθενών με ηλικία από 18 έως 30 ετών από τέσσερις διαφορετικές πόλεις των ΗΠΑ [17]. Στους ασθενείς μετρήθηκαν οι τιμές LDL και HDL. Η συνολική έκθεση σε λιπίδια υπολογίστηκε με βάση συχνές διαδοχικές τιμές λιπιδίων ορού κατά τη διάρκεια της μελέτης. Συνολικά το 87% του συνολικού αριθμού παρουσίασε δυσλιπιδαιμία στη νεαρή ενήλικη ζωή. Σε μετρήσεις που έλαβαν χώρα είκοσι χρόνια αργότερα ο επιπολασμός της ασβέστωσης στεφανιαίων ήταν 8 % σε αυτούς με φυσιολογικά επίπεδα λιπιδίων και 44% σε αυτούς με αυξημένες τιμές LDL χοληστερόλης. Η συσχέτιση γινόταν ισχυρότερη όσο τα επίπεδα LDL αυξάνονταν. Τόσο η LDL όσο και η HDL συσχετίστηκαν ανεξάρτητα με την ασβέστωση των στεφανιαίων. Συμπερασματικά, αυξημένα επίπεδα LDL και HDL χοληστερόλης στη νεαρή

ενήλικο ζωή συσχετίστηκαν ανεξάρτητα με αθηρωμάτωση των στεφανιαίων δύο δεκαετίες αργότερα.

Σε άλλη μελέτη παρακολούθηθηκαν 1675 ασθενείς, οι οποίοι είχαν υποστεί καρδιαγγειακή επέμβαση και μετρήθηκαν οι τιμές της κρεατινίνης τους, η αύξηση της οποίας έπειτα από την επέμβαση συσχετίστηκε με μετέπειτα εμφράγματα του μυοκαρδίου [18]. Από τους 1675 ασθενείς, σε 313 ανιχνεύτηκαν υψηλά επίπεδα κρεατινίνης. Αυτή η μεταβολή παρατηρήθηκε κυρίως σε ασθενείς με αθηροσκλήρωση, γεγονός που συνδέει τον συγκεκριμένο βιοχημικό δείκτη με τη νόσο. Αύξηση των τιμών της κινάσης κρεατινίνης αποτελεί ένδειξη πιθανής ύπαρξης αθηρωμάτωσης. Επιπλέον η συσχέτιση της νόσου και ενός άλλου βιοχημικού δείκτη, της ουρίας, παρατηρήθηκε σε μελέτη που έγινε σε ασθενείς που έπασχαν από χρόνια νεφρική ανεπάρκεια [19]. Έχει παρατηρηθεί ότι ασθενείς με χρόνια νεφρική ανεπάρκεια, εμφανίζουν καρδιαγγειακές νόσους τριάντα φορές συχνότερα από τον γενικό πληθυσμό. Η συνέχιση της προόδου της νεφρικής βλάβης έχει ως αποτέλεσμα την μείωση των επιπέδων φετουΐνης, μιας προστατευτικής πρωτεΐνης για ανάπτυξη ασβεστώσεων και οξείδωση των LDL μορίων μέσω χημικής τροποποίησης του πρωτεϊνικού συστατικού της LDL με οξύ που προέρχεται από την ουρία. Ως αποτέλεσμα υπάρχει ρήξη μεταξύ των παραγόντων που προάγουν την αγγειακή βλάβη με αυτούς που την αναστέλλουν. Τελικά, αναπτύσσονται σοβαρές αγγειακές ασβεστώσεις και αρτηριοσκλήρυνση όταν υπάρχει ο συνδυασμός νεφρικής ανεπάρκειας και υψηλών επιπέδων ουρίας.

Επιπρόσθετα, ένας άλλος συμβατικός βιοχημικός δείκτης που σχετίζεται με την αθηρωμάτωση είναι ο πληθυσμός των λεμφοκυττάρων. Σε πειράματα που πραγματοποιήθηκαν, αποδείχθηκε ότι η μεταβολή του πληθυσμού των λεμφοκυττάρων μπορεί να χρησιμοποιηθεί ως εργαλείο διάγνωσης και πρόγνωσης της αθηροσκλήρωσης [20]. Η έρευνα αφορούσε δύο ομάδες, μια ομάδα ασθενών σε προχωρημένο στάδιο της νόσου HIV και μια ομάδα μη μολυσμένων από τον ιό ασθενών αλλά με θετική διάγνωση πάχυνσης του έσω χιτώνα της καρωτίδας όπως επιβεβαιώθηκε από υπερηχογράφημα Doppler. Όταν συγκρίθηκαν τα λαμβανόμενα από τις δύο ομάδες δεδομένα παρατηρήθηκε ότι οι πληθυσμοί λεμφοκυττάρων μεταβλήθηκαν στην ίδια κατεύθυνση (δηλαδή αυξημένη ή μειωμένη συχνότητα στους ασθενείς που ανέπτυξαν πλάκα έναντι εκείνων που δεν εμφάνισαν πλάκα) και στην ομάδα των

προσβεβλημένων από τον ιό HIV ασθενών και σε μη μολυσμένους ασθενείς. Οι ασθενείς με HIV ήταν σε προχωρημένο στάδιο και έλαβαν θεραπεία με αναστολείς πρωτεάσης, η οποία συνδέεται στατιστικά με σημαντική αύξηση του κινδύνου καρδιαγγειακής νόσου. Συμπερασματικά, και στις δύο ομάδες οι ασθενείς που ανέπτυξαν πλάκα εμφάνισαν κοινές μεταβολές στους πληθυσμούς λεμφοκυττάρων, συνδέοντας με αυτόν τον τρόπο αυτόν τον βιοχημικό δείκτη με την αθηρωμάτωση. Αντίστοιχα, έχει μελετηθεί και η συσχέτιση της αθηρωματικής πλάκας με εξειδικευμένους βιοχημικούς δείκτες. Αρχικά, παρακολούθηθηκαν 258 ασθενείς οι οποίοι παρουσίαζαν στένωση μεγαλύτερη του 50% [21]. Μετρήθηκε η CRP και ένα είδος γλυκοπρωτεΐνης (ogrosomuroid) καθώς επίσης έγινε χρήση υπερήχου για την εκτίμηση των χαρακτηριστικών της πλάκας. Τα αποτελέσματα της μελέτης συσχέτισαν την αυξημένη παρουσία των δυο παραπάνω εξειδικευμένων βιοχημικών δεικτών με τα αυξημένα επίπεδα τριγλυκεριδίων και με την αύξηση της πλάκας. Επιπλέον, η αυξημένη παρουσία των δυο παραπάνω εξειδικευμένων βιοχημικών δεικτών προδιάθετε για πλάκα επιρρεπή σε ρήξη.

Ακόμη ένας εξειδικευμένος βιοχημικός δείκτης που συσχετίστηκε με τη νόσο, είναι η γαλεκτίνη-3. Η συσχέτιση του δείκτη αυτού με την αθηρωμάτωση παρατηρήθηκε στο πλαίσιο έρευνας που συμπεριλάμβανε 78 ασθενείς (40 συμπτωματικούς και 38 ασυμπτωματικούς) [22]. Στους ασθενείς αυτούς πραγματοποιήθηκε υπολογισμός του δείκτη GSM και ιστολογική ανάλυση της καρωτίδας μετά από ενδαρτηρεκτομή μέσω υπερήχου. Οι συμπτωματικοί ασθενείς είχαν μικρότερη τιμή GSM και μικρότερη τιμή γαλεκτίνης-3 σε αντίθεση με τους ασυμπτωματικούς. Χαμηλά επίπεδα γαλεκτίνης-3 συσχετίστηκαν με ασταθή καρωτιδική πλάκα και προέκυψε θετική συσχέτιση μεταξύ γαλεκτίνης-3 και GSM.

Η συσχέτιση μεταξύ του βιοχημικού δείκτη retinol-binding protein-4 (RBP4) και της νόσου πραγματοποιήθηκε μέσω της μελέτης 225 ασθενών που χαρακτηρίζονταν από υψηλό βαθμό στένωσης και 75 ασθενείς με χαμηλό βαθμό στένωσης [23]. Ακόμη ορίστηκε ένα control group (ομάδα αναφοράς) που περιλάμβανε 73 άτομα χωρίς συμπτώματα. Μέσω καρωτιδικού υπερήχου υπολογίστηκε το εμβαδό των πλακών και οι αντίστοιχοι δείκτες GSM. Παρατηρήθηκε πως η RBP4 ήταν αυξημένη στους ασθενείς με υψηλό βαθμό στένωσης σε σχέση με αυτούς με χαμηλό βαθμό στένωσης και χαμηλότερη από τους ασθενείς αναφοράς. Τα υψηλά αυτά επίπεδα συσχετίστηκαν με μεγαλύτερης σοβαρότητας αθηροσκλήρωση.

Στο πλαίσιο έρευνας 246 ασθενών που έπασχαν από τη νόσο της καρωτιδικής αθηρωματικής πλάκας πραγματοποιήθηκε συσχέτιση της νόσου με άλλο ένα βιοχημικό δείκτη. Συγκεκριμένα μετρήθηκαν τα επίπεδα της ιντερλευκίνη-6 (IL-6, interleukin-6) καθώς και οι δείκτες GSM των αντίστοιχων αθηρωματικών πλακών. Αποτέλεσμα της έρευνας αποτέλεσε η συσχέτιση των υψηλών επιπέδων της ιντερλευκίνη-6 με χαμηλές τιμές GSM [24].

Ακόμη ένας εξειδικευμένος δείκτης που σχετίζεται με την νόσο της αθηρωματικής πλάκας αποτελεί ο δείκτης TIMP1. Σε έρευνα που διεξήχθη σε 1069 ασθενείς μέσης ηλικίας 59 ετών, εκ των οποίων το 58% αποτελούνταν από γυναίκες, πραγματοποιήθηκε συσχέτιση του δείκτη αυτού με την νόσο της αθηρωμάτωσης [25]. Οι τιμές του TIMP-1 βρέθηκαν υψηλότερες σε ασθενείς με μεγαλύτερη δείκτη μάζας σώματος και επίπεδα, με αυξημένα επίπεδα HDL χοληστερόλης καθώς και με τον διαβήτη. Λόγω των περισσότερων λιπιδίων, παρατηρείται αύξηση του πάχους της πλάκας και μεγαλύτερος κίνδυνος καρδιαγγειακών παθήσεων. Ο δείκτης TIMP1 έχει θετική συσχέτιση με το πάχος και την επιφάνεια της πλάκας και αρνητική με το GSM, δεδομένου ότι μεγαλύτερο πάχος συνεπάγεται με μικρότερη ένταση ηχούς. Τέλος, ακόμη μια έρευνα [26] που έγινε σε 2191 ασθενείς, από τους οποίους οι 351 παρουσίασαν, σε βάθος 8 χρόνων, καρδιαγγειακά επεισόδια μετρήθηκαν και στη συνέχεια συσχετίστηκαν οι δείκτες CRP, IL-6 και ο παράγοντας νέκρωσης όγκου-α (TNF-α) με τη νόσο της αθηρωμάτωσης.

Πίνακας 1.2 : Συνοπτική παρουσίαση εργασιών που συσχέτισαν βιοχημικούς δείκτες με τη νόσο της αθηρωμάτωσης.

Έτος	Δεδομένα	Δείκτες	Συμπεράσματα
<b>δημοσίευσης</b>			
<b>1998 [15]</b>	Διαβητικοί/Μη διαβητικοί ασθενείς	Αντίσταση Ινσουλίνης	Βελτίωση της αντίστασης της ινσουλίνης οδηγεί σε μειωμένη σύνθεση χοληστερόλης και μείωση του κινδύνου εμφάνισης της νόσου.

Έτος δημοσίευσης	Δεδομένα	Δείκτες	Συμπεράσματα
2001 [16]	750 Ασθενείς με αθηρωμάτωση	Ομοκυστεΐνη	Υψηλά επίπεδα ομοκυστεΐνης στο πλάσμα αυξάνουν τον κίνδυνο εμφάνισης της αθηρωματικής νόσου.
2010 [17]	3258 ενήλικες	LDL, HDL	Αυξημένα επίπεδα LDL και HDL χοληστερόλης στη νεαρή ενήλικη ζωή συσχετίστηκαν ανεξάρτητα, με τη νόσο της αθηρωμάτωσης δύο δεκαετίες αργότερα.
1999 [18]	1675 Ασθενείς	Κινάση κρεατινίνης	Παρατηρήθηκε ότι αυξημένες τιμές κινάσης κρεατινίνης εμφανίζονταν έπειτα από καρδιαγγειακή επέμβαση σε ασθενείς με αθηρωμάτωση.
2011 [19]	Ασθενείς με χρόνια νεφρική ανεπάρκεια	Ουρία	Οι βλάβες που προκαλούνται από νεφρική ανεπάρκεια σε συνδυασμό με την ουρία οδηγούν σε ανάπτυξη σοβαρών αγγειακών ασβεστώσεων και αρτηριοσκλήρυνσης
2006 [20]	Φορείς και μη του ιού HIV	Λεμφοκύτταρα	Η μεταβολή του πληθυσμού των λεμφοκυττάρων μπορεί να χρησιμοποιηθεί ως προγνωστικός δείκτης της νόσου.

Έτος δημοσίευσης	Δεδομένα	Δείκτες	Συμπεράσματα
<b>2001 [21]</b>	258 Ασθενείς με στένωση >50%	CRP, orosomucoid (γλυκοπρωτεΐνη)	Η αυξημένη τιμή των δεικτών αυτών σχετίζεται με την αύξηση των τριγλυκεριδίων και της πλάκας.
<b>2015 [22]</b>	78 Ασθενείς (40 συμπτωματικοί και 38 ασυμπτωματικοί)	Γαλεκτίνη	Συσχέτιση χαμηλής συγκέντρωσης της γαλεκτίνης με ασταθείς πλάκες. Θετική συσχέτιση GSM και γαλεκτίνης-3.
<b>2014 [23]</b>	225 Ασθενείς με υψηλό βαθμό στένωσης 79 Ασθενείς με χαμηλό βαθμό στένωσης 73 Άτομα ως control group	RBP4 (retinol-binding protein-4)	Η RBP4 ήταν αυξημένη στους ασθενείς με υψηλό βαθμό στένωσης περισσότερο από ότι στους ασθενείς με χαμηλό βαθμό στένωσης.
<b>2004 [24]</b>	246 Ασθενείς που έπασχαν από καρωτιδική αθηρωμάτωση.	IL-6 (interleukin-6)	Συσχέτιση υψηλών επιπέδων IL-6 με χαμηλές τιμές GSM και κίνδυνος ύπαρξης ασταθούς πλάκας.
<b>2015 [25]</b>	1069 Ασθενείς	TIMP-1	Θετική συσχέτιση του πάχους της πλάκας την τιμή του TIMP-1 και αρνητική συσχέτιση με το GSM.
<b>2010 [26]</b>	2191 Ασθενείς	CRP, TNF-α, IL-6	Αύξηση της τιμής αυτών των δεικτών συνεπάγεται αύξηση της επικινδυνότητας.

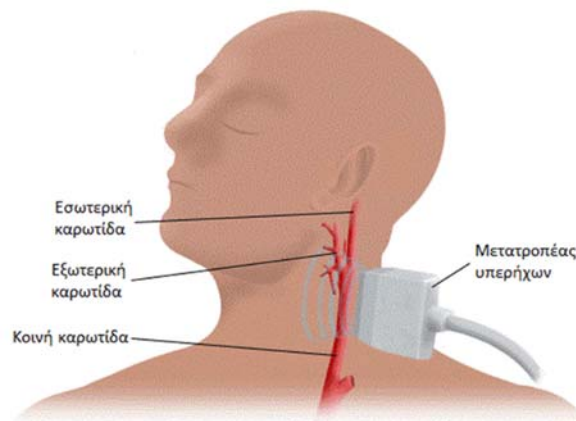
## 1.4 Ο ρόλος και η συνεισφορά του υπερήχου

### 1.4.1 Απεικονιστικές μέθοδοι

Αναμφίβολα, η ραγδαία εξέλιξη της τεχνολογίας έχει συμβάλει σημαντικά στην προαγωγή των ιατρικών επιστημών. Η ανάγκη για ταχεία και έγκαιρη διάγνωση είναι επιτακτική στις μέρες μας και παρουσιάζει μια συνεχώς εξελισσόμενη πορεία. Οι κύριες απεικονιστικές μέθοδοι για τη νόσο της αθηρωματικής πλάκας είναι το υπερηχογράφημα, η αξονική τομογραφία, η μαγνητική τομογραφία και η τομογραφία εκπομπής ποζιτρονίου [29].

Ωστόσο, κυρίαρχη μέθοδος είναι το υπερηχογράφημα (Εικόνα 1.6) το οποίο υπερισχύει των άλλων μεθόδων λόγω των ποικίλων πλεονεκτημάτων που έχει έναντι των άλλων μεθόδων. Τα βασικότερα πλεονεκτήματα αποτελούν η μη επεμβατική φύση του, η μη έκθεση του ασθενή σε ακτινοβολία, το χαμηλό κόστος, το γεγονός ότι είναι μια ταχεία μέθοδος και τέλος του απλού εξοπλισμού που απαιτεί για τη εξέταση και τη διάγνωση. Εκτός του απλού υπερηχογραφήματος, γίνεται χρήση και του υπερηχογραφήματος έγχρωμης ροής (triplex) το οποίο μπορεί και συνδυάζει το απλό υπερηχογράφημα με την ανάλυση της αιματικής ροής. Η χρήση υπερήχων β-σάρωσης είναι η πιο σύνηθες μέθοδος για την εξέταση, απεικόνιση και διάγνωση της νόσου, μέθοδος η οποία συναντάται σε όλα συστήματα καταγραφής υπερήχων [30].

Στη συνέχεια του παρόντος κεφαλαίου θα γίνει εκτενής αναφορά στη χρήση του υπερηχογραφήματος για τη διάγνωση της καρωτιδικής αθηρωμάτωσης, στην αρχή λειτουργίας του υπερήχου, στα πλεονεκτήματα και μειονεκτήματά και τέλος στο ρόλο και στη συνεισφορά του υπερήχου.



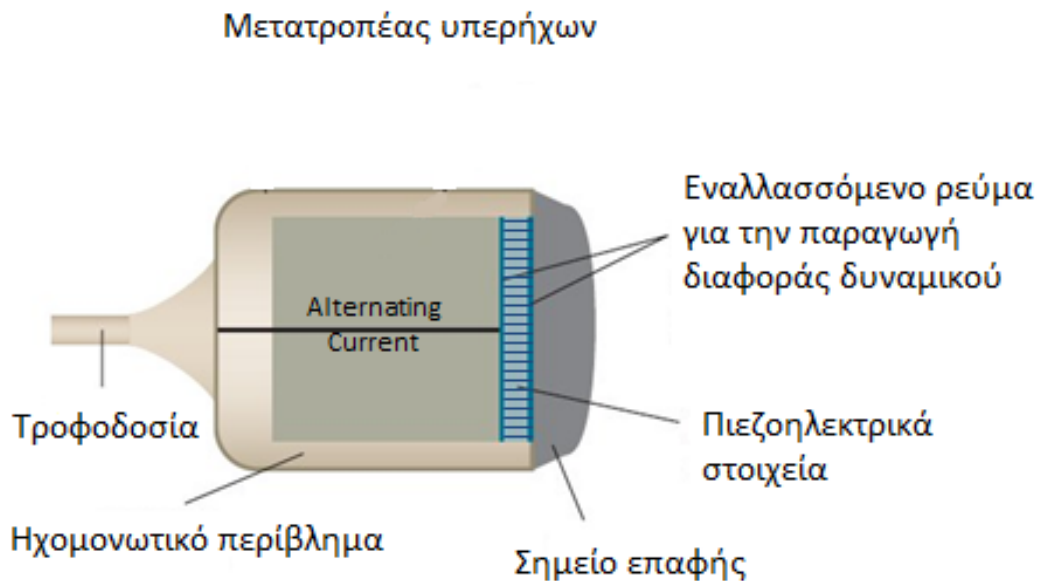
Εικόνα 1.6 : Το υπερηχογράφημα της καρωτιδας [31].



### 1.4.2 Απεικόνιση καρωτίδας με χρήση υπερήχων.

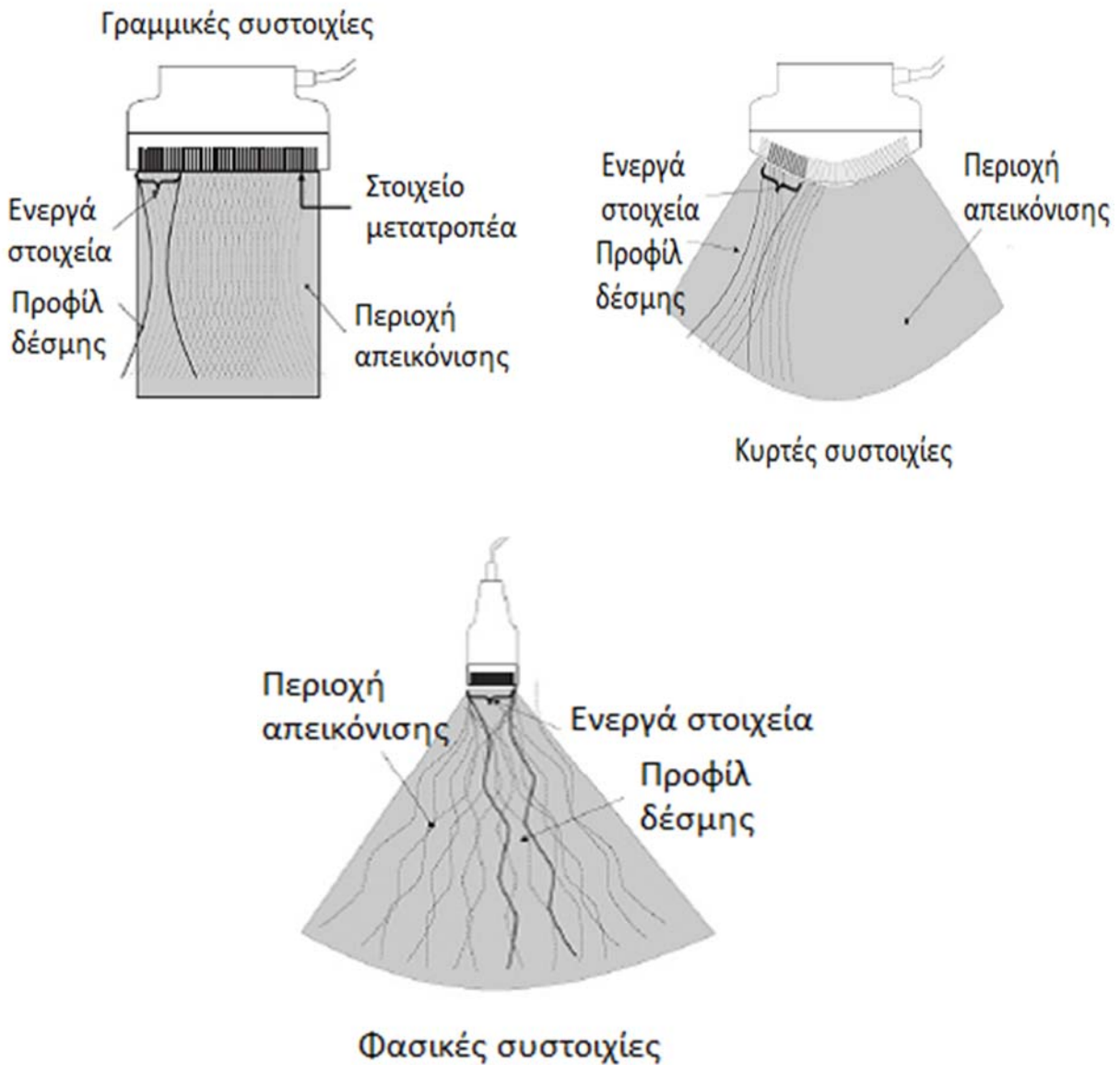
Όπως αναφέρθηκε, η χρήση υπερήχων β-σάρωσης (B-mode Ultrasound) είναι η πιο συνηθισμένη μέθοδος απεικόνισης της καρωτίδας. Αποτελεί μια μη επεμβατική μέθοδο άρα εξαλείφεται η πιθανότητα επιπλοκών, ο ασθενής δεν ακτινοβολείται με ιοντίζουσα ακτινοβολία και δεν ταλαιπωρείται λόγω της ταχείας φύσεως της εξέτασης. Ωστόσο, βασικό μειονέκτημα της μεθόδου αποτελεί το γεγονός της μειωμένης διακριτικής ικανότητας και της χαμηλής ποιότητας της εξαγόμενης απεικόνισης [30].

Η αρχή λειτουργίας του υπερήχου είναι το πιεζοηλεκτρικό φαινόμενο δηλαδή η χρήση κρυστάλλων που έχουν την ιδιότητα να μετατρέπουν την μηχανική ενέργεια σε ηλεκτρική και το αντίστροφο. Οι κρύσταλλοι αυτοί είναι εφικτό να συντονιστούν σε υψηλές ηχητικές συχνότητες της τάξης των 2-15 MHz [32]. Όταν επιτευχθεί ο συντονισμός τους, πραγματοποιείται μετατόπιση του εσωτερικού τους φορτίου και έτσι δημιουργείται ένα δίπολο δηλαδή μια διαφορά δυναμικού στις δύο πλευρές. Αυτή τη διαφορά δυναμικού εκμεταλλεύεται ο μετατροπέας υπερήχων (ultrasound transducer) (Εικόνα 1.7).



Εικόνα 1.7 : Ο μετατροπέας υπερήχων [28]

Τα σύγχρονα απεικονιστικά συστήματα απαρτίζονται από πολλαπλούς μετατροπείς, οι οποίοι οργανώνονται σε γραμμικές (linear array), φασικές (phased-array) και κυρτές (convex array) συστοιχίες (Εικόνα 1.8). Η διαφοροποίηση αυτών των ειδών οφείλεται στην εστίαση της δέσμης και στον τρόπο σάρωσης [33] [34].

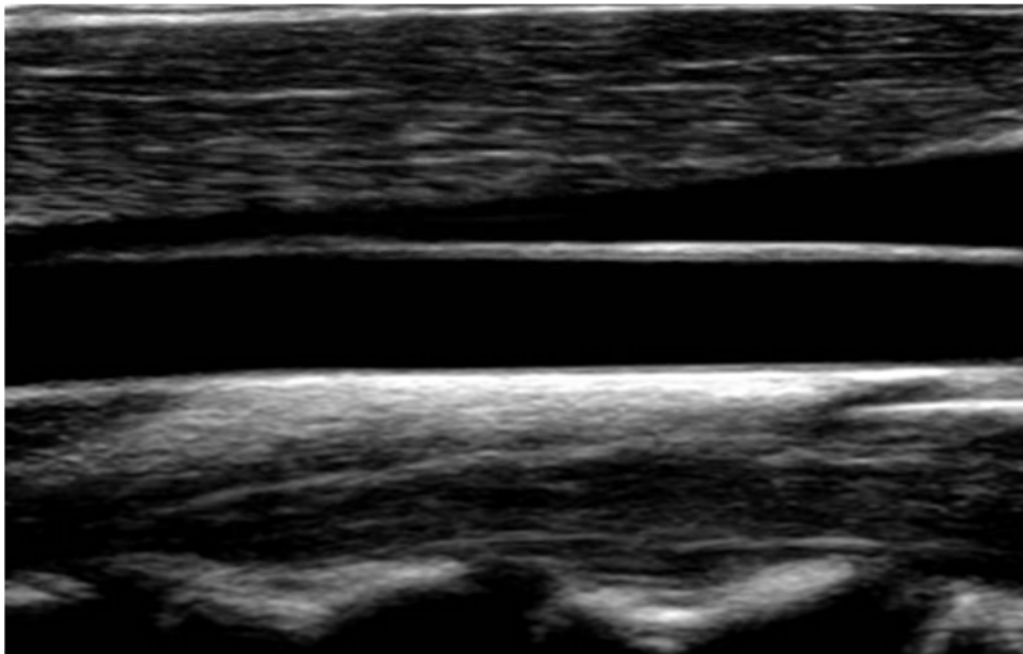


Εικόνα 1.8 : Τα διαφορετικά είδη συστοιχιών του μετατροπέα υπερήχων [1].

Συγκεκριμένα, όσον αφορά την απεικόνιση της καρωτίδας, γίνεται χρήση κυρίως γραμμικής συστοιχίας (linear array) μετατροπέων και τα πιεζοηλεκτρικά στοιχεία (κρύσταλλοι) τοποθετούνται σε σειρά όπως φαίνεται στην Εικόνα 1.5. Τα πιεζοηλεκτρικά αυτά στοιχεία όταν διεγερθούν, δημιουργούν μια δέσμη, η οποία μεταφέρεται στα γειτονικά στοιχεία (σειριακή μεταφορά) και έτσι πραγματοποιείται η σάρωση με την αποστολή ηχητικών σημάτων από όλες τις ομάδες των πιεζοηλεκτρικών στοιχείων [35]. Κάθε σήμα που εκπέμπεται προς το στόχο-περιοχή ενδιαφέροντος σκεδάζεται και ανακλάται.

Η σκεδαζόμενη και ανακλώμενη ενέργεια λαμβάνεται από τον μετατροπέα υπερήχων και έτσι πραγματοποιείται η απεικόνιση του στόχου σε διαβαθμίσεις του γκρι (grey scale). Στη β-σάρωση (b-mode), το πλάτος του σήματος που λαμβάνεται αντιπροσωπεύεται με διαφορετική φωτεινότητα στην λαμβανόμενη απεικονιστική εικόνα, γεγονός που οφείλεται στη διαφορά της ακουστικής εμπέδησης που παρουσιάζει κάθε ιστός λόγω της διαφορετικής του σύστασης και μορφής [33] [34].

Το αποτέλεσμα της β-σάρωσης είναι η απεικόνιση μίας τομής του επιπέδου σε δύο διαστάσεις (2D). Η τομή αυτή (για την ίδια περιοχή πάντα) καταγράφεται σε διαδοχικά στιγμιότυπα ώστε στο τέλος να παράγεται μία ακολουθία βίντεο [36]. Ένα στιγμιότυπο αυτής της ακολουθίας παρουσιάζεται παρακάτω (Εικόνα 1.9).



Εικόνα 1.9 : Στιγμιότυπο κοινής καρωτίδας από β-σάρωση υπερήχου [37]

Όπως αναφέρθηκε ήδη, η ποιότητα της εξαγόμενης εικόνας είναι χαμηλή με αποτέλεσμα η διάγνωση και η ερμηνεία της να εξαρτάται από την εξοικείωση του χρήστη και από το γνωστικό του επίπεδο. Αυτό έχει ως απώτερο αποτέλεσμα οι εκτιμήσεις διαφορετικών ιατρών που αφορούν όμως στην ίδια εικόνα, να παρουσιάζουν σημαντικές αποκλίσεις.

Βασική αιτία για την χαμηλή ποιότητα εικόνας αποτελεί ο speckle θόρυβος που αποτελεί είδος χωρικού θορύβου. Ωστόσο, η ποιότητα της εικόνας μπορεί να βελτιωθεί με τη χρήση ηλεκτρονικής κατευθυνόμενης δέσμης ώστε μια περιοχή-στόχος να μπορεί να απεικονιστεί από διαφορετικές γωνίες χωρίς τη μετακίνηση του μετατροπέα υπερήχων μειώνοντας έτσι τον θόρυβο τύπου speckle. Επίσης μπορεί να πραγματοποιηθεί αποθορυβοποίηση με χρήση υπολογιστικών αλγορίθμων αφαίρεσης θορύβου, κάνοντας έτσι τα αποτελέσματα πιο έγκυρα.

Τέλος, η απεικόνιση της καρωτίδας με χρήση υπερήχων επιτρέπει τον υπολογισμό του βαθμού στένωσης και την εξαγωγή διαφόρων μορφολογικών χαρακτηριστικών της πλάκας [38]. Η ανάλυση μπορεί να πραγματοποιηθεί με τη χρήση κατάλληλων λογισμικών επεξεργασίας εικόνας σε υπολογιστικά συστήματα προκειμένου έτσι, να εξαχθούν διάφορα χαρακτηριστικά.

### **1.4.3 Ανάλυση εικόνων υπερήχου**

#### **Κατάτμηση εικόνας του αρτηριακού τοιχώματος.**

Κατάτμηση εικόνας (image segmentation) ονομάζεται η αυτόματη διαδικασία διαχωρισμού μίας εικόνας, η οποία μπορεί να πραγματοποιηθεί με ελάχιστη ή ακόμα και μηδενική ανθρώπινη παρέμβαση.

Η κατάτμηση της εικόνας μπορεί να πραγματοποιηθεί με δύο διαφορετικούς τρόπους, δηλαδή βασιζόμενη σε περιοχή (region based) ή βασιζόμενη σε όριο (boundary based). Οι προσεγγίσεις που βασίζονται σε περιοχή, μπορούν να κατηγοριοποιηθούν στις μεθόδους κλίσης και αύξησης περιοχών ενώ οι προσεγγίσεις βασιζόμενες σε όριο, αντιπροσωπεύουν τις μεθόδους ανίχνευσης ακμών (edge detection), εφαρμογής κατωφλίου (threshold) και τις μεθόδους ενεργών περιγραμμάτων (snakes κτλ.). Γενικά για την κατάτμηση - τμηματοποίηση των εικόνων, έχουν σχεδιαστεί πολλές τεχνικές. Ωστόσο δεν υπάρχει κάποια γενική μέθοδος ή

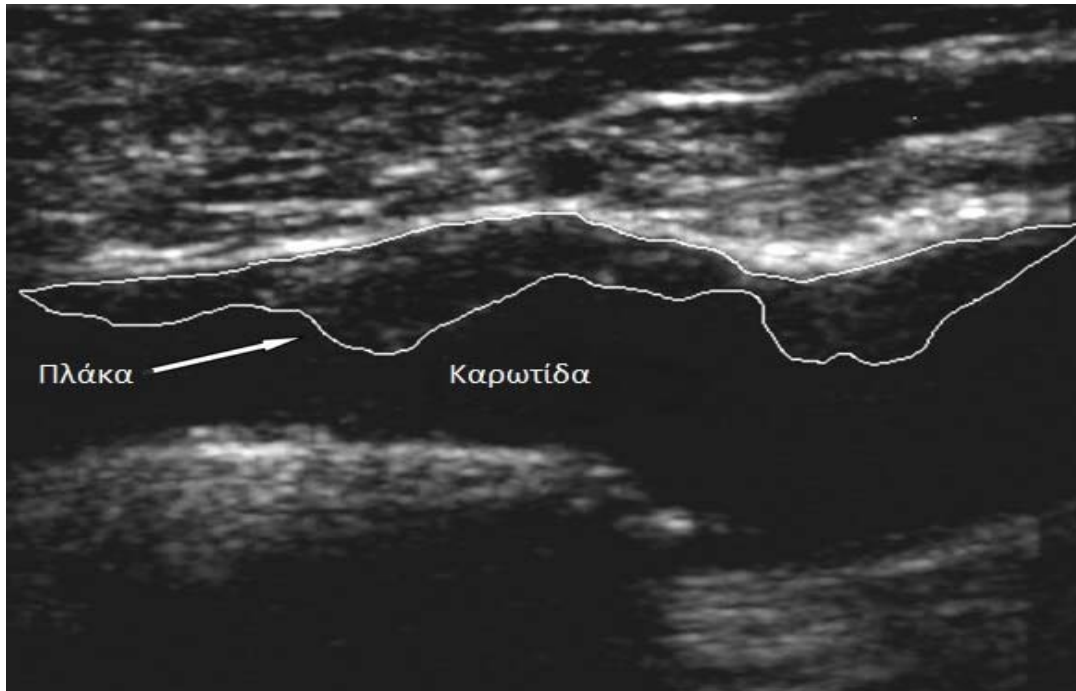
προδιαγραφές που να καθορίζουν εκ των προτέρων την αποτελεσματικότητα και την ευστοχία αυτών των τεχνικών.

Όσον αφορά τη διάγνωση της αθηρωματικής πλάκας, χρησιμοποιείται ένα στιγμιότυπο από την ακολουθία βίντεο της απεικόνισης της καρωτίδας με σκοπό να απομονωθεί η πλάκα (Εικόνα 1.10) και να εξαχθούν περιοχές ενδιαφέροντος του αγγείου. Πολλές φορές προκειμένου να γίνει η αρχικοποίηση του αλγορίθμου κατάτμησης, απαιτείται ο εντοπισμός μιας περιοχής ενδιαφέροντος ROI (Region Of Interest) από ειδικό ή γιατρό.

Μέσω της ανάλυσης των εικόνων του υπερηχογραφήματος, είναι εφικτό να υπολογιστούν χαρακτηριστικά και δείκτες που σχετίζονται άμεσα με τη νόσο. Χαρακτηριστικό παράδειγμα τέτοιου δείκτη αποτελεί ο βαθμός στένωσης του αγγείου δηλαδή με άλλα λόγια, το ποσοστό του εμβαδού του αγγείου που καταλαμβάνεται από την αθηρωματική πλάκα. Από το βαθμό στένωσης μπορεί να γίνει εκτίμηση για το αν χρειάζεται ο ασθενής να υποστεί επέμβαση ή όχι.

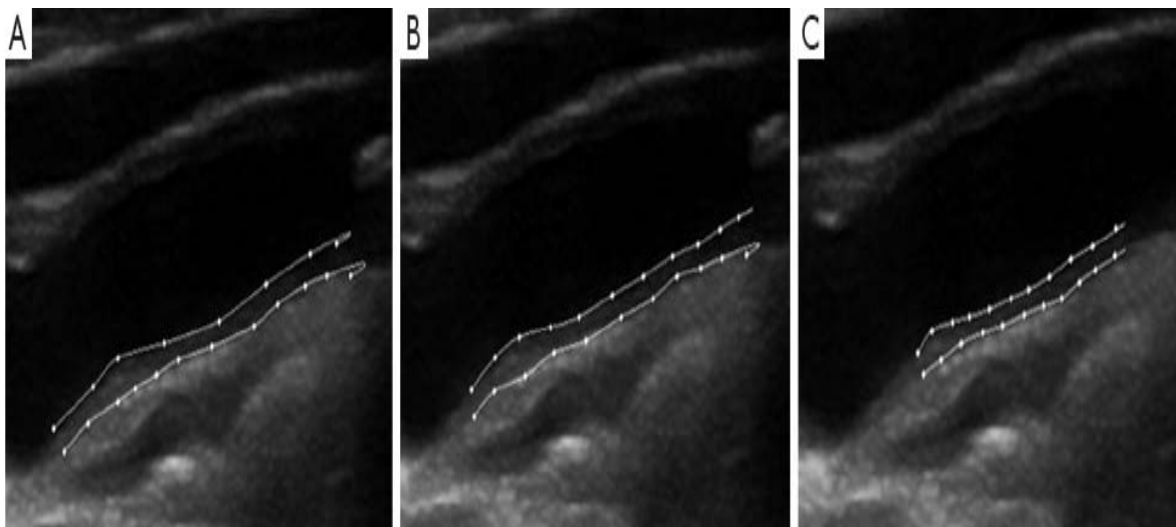
Επιπρόσθετα, ακόμα πιο περίπλοκοι αλγόριθμοι τμηματοποίησης, μπορούν να υπολογίσουν το πάχος του ενδιάμεσου στρώματος (IMT – Intima Media Thickness) και του εσωτερικού στρώματος καθώς επίσης τη διάμετρο και τον όγκο των αγγείων [24]. Το πάχος του ενδιάμεσου στρώματος χρησιμοποιείται για τον έλεγχο της αποτελεσματικότητας κάποιας αγωγής αποκατάστασης και αποτελεί πολλές φορές δείκτη πιθανής ανάπτυξης καρδιοαγγειακών παθήσεων [30].

Επιπλέον, οι αλγόριθμοι κατάτμησης εικόνων χρησιμοποιούνται για την εξαγωγή χαρακτηριστικών κίνησης και υφής και έχουν αντικαταστήσει πλέον τη χειροκίνητη επιλογή της περιοχής ενδιαφέροντος (πλάκα), η οποία απαιτούσε μεγάλη εξειδίκευση και εμπειρία όπως φαίνεται στην εικόνα (Εικόνα 1.11) προκειμένου να αποφευχθούν λάθος επιλογές και κατά επέκταση λανθασμένες μετρήσεις και εκτιμήσεις [39].



Εικόνα 1.10 : Εντοπισμός της αθηρωματικής πλάκας σε εικόνα υπερήχου β-σάρωσης [40].

Όπως φαίνεται από την Εικόνα 1.11 ανάλογα με την εμπειρία του ειδικού επιλέγεται διαφορετική περιοχή ενδιαφέροντος γεγονός που σημαίνει διαφορετική αρχικοποίηση του αλγορίθμου και άρα διαφορετικά αποτελέσματα με αποκλίσεις.



Εικόνα 1.11 : Ανάλυση του πάχους του ενδιάμεσου στρώματος (IMT) από τρεις παρατηρητές A, B και C [41].

## Ανάλυση Κίνησης

Η εφαρμογή κατάλληλων αλγορίθμων σε εικόνες που προέρχονται από υπερηχογράφημα β-σάρωσης, παρέχουν τη δυνατότητα μελέτης της κίνησης του αγγειακού τοιχώματος [42]. Η εφαρμογή αυτών των αλγορίθμων δεν γίνεται σε στιγμιότυπο ακολουθίας όπως οι αλγόριθμοι κατάτμησης αλλά αντιθέτως σε μια ακολουθία-βίντεο. Το μέγεθος που υπολογίζεται από την εφαρμογή αυτών των αλγορίθμων είναι το δισδιάστατο διανυσματικό πεδίο ταχυτήτων ανάμεσα σε διαδοχικές εικόνες-στιγμιότυπα. Αυτού του είδους αλγόριθμοι χωρίζονται σε δύο βασικές κατηγορίες, στους διακριτούς και στους συνεχείς, και διαφέρουν αρκετά [43].

Οι διακριτοί αλγόριθμοι στηρίζουν τη λειτουργία τους στην αντιστοίχιση τμημάτων των εικόνων. Ο πιο γνωστός αλγόριθμος αυτής της κατηγορίας είναι ο αλγόριθμος ταύτισης περιοχών, ο οποίος χρησιμοποιεί σαν περιοχή αναφοράς ένα τμήμα από το πρώτο στιγμιότυπο της ακολουθίας-βίντεο. Στη συνέχεια, με επαναληπτικό τρόπο, επιχειρεί να εντοπίσει στις επόμενες εικόνες, την περιοχή που ταιριάζει περισσότερο στην περιοχή αναφοράς. Αυτό γίνεται μέχρι να προσπελαστεί από τον αλγόριθμο και το τελευταίο στιγμιότυπο της ακολουθίας. Το αποτέλεσμα του συγκεκριμένου αλγορίθμου είναι το διάνυσμα μετατόπισης το οποίο αντιπροσωπεύει την καλύτερη ταύτιση που πραγματοποιήθηκε [43].

Αντίστοιχα, στην κατηγορία των συνεχών αλγορίθμων ανήκει ο αλγόριθμος οπτικής ροής, ο οποίος κυριαρχεί έναντι άλλων αλγορίθμων για την ανάλυση της κίνησης του αρτηριακού τοιχώματος. Οι αλγόριθμοι αυτής της κατηγορίας βασίζονται στην παραγωγή του στιγμιότυπου (εικόνας) και κάνουν χρήση σχέσεων που σχετίζονται και εξαρτώνται από τις χρονικές και χωρικές παραγώγους της εικόνας [30]. Συγκρίνοντας τους διακριτούς με τους συνεχείς αλγορίθμους, μπορεί να διαπιστωθεί πως οι διακριτοί παρουσιάζουν πλεονέκτημα σε σχέση με τους συνεχείς αφού η χρήση παραγώγων που γίνεται από τους συνεχείς αυξάνουν την επίδραση του θορύβου στο τελικό αποτέλεσμα [43].

Συνοψίζοντας, η κίνηση του καρωτιδικού τοιχώματος μπορεί να μελετηθεί με χρήση αλγορίθμων ανάλυσης κίνησης. Μέσω αυτής της μελέτης, μπορούν να εξαχθούν πληροφορίες για την ταχύτητα καθώς και για το πλάτος και τη διεύθυνση της κίνησης του τοιχώματος.

## Ανάλυση υφής

Πριν εφαρμοστεί η ανάλυση υφής, πρέπει να έχει επιλεγθεί η περιοχή της πλάκας μέσω της κατάτμησης της εικόνας (περιγράφηκε προηγουμένως) όπως φαίνεται στην Εικόνα 1.10. Με την ανάλυση υφής εξάγονται δεδομένα που σχετίζονται με την κατανομή της φωτεινότητας της αθηρωματικής πλάκας. Ένα από τα βασικότερα δεδομένα που εξάγονται είναι ο δείκτης της μεσαίας τιμής της φωτεινότητας των εικονοστοιχείων της πλάκας γνωστός και ως GSM (Gray-Scale Median). Η τιμή της GSM συνδέεται άμεσα με τον τύπο της πλάκας (σταθερή ή ασταθής πλάκα).

Πιο συγκεκριμένα, οι ασταθείς πλάκες, αποτελούνται κυρίως από λιπίδια ενώ οι σταθερές πλάκες περιέχουν ασβέστιο σε μεγάλη συγκέντρωση. Η ακουστική εμπέδηση των ασταθών πλακών είναι παραπλήσια με την ακουστική εμπέδηση των γειτονικών μέσων διάδοσης με αποτέλεσμα να απορροφάται μεγάλο ποσοστό των ακτίνων του υπερήχου. Αντίθετα, οι σταθερές πλάκες έχουν διαφορετική ακουστική εμπέδηση σε σχέση με αυτή των γειτονικών μέσων διάδοσης και κατά συνέπεια απορροφάται μικρότερο ποσοστό των ακτίνων. Αυτό έχει σαν αποτέλεσμα, οι ασταθείς πλάκες να μην ανακλούν τις ακτίνες του υπερήχου και να παρουσιάζονται πιο σκοτεινές σε σχέση με τις σταθερές, οι οποίες παρουσιάζονται πιο λευκές αφού ανακλούν τις ακτίνες σε μεγαλύτερο βαθμό [27].

Επομένως, μια πλάκα με χαμηλή φωτεινότητα (σκοτεινή εικόνα) σημαίνει χαμηλή τιμή του δείκτη GSM και κατά συνέπεια αυξημένο κίνδυνο εμφάνισης συμπτωμάτων αφού προειδοποιεί για πιθανή ύπαρξη ασταθούς πλάκας [43]. Οι σταθερές πλάκες αποτελούνται κυρίως από ασβέστιο και εν γένει δεν αποτελούν άμεση απειλή για την υγεία. Επίσης, η τιμή της GSM θα είναι μεγαλύτερη συγκριτικά με των ασταθών [44] [28]. Ένα επιπλέον χαρακτηριστικό που μπορεί να εξαχθεί από την ανάλυση υφής αποτελεί η φράκταλ διάσταση της εικόνας (fractal dimension). Η φράκταλ διάσταση μεγαλώνει αναλόγως με το πόσο τραχεία είναι η επιφάνεια της εικόνας [45]. Τέλος, οι S. Aly et al [28] κατάφεραν να συσχετίσουν τη τιμή της μεσαίας τιμής της φωτεινότητας (GSM) με τις συμπτωματικές και ασυμπτωματικές ομάδες ασθενών. Συγκεκριμένα, διαπίστωσαν πως στους συμπτωματικούς ασθενείς ο δείκτης GSM μειώθηκε καθώς η περιεκτικότητα σε λίπος αυξήθηκε, ενώ στους ασυμπτωματικούς ο δείκτης GSM αυξήθηκε όταν αυξήθηκε η περιεκτικότητα σε λίπος. Από όλα τα παραπάνω γίνεται εμφανής η



τεράστια *διαγνωστική δύναμη* που παρέχεται μέσω του δείκτη GSM όσο αφορά τη νόσο της καρωτιδικής αθηρωματικής πλάκας.

Τα τελευταία χρόνια, οι μελέτες σχετικά με την ανάλυση υφής πραγματοποιούνται με τη χρήση προηγμένων μεθόδων, οι οποίες βασίζονται κατά κύριο λόγο στη μορφολογική επεξεργασία των εικόνων και έχει παρατηρηθεί ότι παρέχουν πιο αποτελεσματική διάκριση μεταξύ σταθερής και ασταθούς πλάκας από ότι οι κλασικές μέθοδοι. Ενδεικτικά, ένα παράδειγμα προηγμένης μεθόδου αποτελεί η εφαρμογή του μορφολογικού τελεστή “άνοιγμα”, ο οποίος εξομαλύνει το περίγραμμα της περιοχής ενδιαφέροντος και εξαλείφει τις προεξοχές [46].

#### **1.4.4 Καθιερωμένοι και ευρέως χρησιμοποιούμενοι απεικονιστικοί δείκτες**

Στη συγκεκριμένη εργασία δόθηκε μεγάλη έμφαση στους δείκτες που σχετίζονται με το βαθμό στένωσης, το εμβαδό της πλάκας και με τη τιμή GSM (Gray-scale median). Το εμβαδόν της πλάκας σχετίζεται άμεσα με το πλάτος και την επιφάνειά της πλάκας. Η συνεχής συσσώρευση λιπιδίων όπως τα τριγλυκερίδια στον εσωτερικό χιτώνα της αρτηρίας έχει σαν αποτέλεσμα τη συνεχή αύξηση του πάχους της πλάκας. Αυτό με τη σειρά του οδηγεί στην αύξηση του βαθμού στένωσης και του εμβαδού της πλάκας. Ακριβώς το ίδιο ισχύει και για τη συσσώρευση του λιπώδους συστατικού της χοληστερόλης στα τοιχώματα των αρτηριών. Το γεγονός αυτό επιβεβαιώνεται από έρευνα που έγινε με 76 ασθενείς, οι οποίοι χωρίστηκαν σε δυο ομάδες ανάλογα με το πάχος της καρωτιδικής πλάκας σε υψηλού κινδύνου και χαμηλού κινδύνου. Συσχετίστηκε η HDL χοληστερόλη με την αύξηση του πάχους της πλάκας και κατά επέκταση με την καρωτιδική αλλοίωση [13].

Η λαμβανόμενη ηχώ από την καρωτιδική πλάκα που ποσοτικοποιείται από την τιμή του gray-scale median (GSM) έχει συσχετιστεί με το βαθμό σταθερότητας της πλάκας. Παρατηρήθηκε ότι η εντατική θεραπεία για μείωση λιπιδίων σε ασθενείς βελτίωσε τη τιμή του GSM. Το πλήθος των ασθενών στη συγκεκριμένη μελέτη ήταν 97 με στένωση καρωτίδας και η διάρκεια της θεραπείας διήρκεσε 6 μήνες. Δείγματα γενικών βιοχημικών δεικτών και GSM λήφθηκαν στην αρχή και μετά το πέρας 6 μηνών. Το αποτέλεσμα της μελέτης ήταν οι τιμές του

GSM να αυξηθούν σημαντικά λόγω της μείωσης των λιπιδίων. Επομένως η αντίστροφη σχέση μεταξύ λιπιδίων και τιμής του GSM αναδείχθηκε μέσω αυτής της μελέτης [13].

Επιπρόσθετα μελετήθηκε στα πλαίσια έρευνας και η σχέση μεταξύ αντιποκίνων και retinol-binding πρωτεΐνη 4 (RBP4) με τη σταθερότητα της πλάκας και κατά επέκταση με τη τιμή του GSM. Στη συγκεκριμένη μελέτη συμμετείχαν 225 ασθενείς με υψηλά επίπεδα στένωσης της καρωτίδας και 75 ασυμπτωματικοί ασθενείς με χαμηλά επίπεδα στένωσης της καρωτίδας. Σε όλους τους συμμετέχοντες έγινε εκτίμηση του GSM μέσω υπερήχου. Η RBP4 βρέθηκε πιο αυξημένη στους ασθενείς με υψηλή στένωση της καρωτίδας σε σχέση με αυτούς που είχαν χαμηλή. Συμπερασματικά, η RBP4 συσχετίστηκε θετικά με τη σοβαρότητα της αθηροσκλήρωσης και με την τιμή του δείκτη GSM [23].

Με την συσσώρευση μεγάλων συγκεντρώσεων λιπιδίων στο εσωτερικό της αρτηρίας, παρατηρείται μεγαλύτερη εξασθένιση του υπερήχου με αποτέλεσμα την ελάττωση του δείκτη της μεσαίας τιμής της φωτεινότητας των εικονοστοιχείων της πλάκας (GSM) συγκριτικά με την τιμή που λαμβάνει σε σταθερές πλάκες. Συμπερασματικά, ο δείκτης GSM παρουσιάζει μικρότερη τιμή για ασταθή πλάκα, παράμετρος που συντελεί στην αύξηση του βαθμού επικινδυνότητας για τον ασθενή [14].

Τέλος, διαπιστώθηκε ότι σε συμπτωματικούς ασθενείς, η αύξηση της περιεκτικότητας σε λίπος οδηγεί στην ελάττωση της τιμής του δείκτη GSM. Αντίθετα, σε ασυμπτωματικούς ασθενείς παρατηρήθηκε αύξηση του δείκτη με αύξηση σε περιεκτικότητας λίπους, γεγονός που οδήγησε στην συσχέτιση του δείκτη με συμπτωματικές και ασυμπτωματικές ομάδες ασθενών [15].

## 1.5 Σκοπός της παρούσας εργασίας

Με βάση τα παραπάνω στοιχεία, τα οποία αναδεικνύουν τόσο την πολυπλοκότητα της νόσου, όσο και τον σημαντικό ρόλο της πολυπαραγοντικής περιγραφής της νόσου, σκοπός της παρούσας εργασίας είναι η αναγνώριση διαφορετικών φαινοτύπων της αθηρωματικής πλάκας βασιζόμενοι σε συμβατικούς και πιο εξειδικευμένους βιοχημικούς δείκτες που έχουν χρησιμοποιηθεί στη σχετική βιβλιογραφία. Για το σκοπό αυτό, συνδυάστηκαν τεχνικές συσταδοποίησης και παραγωγής δένδρογράμματος για την ανίχνευση των φαινοτύπων. Οι

κατηγοριοποιήσεις ερμηνεύτηκαν μέσω συσχετίσεων με καθιερωμένους απεικονιστικούς δείκτες αξιολόγησης της νόσου (βαθμός στένωσης, GSM και συνολικό εμβαδό αθηρωματικών αλλοιώσεων), καθώς και με τη συμπτωματολογία των αντίστοιχων κλινικών περιπτώσεων.

Στη συνέχεια, παρουσιάζεται η μεθοδολογία και τα δεδομένα που χρησιμοποιήθηκαν και ακολουθούν τα αποτελέσματα της μελέτης. Τέλος, συνοψίζονται τα συμπεράσματά μας και παρουσιάζονται πιθανές μελλοντικές προοπτικές της εργασίας.

# Κεφάλαιο 2 : Μεθοδολογία

## Περίληψη

Στο παρόν κεφάλαιο παρουσιάζεται αρχικά η έννοια της ανάλυσης κατά συστάδες ή ανάλυση σε ομάδες (Cluster Analysis). Στη συνέχεια γίνεται αναφορά στα μέτρα απόστασης , στα μέτρα ομοιότητας και στην ιεραρχική μέθοδο συσταδοποίησης (Hierarchical Clustering), η οποία χρησιμοποιήθηκε στα πλαίσια αυτής της εργασίας ώστε να πραγματοποιηθεί η ομαδοποίηση των ασθενών και η αναγνώριση των φαινοτύπων. Επιπλέον αναλύονται τα βήματα του αλγορίθμου τα οποία οδήγησαν στο τελικό αποτέλεσμα. Τέλος περιγράφεται η συλλογή δεδομένων (dataset) που χρησιμοποιήθηκε σαν είσοδος του αλγορίθμου.

## 2.1 Θεωρητικό υπόβαθρο του αλγορίθμου

### 2.1.1 Ανάλυση κατά συστάδες

Με τον όρο ανάλυση κατά συστάδες (Cluster Analysis) αναφερόμαστε στη μέθοδο η οποία έχει σκοπό την κατάταξη υπάρχουσών παρατηρήσεων σε ομάδες. Συγκεκριμένα, η ανάλυση κατά συστάδες εξετάζει την ομοιότητα που παρουσιάζουν κάποιες παρατηρήσεις, με σκοπό να δημιουργήσει ομάδες από παρατηρήσεις που μοιάζουν μεταξύ τους δηλαδή να κάνει ένα είδος ομαδοποίησης όμοιων χαρακτηριστικών [48].

Η ανάλυση κατά συστάδες έχει χρησιμοποιηθεί ευρέως σε όλα τα επιστημονικά πεδία προκειμένου να ομαδοποιήσει δεδομένα και ο ρόλος της είναι καθοριστικός στην ανάπτυξη πολλών επιστημονικών κλάδων όπως στη συγκεκριμένη περίπτωση, της βιοιατρικής.

Βασικές έννοιες της μεθόδου αποτελεί η έννοια της απόστασης και της ομοιότητας. Παρατηρήσεις που μοιάζουν μεταξύ τους θα έχουν μεγάλη τιμή ομοιότητας και η απόσταση μεταξύ τους θα είναι μικρή. Επομένως, οι δύο αυτές έννοιες αποτελούν σημαντικό κομμάτι της μεθόδου αφού μας επιτρέπουν να μετρήσουμε πόσο μοιάζουν οι παρατηρήσεις- δεδομένα μεταξύ τους και να τα κατηγοριοποιήσουμε στην ίδια ομάδα [49].

Η ομαδοποίηση (Clustering), διαφοροποιείται από τις τεχνικές ταξινόμησης. Οι τεχνικές ταξινόμησης αναφέρονται σε ένα γνωστό αρχικό πλήθος ομάδων και κατατάσσονται στα συστήματα επιβλεπόμενης μάθησης (Supervised learning) δηλαδή γνωρίζουμε στην ουσία χαρακτηριστικά της εισόδου και της εξόδου και προσπαθούμε να μοντελοποιήσω το σύστημα για την συγκεκριμένη είσοδο και έξοδο. Τέτοιου είδους τεχνικές μπορούν να χρησιμοποιηθούν για την μελέτη της πάθησης Alzheimer στην οποία είναι γνωστές οι επιπτώσεις (έξοδος) καθώς και τις μεταλλάξεις που την δημιουργούν (είσοδος) και στόχος είναι να βρεθεί το μοντελοποιημένο σύστημα της νόσου [50].

Στη συσταδοποίηση δεν γίνεται καμία αρχική υπόθεση σχετικά με τον αριθμό των ομάδων ή τη δομή των ομάδων δηλαδή δεν παρέχεται στον αλγόριθμο η μεταβλητή εξόδου. Η ανάλυση σε συστάδες υπάγεται στην μάθηση χωρίς επίβλεψη (unsupervised learning). Η κατηγοριοποίηση γίνεται με βάση την ομοιότητα ή την απόσταση των αντικειμένων [51] [52].

Η συσταδοποίηση μπορεί να πραγματοποιηθεί με αρκετές μεθόδους. Οι κυρίαρχες από αυτές τις μεθόδους είναι οι ιεραρχικές και οι μη ιεραρχικές μέθοδοι.

Κατά την ιεραρχική μέθοδο (hierarchical cluster analysis) ο αριθμός των ομάδων δεν είναι γνωστός από την αρχή. Ξεκινάμε θεωρώντας ότι κάθε αντικείμενο αποτελεί από μόνο του μια ομάδα. Σε κάθε βήμα του αλγορίθμου συγχωνεύονται τα δύο αντικείμενα που έχουν την πιο μικρή απόσταση αφού προηγουμένως έχει υπολογιστεί ο πίνακας αποστάσεων δηλαδή οι αποστάσεις όλων των παρατηρήσεων από τις υπόλοιπες. Η συγκεκριμένη μέθοδος θα αναλυθεί εκτενώς στην επόμενη ενότητα του παρόντος κεφαλαίου [53].

Αντίθετα, στην μη ιεραρχική μέθοδο (Non-hierarchical cluster analysis) το πλήθος των ομάδων είναι γνωστό εξ αρχής. Με χρήση αλγορίθμων κατηγοριοποιούμε τα αντικείμενα (παρατηρήσεις) στις ομάδες αναλόγως με το ποια ομάδα είναι πιο κοντά στο εκάστοτε αντικείμενο. Με άλλα λόγια ο στόχος των μη ιεραρχικών μεθόδων είναι η ομαδοποίηση η ομάδων σε έναν προκαθορισμένο αριθμό ομάδων. Σε αυτή την κατηγορία μεθόδου ανήκει και ο αλγόριθμος K-mean.

Μεγάλο και διαχρονικό πρόβλημα αποτελεί η εύρεση του βέλτιστου αριθμού των ομάδων (Optimal number of clusters) ενός συνόλου δεδομένων-αντικειμένων. Οι αλγόριθμοι που αναφέρθηκαν παραπάνω ομαδοποιούν τα δεδομένα σε ομάδες χωρίς όμως να προσδιορίζουν τον ακριβή αριθμό ομάδων από τις οποίες απαρτίζονται τα δεδομένα [54] [55]. Για την αντιμετώπιση αυτού του προβλήματος έχουν αναπτυχθεί αρκετές μέθοδοι ωστόσο κάθε πρόβλημα ομαδοποίησης σε ομάδες, χρειάζεται τη δική του προσέγγιση. Στη συγκεκριμένη διπλωματική εργασία ο βέλτιστος αριθμός των ομάδων υπολογίστηκε βάσει τη μεθοδολογία που περιγράφεται στο άρθρο που αποτέλεσε έμπνευση της παρούσας εργασίας [47].

Τέλος υπάρχουν αλγόριθμοι οι οποίοι μπορούν να δεχτούν σαν είσοδο ,εκτός από το σύνολο των παρατηρήσεων, τον βέλτιστο αριθμό συστάδων και έτσι να εξάγουν ένα αποτέλεσμα ομαδοποίησης σε περιορισμένο αριθμό ομάδων.

### 2.1.2 Μέτρα Απόστασης

Τα μέτρα απόστασης χρησιμοποιούνται ευρέως στις διαδικασίες ομαδοποίησης και ουσιαστικά εκφράζουν το πόσο δύο αντικείμενα είναι όμοια ή ανόμοια μεταξύ τους. Αντικείμενα τα οποία μοιάζουν μεταξύ τους θα παρουσιάζουν μικρή τιμή απόστασης και έτσι θα συγχωνεύονται στην ίδια ομάδα (cluster).

Τα δεδομένα (data) που χρησιμοποιούνται για ομαδοποίηση αποτελούνται από  $n$  παρατηρήσεις (subjects). Κάθε μία από τις παρατηρήσεις αναπαρίσταται από ένα διάνυσμα της μορφής  $x_i = (x_{i1}, x_{i2}, \dots, x_{ik})$ , όπου  $k$  είναι ο αριθμός των μεταβλητών που περιγράφουν μία παρατήρηση [28]. Ένα σύνολο παρατηρήσεων (subjects) μπορεί να παρασταθεί ως ένα σύνολο  $n$  σημείων στο χώρο. Έτσι μπορεί να κατασκευαστεί ένα πίνακας με  $n$  γραμμές και  $k$  στήλες δηλαδή τις γραμμές θα αποτελούν οι παρατηρήσεις-αντικείμενα και τις στήλες οι μεταβλητές που περιέχονται σε κάθε παρατήρηση-αντικείμενο. Συγκεκριμένα στην παρούσα διπλωματική εργασία, οι γραμμές αντιπροσωπεύονται από τους ασθενείς και οι στήλες από δείκτες εξετάσεων (συμβατικών, εξειδικευμένων βιοχημικών δεικτών κτλ) κάθε ασθενή. Ο πίνακας αυτός ονομάζεται πίνακας δεδομένων και θα αναλυθεί στη συνέχεια του κεφαλαίου.

Η απόσταση ( $D(X, \Psi)$ ) μεταξύ δύο αντικειμένων  $X$  και  $\Psi$  ονομάζεται μετρική (metric) όταν ικανοποιεί τις παρακάτω ιδιότητες [56]:

- 1) Αν  $X = \Psi$ , τότε  $D(X, \Psi) = 0$  δηλαδή η απόσταση ενός αντικειμένου από τον εαυτό του είναι μηδενική.
- 2) Αν  $X \neq \Psi$ , τότε  $D(X, \Psi) > 0$  δηλαδή η απόσταση μεταξύ δυο διαφορετικών αντικειμένων δεν είναι ποτέ αρνητική.
- 3) Η απόσταση είναι συμμετρική και δεν έχει σημασία η σειρά υπολογισμού δηλαδή  $D(X, \Psi) = D(\Psi, X)$ .
- 4) Αν ισχύει η τριγωνική ιδιότητα:  $D(X, \Psi) + D(\Psi, Z) \geq D(X, Z)$  δηλαδή όταν το μήκος της διαδρομής από το πρώτο στο τρίτο αντικείμενο ( $D(X, Z)$ ) δεν μπορεί να είναι μεγαλύτερο από το μήκος της διαδρομής που διέρχεται από το δεύτερο αντικείμενο [46].

Μια απόσταση ονομάζεται ημι-μετρική (semi-metric) όταν ικανοποιεί τις τρεις πρώτες ιδιότητες που αναφέρθηκαν παραπάνω.

Στην ιεραρχική μέθοδο ομαδοποίησης χρησιμοποιούνται μετρικές αποστάσεις προκειμένου να υπολογιστούν οι ομοιότητες και οι ανομοιότητες των αντικειμένων. Οι βασικότερες μετρικές αποστάσεις που χρησιμοποιούνται στα προβλήματα ομαδοποίησης παρατίθενται παρακάτω.

### Μετρικές αποστάσεις

#### ➤ Ευκλείδεια Απόσταση (Euclidean distance)

Το πιο γνωστό μέτρο απόστασης για την μέτρηση της ομοιότητα ή ανομοιότητας δυο αντικειμένων είναι η Ευκλείδεια απόσταση, η οποία εκφράζει τη γεωμετρική απόσταση στο χώρο [39]. Η Ευκλείδεια απόσταση ορίζεται από τη σχέση :

$$d(x_i, x_j) = \sqrt{\sum_{r=1}^k (x_{ir} - x_{jr})^2} \quad (2.1)$$

Όπου :  $x_i, x_j$  είναι τα αντικείμενα (οι ασθενείς στη συγκεκριμένη περίπτωση) και  $x_{ir}, x_{jr}$  είναι τα δεδομένα που περιέχονται στα αντικείμενα (οι τιμές των εξετάσεων κάθε ασθενή στην προκείμενη περίπτωση) και  $k$  είναι ο αριθμός των μεταβλητών που περιγράφουν κάθε αντικείμενο.

Για την κατασκευή του πίνακα αποστάσεων  $n$  αντικειμένων  $x_1, x_2, \dots, x_n$ , υπολογίζουμε αρχικά τις αποστάσεις τους  $d(x_i, x_j)$  για  $i, j = 1, 2, \dots, n$  και τις τοποθετούμε σε ένα πίνακα  $D$  με  $n$  στήλες και  $n$  γραμμές. Τα στοιχεία της διαγώνιου του πίνακα  $D$  θα είναι μηδέν λόγω της ιδιότητας 1 που αναφέρθηκε παραπάνω και ο πίνακας θα είναι συμμετρικός [47].

Η ευκλείδεια απόσταση αποτελεί τη πιο γνωστή και απλή μέθοδο υπολογισμού των αποστάσεων δυο αντικειμένων. Ωστόσο ένα βασικό μειονέκτημα της μεθόδου αποτελεί το



γεγονός ότι οι αποστάσεις επηρεάζονται σε μεγάλο βαθμό από την διαφοροποίηση της κλίμακας μεταξύ των στοιχείων από τις οποίες υπολογίζονται. Αυτό το γεγονός γεννά την ανάγκη του μετασχηματισμού των τιμών των στοιχείων ώστε να αποκτήσουν συγκρίσιμη κλίμακα δηλαδή ενός είδους κανονικοποίησης των τιμών [58]. Πλέον για την αντιμετώπιση του συγκεκριμένου προβλήματος έχουν αναπτυχθεί πολλές μέθοδοι όπως ο υπολογισμός της απόστασης Pearson (Pearson correlation).

### ➤ Απόσταση Pearson

Η απόσταση Pearson παρέχει καλύτερες συγκρίσεις σε σχέση με την ευκλείδεια απόσταση όταν οι μεταβλητές παρουσιάζουν μεγάλη διαφοροποίηση κλίμακας. Για να φέρουμε τις μεταβλητές σε συγκρίσιμη κλίμακα, κάθε μεταβλητή διαιρείται με την τυπική της απόκλιση. Έτσι η απόσταση Pearson περιγράφεται από την παρακάτω σχέση [60]:

$$d(x_i, x_j) = \sqrt{\sum_{r=1}^k \frac{(x_{ir} - x_{jr})^2}{s_r^2}} \quad (2.2)$$

Όπου  $s_r$  η διακύμανση της  $r$  μεταβλητής και  $k$  είναι ο αριθμός των μεταβλητών που περιγράφουν κάθε αντικείμενο.

### ➤ Απόσταση Mahalanobis

Η απόσταση Mahalanobis ανταποκρίνεται ορθά σε δεδομένα με εμφανείς διακυμάνσεις [61]. Εκφράζεται από τη σχέση :

$$(d(x_i, x_j))^2 = (x_i - x_j) * S^{-1}(x_i - x_j)' \quad (2.3)$$

Όπου  $S$  είναι ο πίνακας διακύμανσης. Όταν ο  $S$  είναι διαγώνιος με διαγώνια στοιχεία τα  $s_j^2$ , τότε η απόσταση που προκύπτει συμπίπτει με την απόσταση Pearson.

### ➤ Απόσταση Manhattan

Η απόσταση Manhattan εκφράζεται από τη σχέση :

$$d(x_i, x_j) = \sum_{r=1}^k |x_{ir} - x_{jr}| \quad (2.4)$$

Όπου :  $x_i, x_j$  είναι τα αντικείμενα (οι ασθενείς στη συγκεκριμένη περίπτωση),  $x_{ir}, x_{jr}$  είναι τα δεδομένα που περιέχονται στα αντικείμενα (οι τιμές των εξετάσεων κάθε ασθενή στην προκείμενη περίπτωση) και  $k$  είναι ο αριθμός των μεταβλητών που περιγράφουν κάθε αντικείμενο.

Η απόσταση Manhattan ισούται ουσιαστικά με τη μέση διαφορά μεταξύ των μεταβλητών που χαρακτηρίζουν κάθε αντικείμενο  $X$  και τα αποτελέσματα που παρέχει είναι κοντινά στα αποτελέσματα της ευκλείδειας απόστασης [62].

### ➤ Απόσταση Minkowski

Η απόσταση Minkowski δίνεται από τη σχέση :

$$d(x_i, x_j) = \left( \sum_{r=1}^k |x_{ir} - x_{jr}|^\lambda \right)^{\frac{1}{\lambda}} \quad (2.5)$$

Όπου :  $x_i, x_j$  είναι τα αντικείμενα,  $x_{ir}, x_{jr}$  είναι τα δεδομένα που περιέχονται στα αντικείμενα και  $k$  είναι ο αριθμός των μεταβλητών που περιγράφουν κάθε αντικείμενο όπως ακριβώς και στην προηγούμενη περίπτωση. Για τιμή του  $\lambda$  ίση με 1 η απόσταση Minkowski ταυτίζεται με την απόσταση Manhattan ενώ για  $\lambda$  ίσο με 2, ταυτίζεται με την Ευκλείδεια απόσταση [63].

### ➤ Η απόσταση μεγίστου ή απόσταση Chebyshev

Η απόσταση μεγίστου ορίζει πως η απόσταση μεταξύ δυο αντικειμένων είναι ίση με τη μέγιστη απόσταση των μεταβλητών χαρακτηρίζουν αυτά τα αντικείμενα και δίνεται από την παρακάτω σχέση [63].

$$d(x_i, x_j) = \max_{r=1,2,\dots,k} (|x_{ir} - x_{jr}|) \quad (2.6)$$

Με  $x_i, x_j$  τα αντικείμενα,  $x_{ir}, x_{jr}$  τα δεδομένα που περιέχονται στα αντικείμενα και  $k$  ο αριθμός των μεταβλητών που περιγράφουν κάθε αντικείμενο.

### 2.1.3 Μέτρα Ομοιότητας

Με τη χρήση των μέτρων ομοιότητας μπορεί να γίνει η εκτίμηση δύο παρατηρήσεων για το αν μοιάζουν ή όχι μεταξύ τους. Παρατηρήσεις που μοιάζουν σε μεγάλο βαθμό έχουν μεγάλη τιμή ομοιότητας ενώ παρατηρήσεις που διαφέρουν έχουν μικρή τιμή ομοιότητας. Η πληροφορία αυτή είναι πού χρήσιμη όταν πρόκειται να πραγματοποιηθεί ομαδοποίηση παρατηρήσεων αφού στόχος αποτελεί η ομαδοποίηση όσο πιο όμοιων παρατηρήσεων γίνεται σε κοινές ομάδες.

Ο πιο γνωστός και πιο ευρέως χρησιμοποιούμενος δείκτης υπολογισμού της ομοιότητας είναι ο συντελεστής συσχέτισης Pearson (Pearson's Correlation Coefficient) [64] ο οποίος δίνεται από τη σχέση :

$$s_{ij} = \frac{\sum_{r=1}^k (x_{ir} - \bar{x}_i) * (x_{jr} - \bar{x}_j)}{\sqrt{\sum_{r=1}^k (x_{ir} - \bar{x}_i)^2 * \sum_{r=1}^k (x_{jr} - \bar{x}_j)^2}} \quad (2.7)$$

Με  $\bar{x}_i = \frac{1}{k} * \sum_{r=1}^k (x_{ir})$  ,  $\bar{x}_j = \frac{1}{k} * \sum_{r=1}^k (x_{jr})$  και  $x_i, x_j$  οι παρατηρήσεις,  $x_{ir}, x_{jr}$  είναι τα δεδομένα που περιέχονται στις παρατηρήσεις και  $k$  είναι ο αριθμός των μεταβλητών που περιγράφουν κάθε παρατήρηση.

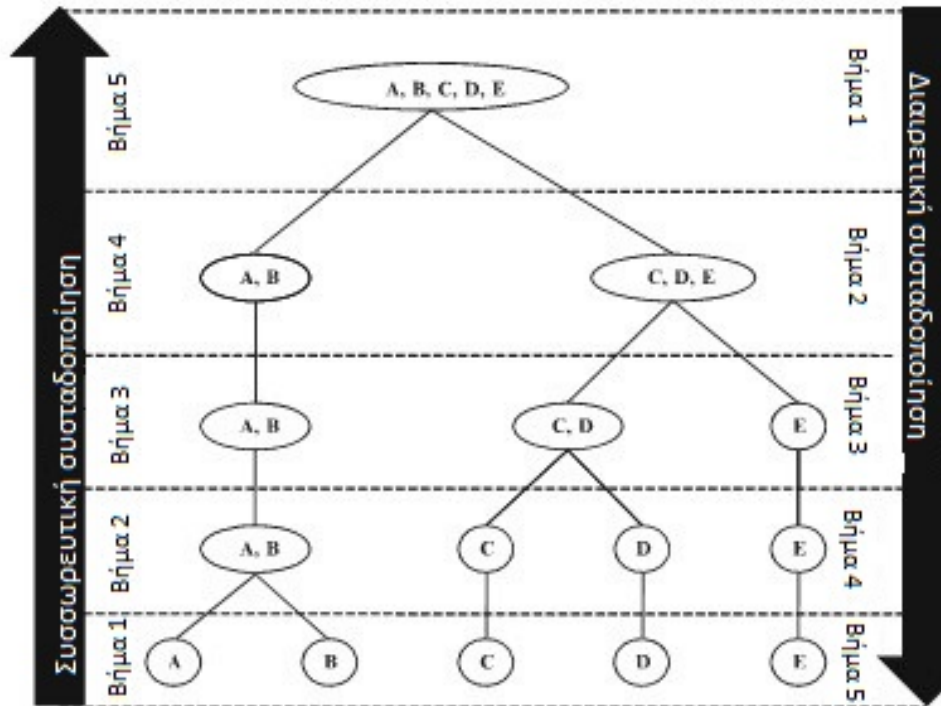
Ο συντελεστής συσχέτισης Pearson χρησιμοποιείται κυρίως για την ομαδοποίηση μεταβλητών και θα χρησιμοποιηθεί στη μεθοδολογία που υλοποιεί η παρούσα διπλωματική εργασία. Τέλος, πρέπει να αναφερθεί πως υπάρχουν και άλλοι δείκτες συσχέτισης στη βιβλιογραφία αλλά δεν θα αναλυθούν στην παρούσα εργασία. Ενδεικτικά αναφέρονται ονομαστικά μερικά μέτρα ομοιότητας : Simple matching [66], Rogers and Tanimoto [67], Sokal and Sneath [68], Jaccard coefficient και Dice and Sorensen [67].

### 2.1.4 Κατηγορίες ιεραρχικών μεθόδων συσταδοποίησης

Οι ιεραρχικές μέθοδοι ομαδοποίησης προχωρούν είτε με μία σειρά διαδοχικών διαιρέσεων είτε με μια σειρά διαδοχικών συγχωνεύσεων. Έτσι χωρίζονται αντίστοιχα σε διαιρετικές τεχνικές (divisive) ή συσσωρευτικές (agglomerative).

Οι συσσωρευτικές ιεραρχικές μέθοδοι ομαδοποίησης ξεκινούν θεωρώντας κάθε αντικείμενο σαν ξεχωριστή ομάδα (cluster) επομένως έχουν τόσες αρχικές ομάδες όσες και τα αρχικά αντικείμενα. Τα πιο όμοια αντικείμενα ομαδοποιούνται και οι αρχικές αυτές ομάδες συγχωνεύονται σε μία με αποτέλεσμα να έχουμε μια ομάδα (cluster) λιγότερη. Η διαδικασία επαναλαμβάνεται έως ότου όλες οι αρχικές ομάδες συγχωνευτούν σε μία τελική ομάδα (cluster) [69].

Αντίθετα δουλεύουν οι διαιρετικές μέθοδοι ιεραρχικής ομαδοποίησης, οι οποίες ξεκινούν από μία ενιαία ομάδα και στο πρώτο βήμα την χωρίζουν σε δύο ομάδες. Σκοπός αυτής της διάσπασης-διαίρεσης είναι η παραγωγή δύο νέων ομάδων με τη μεγαλύτερη ανομοιότητα δηλαδή να διαφέρουν μεταξύ τους κατά πολύ. Αυτές οι δύο υποομάδες διαιρούνται με η σειρά τους σε νέες ανόμοιες υποομάδες. Η διαδικασία συνεχίζεται μέχρι να παραχθούν τόσες ομάδες όσες και τα αρχικά μας αντικείμενα δηλαδή μέχρι κάθε αντικείμενο να αποτελεί μία ομάδα [53].

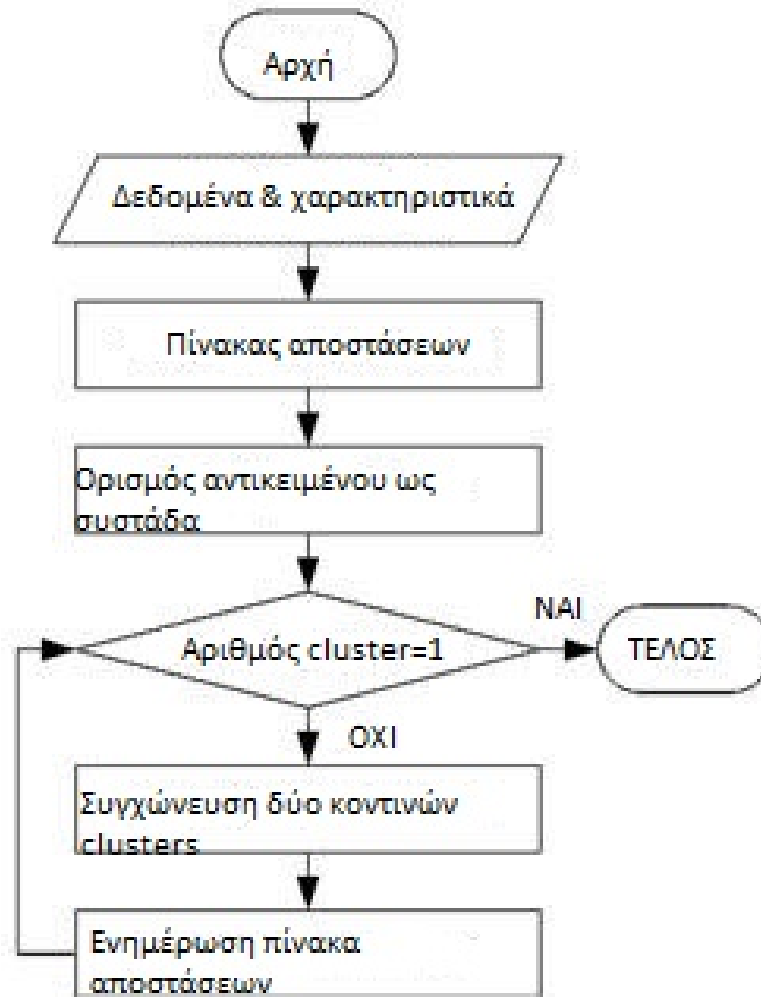


Εικόνα 2.1 : Συσσωρευτική και διααιρετική ομαδοποίηση [70].

Τα επίπεδα ιεραρχίας τόσο για την συσσωρευτική μέθοδο ιεραρχικής ομαδοποίησης όσο και για τη μέθοδο διααιρετικής ιεραρχικής ομαδοποίησης είναι σε πλήθος  $N-1$  όπου  $N$  το πλήθος των αρχικών αντικειμένων (Εικόνα 2.1). Στη παρούσα εργασία έγινε χρήση συσσωρευτικής ιεραρχικής μεθόδου (agglomerative clustering method) για την ομαδοποίηση ασθενών σε ομάδες θεωρώντας στην αρχή κάθε ασθενή να αποτελεί μια ξεχωριστή ομάδα.

### 2.1.5 Συσσωρευτικές ιεραρχικές μέθοδοι ομαδοποίησης

Όπως αναφέρθηκε και προηγουμένως, κατά τη συσσωρευτική ιεραρχική μέθοδο συσταδοποίησης το πρόβλημα ξεκινά με πλήθος ομάδων (clusters) ίσο με τα αντικείμενα (subjects). Έτσι τελικό αποτέλεσμα αποτελεί η ομαδοποίηση των όλων αντικειμένων σε μία ομάδα. Η διαδικασία που ακολουθείται για να πραγματοποιηθεί η παραπάνω μέθοδος απεικονίζεται στην Εικόνα 2.2.



Εικόνα 2.2 : Διάγραμμα ροής συσσωρευτικού αλγορίθμου συσταδοποίησης.

Όσον αφορά τη παρούσα εργασία, αρχικά κάθε ασθενής θεωρήθηκε σαν μία ομάδα (cluster) δηλαδή ξεκινήσαμε με πλήθος ομάδων ίσο με το πλήθος των ασθενών. Στη συνέχεια έγινε κανονικοποίηση (z-scoring) στα δεδομένα εισόδου για να τα φέρουμε σε συγκρίσιμη κλίμακα.

Για τον υπολογισμό του πίνακα αποστάσεως των δεικτών των εξετάσεων (δεδομένα που περιέχονται στα αντικείμενα) των ασθενών (αντικείμενα), έγινε χρήση της μεθόδου συσχέτισης (correlation method). Στη συνέχεια έγινε η ένωση του ζευγαριού (των δύο ασθενών) με τη μεγαλύτερη συσχέτιση δημιουργώντας έτσι μια ομάδα η οποία περιείχε τους δυο αυτούς ασθενείς και το πλήθος των ομάδων μειώθηκε κατά ένα. Τέλος υπολογίστηκαν οι αποστάσεις μεταξύ των συστάδων με την μέθοδο της συσταδοποίησης μέσου συνδέσμου (average linkage),

η οποία θα αναλυθεί στη συνέχεια. Η διαδικασία επαναλήφθηκε μέχρι να γίνει η ομαδοποίηση όλων των ασθενών σε μια ομάδα (cluster) και παράχθηκε το αντίστοιχο δενδρόγραμμα.

Για τον υπολογισμό της απόστασης της νέας ομάδας που δημιουργείται κάθε φορά με τις υπόλοιπες ομάδες, χρησιμοποιούνται οι παρακάτω μέθοδοι.

➤ **Μέθοδος απλής σύνδεσης (Single Linkage Clustering)**

Στην περίπτωση αυτή, η απόσταση μεταξύ δυο συστάδων ισούται με την ελαχίστη απόσταση μεταξύ ενός αντικειμένου που ανήκει στην πρώτη ομάδα και ενός άλλου αντικειμένου της δεύτερης ομάδας. Βασικό μειονέκτημα της μεθόδου αποτελεί το γεγονός ότι μπορεί να συνδέσει απομονωμένα σημεία με ήδη υπάρχουσες ομάδες χωρίς να δημιουργήσει νέες έτσι. Το φαινόμενο αυτό αναφέρεται στη βιβλιογραφία ως φαινόμενο της αλυσίδας (chaining effect) και έχει σαν αποτέλεσμα τη λανθασμένη διαμόρφωση των ομάδων. Τέλος, η μέθοδος αυτή είναι γνωστή και ως μέθοδος του πλησιέστερου γείτονα (nearest neighbor method) [61]. Η μέθοδος αυτή περιγράφεται από την παρακάτω σχέση :

$$D(X, Y) = \min (d(x, y)) \quad (2.8)$$

Όπου  $D(X, Y)$  είναι η απόσταση μεταξύ των ομάδων (cluster)  $X$  και  $Y$  και  $d(x, y)$  η απόσταση μεταξύ των στοιχείων  $x, y$  των ομάδων  $X$  και  $Y$ .

➤ **Μέθοδος της πλήρους σύνδεσης (Complete Linkage Method)**

Στην περίπτωση αυτή, η απόσταση μεταξύ δυο συστάδων ισούται με τη μέγιστη απόσταση μεταξύ δύο αντικειμένων, το πρώτο εκ των οποίων ανήκει στην πρώτη ομάδα (cluster) και το δεύτερο στη δεύτερη ομάδα (cluster) [43]. Η μέθοδος αυτή δημιουργεί συμπαγείς ομάδες-συστάδες όμως χαρακτηρίζεται από την ευαισθησία της στις ακραίες τιμές, γεγονός που αποτελεί και το βασικό μειονέκτημα της μεθόδου. Η μέθοδος περιγράφεται από την παρακάτω σχέση.

$$D(X, Y) = \max (d(x, y)) \quad (2.9)$$

Όπου  $D(X, Y)$  είναι η απόσταση μεταξύ των ομάδων (cluster)  $X$  και  $Y$  και  $d(x, y)$  η απόσταση μεταξύ των στοιχείων  $x, y$  των ομάδων  $X$  και  $Y$ .

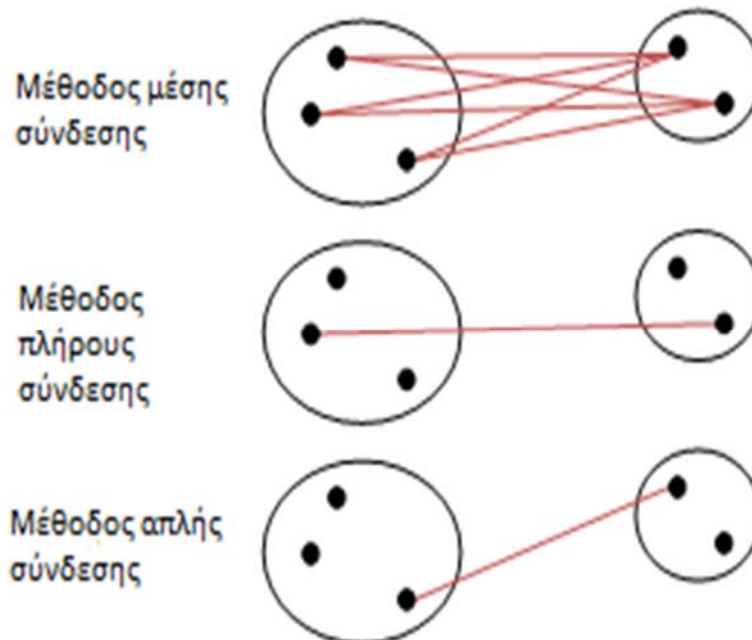
➤ **Μέθοδος της μέσης σύνδεσης (Average Linkage Method)**

Η απόσταση δύο ομάδων σε αυτή την περίπτωση είναι ίση με τη μέση τιμή της απόστασης κάθε αντικείμενου της μίας ομάδας με όλα τα αντικείμενα της άλλης ομάδας [61]. Η συγκεκριμένη μέθοδος χρησιμοποιήθηκε στην παρούσα εργασία για τον υπολογισμό της απόστασης των ομάδων.

$$D(X, Y) = \frac{1}{|X| * |Y|} \sum_{x \in X} \sum_{y \in Y} d(x, y) \quad (2.10)$$

Όπου  $D(X, Y)$  είναι η απόσταση μεταξύ των ομάδων (cluster)  $X$  και  $Y$  και  $d(x, y)$  η απόσταση μεταξύ των στοιχείων  $x, y$  των ομάδων  $X$  και  $Y$ .

Οι τρεις παραπάνω μέθοδοι αναπαρίστανται στην Εικόνα 2.3.



Εικόνα 2.3 : Απόσταση ομάδων βάσει των τριών μεθόδων.

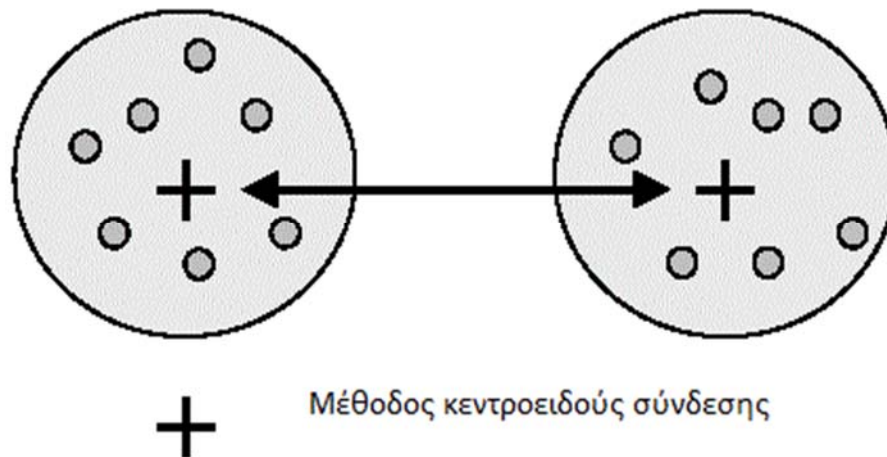


➤ **Μέθοδος της κεντροειδούς σύνδεσης (Centroid Linkage Method)**

Σε αυτή την περίπτωση αρχικά υπολογίζεται η μέση τιμή όλων των αντικειμένων σε κάθε συστάδα δηλαδή το κεντροειδές και στη συνέχεια υπολογίζονται οι αποστάσεις αυτών των κεντροειδών [53]. Το αποτέλεσμα της μεθόδου είναι συνήθως συμπαγείς (Εικόνα 2.4).

$$D(X, Y) = \|\bar{x}_X - \bar{x}_Y\| \quad (2.11)$$

Όπου  $\bar{x}_X$  είναι το κεντροειδές σημείο στο χώρο της X ομάδας (cluster) και  $\bar{x}_Y$  είναι αντίστοιχα το κεντροειδές σημείο στο χώρο της Y ομάδας.



Εικόνα 2.4 : Μέθοδος κεντροειδούς σύνδεσης [73].

➤ **Μέθοδος του Ward (Ward's Method)**

Η μέθοδος του Ward σχηματίζει συστάδες (clusters) μεγιστοποιώντας την ομοιογένεια στο εσωτερικό των συστάδων. Σαν μέτρο ομοιογένειας θεωρείται το άθροισμα τετραγώνων μέσα στην ομάδα. Η μέθοδος αυτή προσπαθεί να ελαχιστοποιήσει το συνολικό άθροισμα των τετραγώνων μέσα στην ομάδα. Οι ομάδες που σχηματίζονται έχουν όσο το δυνατόν μικρότερα

αθροίσματα τετραγώνων άρα κατασκευάζονται συμπαγείς ομάδες. Ουσιαστικά η μέθοδος του Ward ελαχιστοποιεί τις διακυμάνσεις εντός των ομάδων (clusters) [74].

Στην επόμενη παράγραφο του παρόντος κεφαλαίου θα αναλυθούν τα δεδομένα που χρησιμοποιήθηκαν καθώς και η μεθοδολογία που ακολουθήθηκε προκειμένου να φτάσουμε στο επιθυμητό αποτέλεσμα δηλαδή στην ομαδοποίηση των ασθενών με βάση χαρακτηριστικούς δείκτες από ιατρικές εξετάσεις.

## 2.2 Σχεδιασμός της μελέτης

### 2.2.1 Περιγραφή των δεδομένων

Το δεδομένα της μελέτης αντιστοιχούν σε 243 ασθενείς με καρωτιδική αθηρωμάτωση και συλλέχθηκαν στην Αγγειοχειρουργική Κλινική του Πανεπιστημιακού Γενικού Νοσοκομείου «Αττικόν» που εδρεύει στην Αθήνα. Καταγράφηκαν σειρά κλινικών παραμέτρων, όπως η ηλικία, περιβαλλοντικοί παράγοντες στους οποίους ήταν εκτεθειμένοι οι ασθενείς (κάπνισμα και σωματική άσκηση), ο δείκτης μάζας σώματος, η αρτηριακή πίεση, και ο δείκτης λιπώδους μάζας (Πίνακας 2.1).

Επιπλέον ελήφθησαν δείγματα αίματος για σειρά βιοχημικών, συμβατικών (Πίνακας 2.2) και εξειδικευμένων (Πίνακας 2.3), αναλύσεων. Η αιμοληψία διενεργήθηκε το πρωί (8-10 π.μ.) μετά από 12ωρη νηστεία και πριν από τη λήψη οποιουδήποτε φαρμακευτικού σκευάσματος. Προκειμένου να πραγματοποιηθούν οι εξετάσεις, απαιτήθηκε ένα φιαλίδιο γενικής αίματος με τουλάχιστον 2ml και αντιπηκτικό EDTA, ένα φιαλίδιο με παράγοντα πήξης ώστε να απομονωθεί ο ορός καθώς και ένα φιαλίδιο με αντιπηκτικό κιτρικό για την απομόνωση του πλάσματος. Η ανάγκη των επιπλέον φιαλιδίων προέκυψε από τις αρχικές μετρήσεις του ινωδογόνου, οι τιμές του οποίου ήταν πολύ χαμηλές στον ορό και επιλέχθηκε η χρήση πλάσματος απομονωμένου από αίμα με κιτρικό ως αντιπηκτικό. Η μεταφορά του υλικού στο εργαστήριο πραγματοποιήθηκε εντός τριών ωρών από τη λήψη σε θερμοκρασία περιβάλλοντος. Από τα φιαλίδια αίματος συλλέχθηκαν μετά από φυγοκέντρηση ορός, πλάσμα καθώς και αυτούσιο αίμα με EDTA.

Τέλος έγινε υπερηχογραφικός έλεγχος και των δύο καρωτίδων. Καταγράφηκαν ακολουθίες εικόνων υπερήχων β-σάρωσης της καρωτίδας σε διαμήκη τομή. Όλες οι απεικονιστικές καταγραφές έγιναν ακολουθώντας ένα τυποποιημένο πρωτόκολλο, ικανό να διασφαλίσει σχετική ομοιογένεια. Συγκεκριμένα, η εξέταση κάθε ατόμου πραγματοποιήθηκε σε ύπτια θέση, με μια ελαφριά κλίση του κεφαλιού προς τα πίσω και προς την αντίθετη πλευρά της υπό εξέταση καρωτίδας. Πριν ξεκινήσει η διαδικασία του απεικονιστικού ελέγχου, υπήρξε ένα διάστημα ανάπαυσης του ασθενούς διάρκειας τουλάχιστον 5 λεπτών, ώστε να έχουν σταθεροποιηθεί ο καρδιακός ρυθμός και η πίεση του αίματός του. Το κέρδος, που αποτελεί μια υποκειμενική παράμετρο στις ρυθμίσεις της διάταξης, ρυθμίστηκε έτσι ώστε το αίμα να είναι σκούρο και με ομοιόμορφη ηχωγένεια και ο έξω χιτώνας να εμφανίζεται λεπτός, με υψηλή φωτεινότητα και, επίσης, με ομοιόμορφη ηχωγένεια. Η αποθήκευση των απεικονιστικών δεδομένων έγινε σε πρότυπο DICOM (Digital Imaging and Communications in Medicine), το οποίο χρησιμοποιείται ευρέως για την αποθήκευση, διαχείριση και μεταφορά ιατρικών απεικονιστικών δεδομένων, κυρίως επειδή επιτρέπει τη διασύνδεση, συμβατότητα και βελτιστοποίηση της ροής πληροφορίας σε πληροφοριακά συστήματα υγείας. Οι απεικονιστικές καταγραφές χρησιμοποιήθηκαν για τον υπολογισμό των δεικτών του (Πίνακας 2.4) ακολουθώντας τις μεθοδολογίες που περιγράφονται στις εργασίες του Εργαστήριο Βιοϊατρικών Προσομοιώσεων και Απεικονιστικής Τεχνολογίας (BIOSIM) [22] [23] [75].

Τα αγγεία χωρίστηκαν σε δύο κατηγορίες («συμπτωματικά» και «ασυμπτωματικά») με βάση την παρουσία ή μη, αντίστοιχα, τουλάχιστον το τελευταίο εξάμηνο, αγγειακών εγκεφαλικών επεισοδίων που οφείλονται σε αθηρωματικές αλλοιώσεις του αγγείου.

Πίνακας 2.1 : Συνοπτική παρουσίαση του προφίλ των ασθενών.

Ημερομηνία Ελέγχου		24/4/2012 - 4/3/2015	
Χαρακτηριστικά	Πλήθος ή μέση τιμή	%	
Πλήθος ασθενών	243		
Κάπνισμα	Υψηλό: 132	54.30	
	Μεσαίο: 62	26.70	
	Χαμηλό: 49	20.16	

<b>Σωματική Άσκηση</b>	Υψηλή :20	8.23
	Μεσαία: 74	30.45
	Χαμηλή: 128	52.67
	Δεν γνωρίζουμε : 21	8.64
<b>Δείκτης μάζας σώματος</b>	27	
<b>Συστολική πίεση</b>	137	
<b>Διαστολική πίεση</b>	78	
<b>Δείκτης λιπώδους μάζας</b>	31	

Πίνακας 2.2 : Πλήρης ονομασία και συντομογραφία συμβατικών βιοχημικών δεικτών.

<b>Συμβατικοί βιοχημικοί δείκτες</b>	
<b>Πλήρης ονομασία</b>	<b>Συντομογραφία</b>
Ερυθρά Αιμοσφαίρια (Μ/μl)	RBC
Αιμοσφαιρίνη (g/dL)	Hb
Αιματοκρίτης (%)	Hct
Μέσος Όγκος Ερυθρών (fl)	MCV
Μέση Περιεκτικότητα αιμοσφαιρίνης (pg)	MCH
Μέση Συγκέντρωση αιμοσφαιρίνης (g/dL)	MCHC
Εύρος Κατανομής Ερυθρών (%)	RDW-CV
Εύρος κατανομής μεγέθους ερυθρών αιμοσφαιρίων (g/dL)	RDW-SV
Εμπύρηνα ερυθρά # (K/μl)	NRBC
Εμπύρηνα ερυθρά % (/100WBC)	NRBC/100 WBC
Αριθμός Λευκοκυττάρων (K/μl)	WBC
Ουδετερόφιλα (%)	NEUT %
Λεμφοκύτταρα (%)	LYMPH %
Μονοκύτταρα (%)	MONO %
Ηωσινόφιλα (%)	EO %
Βασεόφιλα (%)	BASO %
Αριθμός Ουδετεροφίλων (K/μl)	Neutrophils
Αριθμός Λεμφοκυττάρων (K/μl)	Lymphocytes
Αριθμός Μονοκυττάρων (K/μl)	Monocytes
Αριθμός Ηωσινοφίλων (K/μl)	Eosinophils
Αριθμός Βασεόφιλων (K/μl)	Basophils
Αιμοπετάλια (K/μl)	PLT
Μέσος όγκος PLT (fl)	MPV
Εύρος κατανομής PLT (fl)	PDW
Αιμοπεταλιοκρίτης (%)	PCT

Ποσοστό νεαρών PLT (%)	PLT
Ουρία (mg/dL)	UN
Κρεατινίνη (mg/dL)	Cr
Οξαλοξική Τρανσαμινάση (U/L)	SGOT
Πυροσταφυλική Τρανσαμινάση (U/L)	SGPT
γ-Γλουταμυλοτρανσφεράση (U/L)	γ-GT
Αλκαλική Φωσφατάση (U/L)	ALP
Ολική Χοληστερόλη (mg/dL)	CHO
Τριγλυκερίδια (mg/dL)	TGL
HDL Χοληστερόλη- Υψηλής ποσότητας λιποπρωτεϊνη (mg/dL)	HDL
LDL Χοληστερόλη- Χαμηλής ποσότητας λιποπρωτεϊνη (mg/dL)	LDL
Γλυκόζη (mg/dL)	GLU

Πίνακας 2.3 : Πλήρης ονομασία και συντομογραφία εξειδικευμένων βιοχημικών δεικτών.

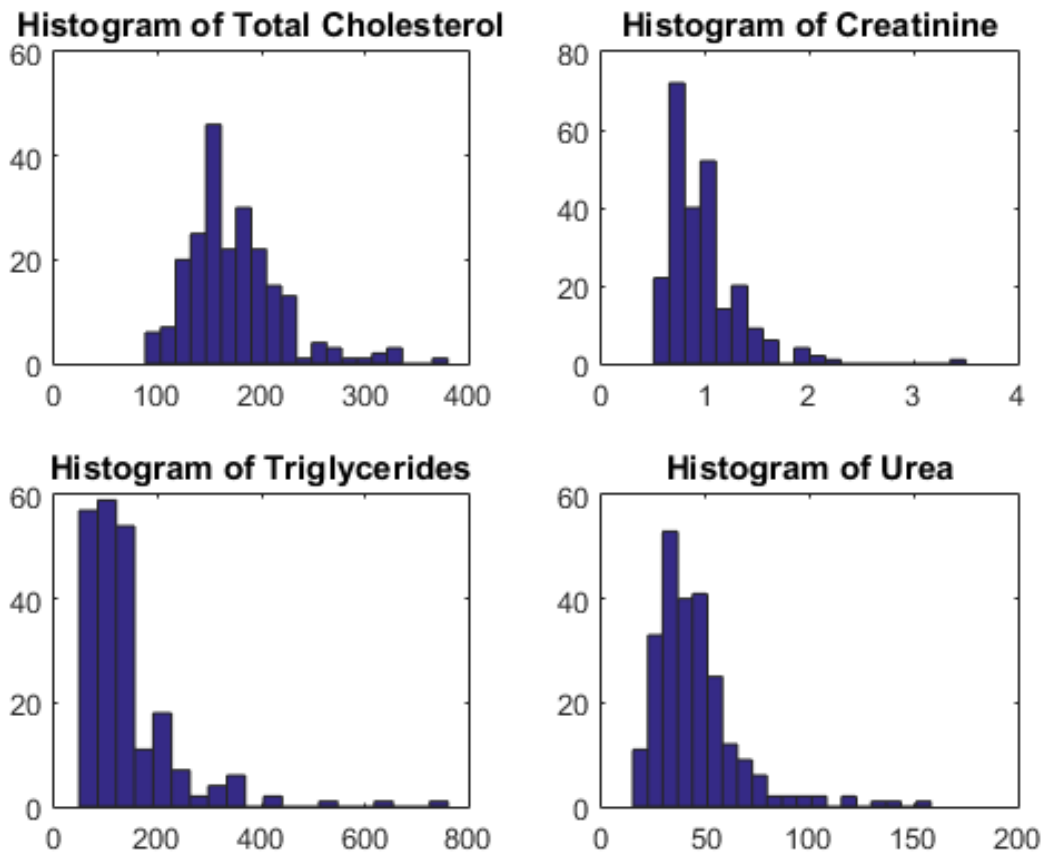
**Εξειδικευμένοι βιοχημικοί δείκτες**

<b>Πλήρης ονομασία</b>	<b>Συντομογραφία</b>
C-αντιδρώσα πρωτεΐνη (ng/ml)	CRP
Fibrinogen (pg/ml)	Fibrinogen
Μεταλλοπρωτεΐναση-1 (pg/ml)	mmp 1
Μεταλλοπρωτεΐναση-2 (pg/ml)	mmp 2
Μεταλλοπρωτεΐναση-7 (pg/ml)	mmp 7
Μεταλλοπρωτεΐναση-9 (pg/ml)	mmp 9
Ιντερλευκίνη 1-β (pg/ml)	IL-1β
Ιντερλευκίνη-6 (pg/ml)	IL-6
Παράγοντας νέκρωσης όγκου-α (pg/ml)	TNFα
Πεταλλοπρωτεΐναση-1 (pg/ml)	TIMP-1
Πεταλλοπρωτεΐναση-2 (pg/ml)	TIMP-2
Πεπτίδιο-C (pg/ml)	C-peptide
Ινσουλίνη (pg/ml)	Insulin
Retinol-Binding Protein (pg/ml)	RBP4
Γαλεκτίνη-3 (ng/ml)	Galectin3

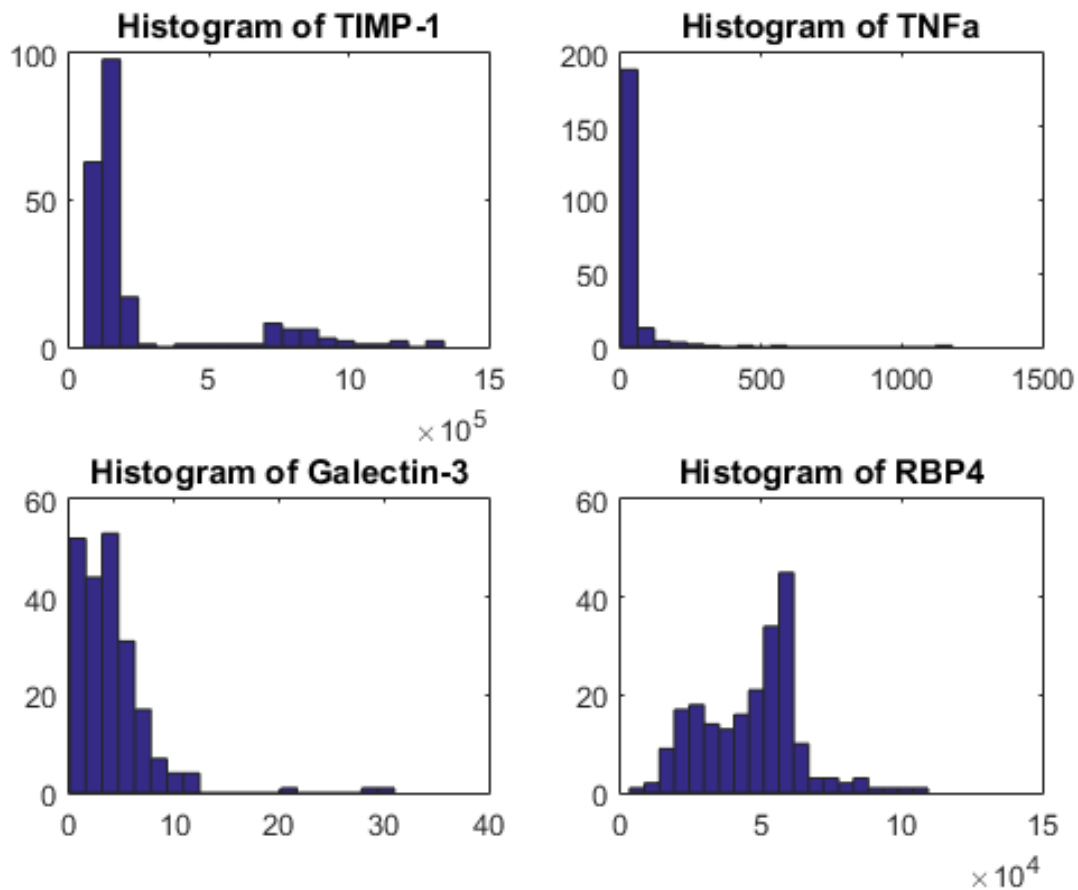
Πίνακας 2.4 : Οι απεικονιστικοί δείκτες

<b>Απεικονιστικοί δείκτες</b>	
<b>Πλήρης ονομασία</b>	<b>Συντομογραφία</b>
Βαθμός στένωσης (%)	ΒΣ
Εμβαδό της πλάκας	Area
Gray-Scale Median	GSM

Στη συνέχεια παρουσιάζονται ενδεικτικά ιστογράμματα που αντιστοιχούν στην κατανομή συμβατικών και εξειδικευμένων βιοχημικών δεικτών στο δείγμα μας.



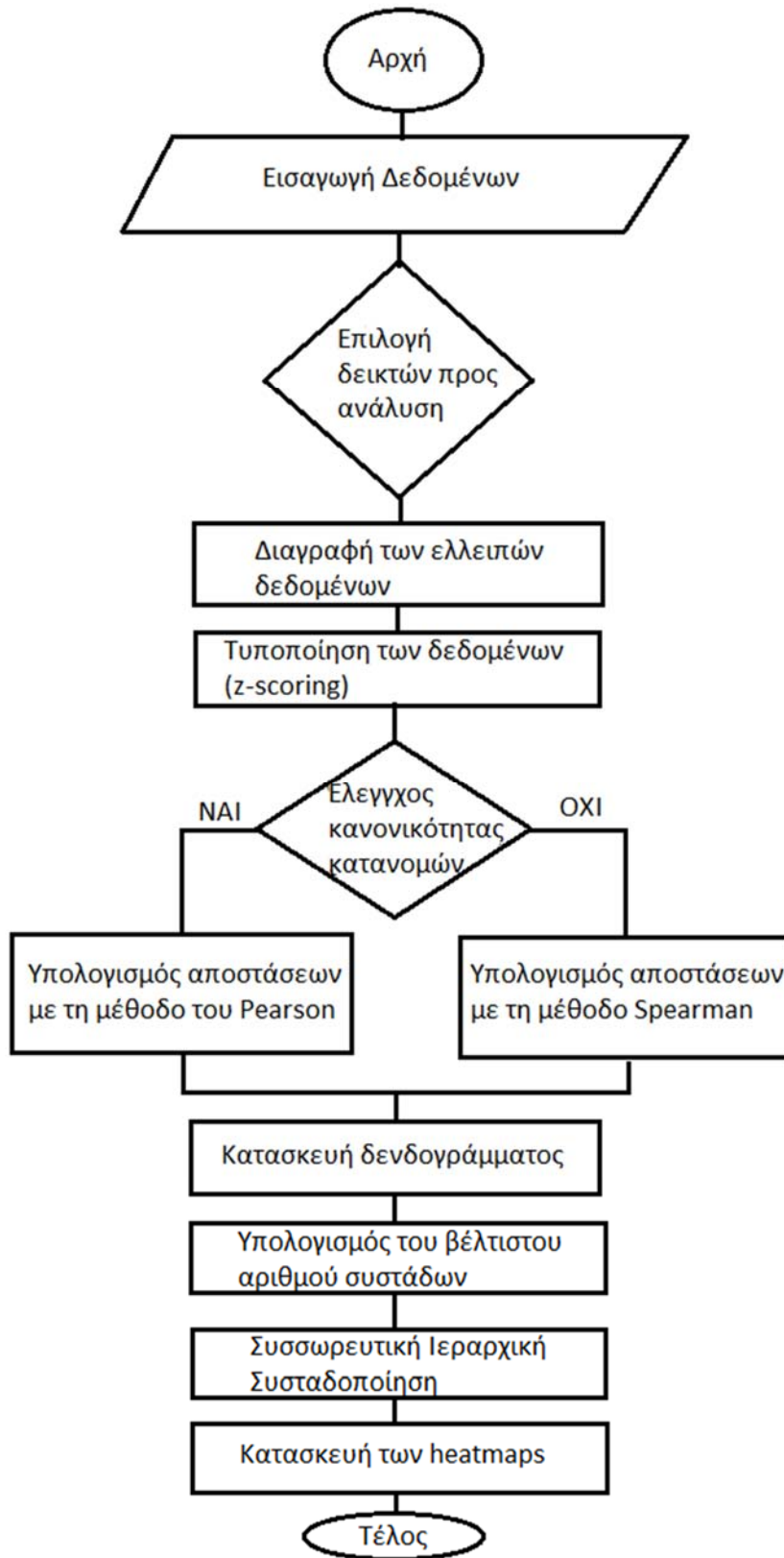
Εικόνα 2.5 : Ιστογράμματα τεσσάρων συμβατικών βιοχημικών δεικτών.



Εικόνα 2.6 : Ιστογράμματα τεσσάρων εξειδικευμένων βιοχημικών δεικτών.

### 2.2.2 Περιγραφή της μεθόδου

Η παρούσα εργασία εμπνεύστηκε από προηγούμενη εργασία η οποία μελέτησε φαινοτύπους του καρκίνου του μαστού, συσχετίζοντας τους με γονιδιακά προφίλ με χρήση τεχνικών συσταδοποίησης [47]. Στα πλαίσια της παρούσας διπλωματικής εργασίας, πραγματοποιήθηκε συσσωρευτική ιεραρχική συσταδοποίηση (Agglomerative Hierarchical Clustering) σε περιβάλλον Matlab με σκοπό την ομαδοποίηση ενός πλήθους ασθενών σε ομάδες βάσει των χαρακτηριστικών τους δηλαδή τις αντίστοιχες τιμές των εξετάσεων κάθε ασθενή.



Εικόνα 2.7 : Διάγραμμα ροής της μεθοδολογίας.



Αρχικά, φορτώθηκαν τα δεδομένα σε μορφή πίνακα που αφορούσαν τις πρώτες εξετάσεις συμβατικών και εξειδικευμένων βιοχημικών δεικτών κάθε ασθενή. Αμέσως μετά υπήρχε η δυνατότητα επιλογής από τον χρήστη τριών επιλογών.

- Πρώτη επιλογή : πραγματοποίηση συσταδοποίησης βάσει μόνο των εξετάσεων των εξειδικευμένων βιοχημικών δεικτών.
- Δεύτερη επιλογή : πραγματοποίηση συσταδοποίησης βάσει μόνο των εξετάσεων των συμβατικών βιοχημικών δεικτών.
- Τρίτη επιλογή : πραγματοποίηση συσταδοποίησης βάσει των εξετάσεων των συμβατικών και εξειδικευμένων βιοχημικών δεικτών.

Στη συνέχεια αφού επιλέχθηκε η επιθυμητή επιλογή, διαγράφηκαν οι γραμμές του πίνακα (ασθενείς) που περιείχαν κενούς δείκτες (missing data) διότι η μέθοδος που αναπτύχθηκε στα πλαίσια της παρούσας εργασίας χρησιμοποιεί μεθόδους συσχέτισης (correlation methods) για τον υπολογισμό των αποστάσεων των συστάδων όπως περιγράφεται και στο άρθρο που αποτέλεσε έμπνευση για την παρούσα εργασία [39]. Η μέθοδος αυτή δεν μπορεί να εφαρμοστεί επιτυχώς όταν υπάρχουν ελλιπή δεδομένα (missing data).

Ο νέος πίνακας δεδομένων χωρίς τα ελλιπή δεδομένα αποτέλεσε την είσοδο της διαδικασίας. Αρχικά πραγματοποιήθηκε τυποποίηση των τιμών του νέου πίνακα ώστε να έχουν τυπική απόκλιση ένα γύρω από το μηδέν (z-scoring) προκειμένου όλα τα δεδομένα να έρθουν σε συγκρίσιμη μορφή. Στη συνέχεια, πραγματοποιήθηκε ένα τεστ κανονικότητας (Kolmogorov-Smirnov normality test) ώστε να ελεγχθεί αν τα δεδομένα μετά την κανονικοποίηση (z-scoring) ακολουθούν την κανονική κατανομή (normal data). Αν τα δεδομένα ακολουθούν κανονική κατανομή το τεστ δίνει αποτέλεσμα ένα διαφορετικά δίνει αποτέλεσμα μηδέν.

Στη περίπτωση όπου τα δεδομένα ακολουθούσαν κανονική κατανομή τότε οι αποστάσεις των δεδομένων εισόδου υπολογίστηκαν με τη μέθοδο συσχέτισης του Pearson, η οποία αναλύθηκε παραπάνω. Αντίθετα, εάν τα δεδομένα δεν ακολουθούσαν κανονική κατανομή τότε οι αποστάσεις των δεδομένων εισόδου υπολογίστηκαν με τη μέθοδο συσχέτισης του Spearman, η οποία χρησιμοποιείται για μη γραμμικές συσχετίσεις δηλαδή για μη κανονικά δεδομένα. Στη συνέχεια αφού υπολογίστηκαν οι αποστάσεις των χαρακτηριστικών των

ασθενών, με χρήση των δυο παραπάνω μεθόδων αναλόγως την περίπτωση, δημιουργήθηκε το αντίστοιχο δενδρόγραμμα.

Επόμενο βήμα αποτέλεσε η πραγματοποίηση συσταδοποίησης επιλέγοντας σαν αριθμό των επιθυμητών ομάδων (number of clusters), το πλήθος των ασθενών, θεωρώντας στην αρχή ότι κάθε ασθενής αποτελεί μια ξεχωριστή ομάδα (cluster). Η διαδικασία αυτή έγινε με επαναληπτικό τρόπο ξεκινώντας με τη θεώρηση ότι θέλουμε όλοι οι ασθενείς να ομαδοποιηθούν στην ίδια ομάδα άρα μετά το πρώτο επαναληπτικό βήμα είχαμε σαν αποτέλεσμα όλοι οι ασθενείς να έχουν ομαδοποιηθεί στην ίδια ομάδα. Στο δεύτερο βήμα θεωρήσαμε ότι θέλουμε όλοι οι ασθενείς να ομαδοποιηθούν σε δύο ομάδες κτλ.

Στόχος αυτής της επαναληπτικής διαδικασίας αποτέλεσε ο υπολογισμός αρχικά και ο σχεδιασμός έπειτα της μέσης απόστασης των στοιχείων κάθε ομάδας (average pairwise distance within the cluster) (Σχέση 2.12) συναρτήσεως του πλήθους των ομάδων (clusters). Αφού παράχθηκε η γραφική παράσταση της μέσης απόστασης των στοιχείων των ομάδων συναρτήσεως του πλήθους των ομάδων, επιλέχθηκε ο βέλτιστος αριθμός ομάδων (optimal number of clusters). Ο βέλτιστος αυτός αριθμός αντιστοιχούσε στο σημείο όπου παρατηρήθηκε η μέγιστη πτώση (μεταβολή) της γραφικής παράστασης όπως περιγράφεται στο άρθρο το οποίο αποτέλεσε έμπνευση της παρούσας εργασίας [47].

$$D_{avg}(c) = \frac{1}{c} \sum_{l=1}^c \sqrt{\frac{\sum_{u=1}^{N_l} \sum_{v=1}^{N_l} \|x_{ul} - x_{vl}\|^2}{N_l(N_l - 1)}} \quad (2.12)$$

Όπου  $c$  είναι ο αριθμός των συστάδων,  $N_l$  το πλήθος μελών της  $l$  συστάδας,  $x_{ul}$  το  $u$ -οστό χαρακτηριστικό της  $l$  συστάδας και  $x_{vl}$  το  $v$ -οστό χαρακτηριστικό της  $l$  συστάδας.

Τέλος κατασκευάζονται τα heatmaps, τα οποία αντιστοιχούν σε χρωματική κωδικοποίηση των τιμών των δεικτών ανά φαινότυπο. Η ερμηνεία των heatmaps επιχειρείται με παράλληλη απεικόνιση των απεικονιστικών δεικτών και της συμπτωματολογίας.

Σε αυτό το σημείο απαραίτητο είναι να γίνει αναφορά στις παραδοχές που έγιναν όσο αφορά τα δεδομένα μας. Αρχικά, για τα δεδομένα των πρώτων συμβατικών και εξειδικευμένων βιοχημικών δεικτών κάθε ασθενή, αν κάποιος δείκτης έχει πολλά ελλιπή δεδομένα (missing

data) >50 %, αφαιρέθηκε. Για τα δεδομένα που αφορούσαν το εμβαδό της πλάκας, το βαθμός στένωσης, το δείκτης GSM κατά τη διαστολή και τη συστολή καθώς και τη συμπτωματολογία κάθε ασθενή, αν υπήρχαν τιμές και για τα δυο αγγεία του ίδιου ασθενή τότε η αντίστοιχη τιμή του μεγέθους θεωρήθηκε η μέση τιμή των τιμών των δυο αγγείων.

Όσο αφορά τη συμπτωματολογία δημιουργήσαμε δύο (2) κατηγορίες. Η 1<sup>η</sup> κατηγορία αφορούσε τους ασυμπτωματικούς ασθενείς και σε αυτήν εντάχθηκαν οι ασθενείς των οποίων και τα δυο αγγεία ήταν ασυμπτωματικά. Αν τουλάχιστον ένα αγγείο ενός ασθενή ήταν συμπτωματικό, τότε ο ασθενής κατατάχθηκε στην 2<sup>η</sup> κατηγορία δηλαδή στη συμπτωματική κατηγορία. Τέλος, αν υπήρχαν δεδομένα μόνο για το ένα αγγείο του ασθενή και αυτό το αγγείο ήταν συμπτωματικό τότε ο ασθενής κατατάχθηκε στην 2<sup>η</sup> κατηγορία (συμπτωματική κατηγορία) και αντίστοιχα αν ήταν ασυμπτωματικό τότε ο ασθενής κατατάχθηκε στην 1<sup>η</sup> κατηγορία (ασυμπτωματική κατηγορία).

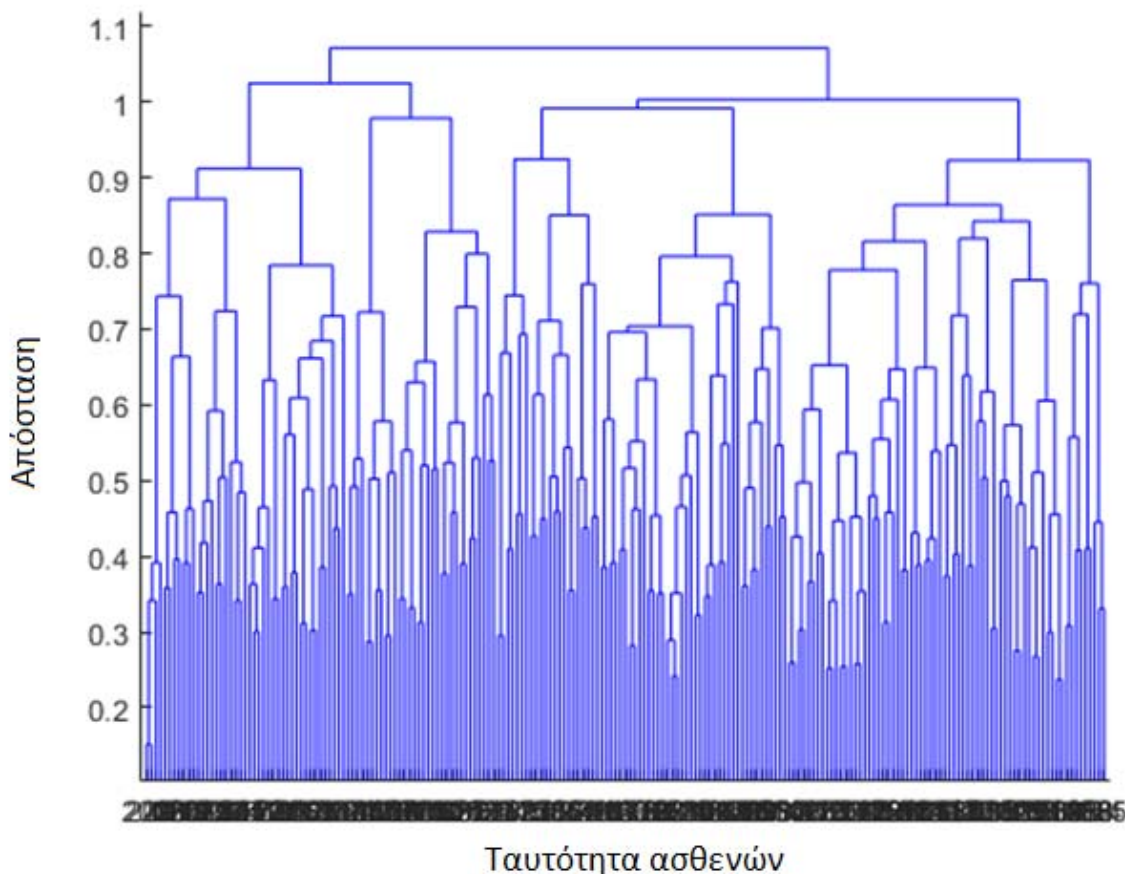
## Κεφάλαιο 3 : Αποτελέσματα

### Περίληψη

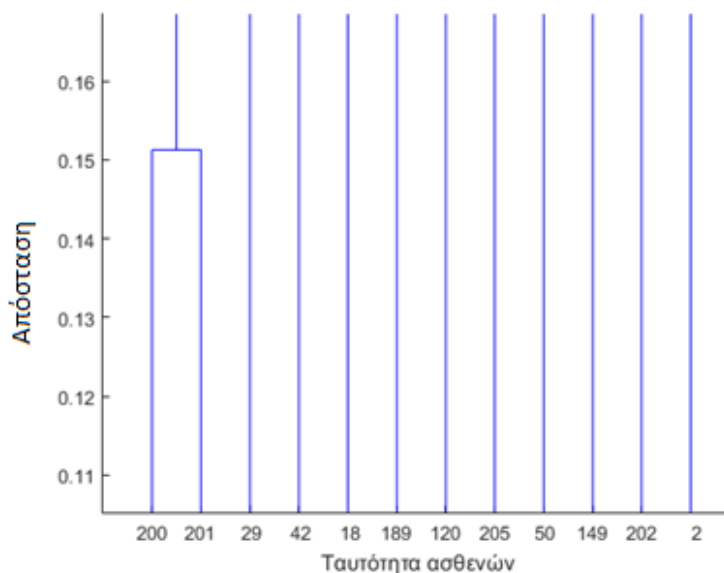
Στο παρόν κεφάλαιο παρουσιάζονται και ερμηνεύονται τα αποτελέσματα της μεθόδου που αναπτύχθηκε στα πλαίσια αυτής της διπλωματικής εργασίας. Όπως αναφέρθηκε ήδη στο προηγούμενο κεφάλαιο υπήρχε η δυνατότητα επιλογής μεταξύ τριών περιπτώσεων, δηλαδή πραγματοποίηση συσταδοποίησης με βάση τις εξετάσεις των εξειδικευμένων βιοχημικών δεικτών ή των συμβατικών βιοχημικών δεικτών ή συνδυασμό και των δυο. Τα αποτελέσματα παρουσιάζονται ξεκινώντας από την πρώτη περίπτωση δηλαδή με βάση τις εξετάσεις εξειδικευμένων βιοχημικών δεικτών των ασθενών.

### 3.1 Χρήση συμβατικών βιοχημικών δεικτών

Βασιζόμενοι στους συμβατικούς βιοχημικούς δείκτες, κατασκευάστηκε το δενδρογράμμα στο οποίο κάθε φύλλο αναπαριστά αρχικά κάθε ασθενή και κάθε κόμβος αναπαριστά τη συγχώνευση των ασθενών και των ομάδων (Εικόνα 3.1). Επιπλέον λόγω πολλαπλής επικάλυψης των ταυτοτήτων των ασθενών στον οριζόντιο άξονα, παρουσιάζεται ένα μεγεθυμένο τμήμα του δενδρογράμματος, στο οποίο ο διαχωρισμός των ταυτοτήτων αυτών είναι εμφανής.

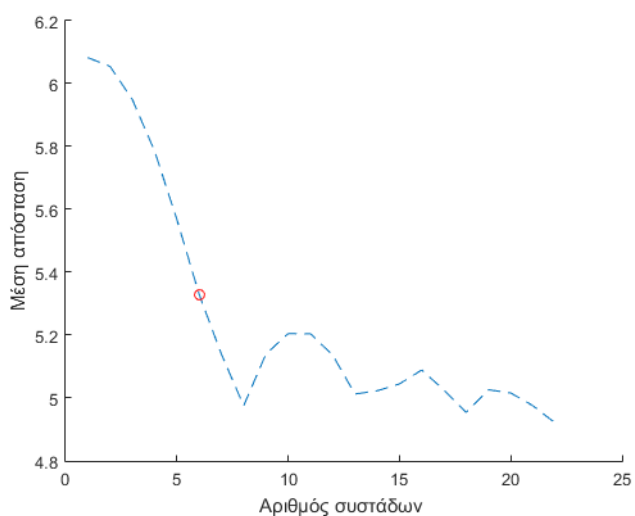


Εικόνα 3.1 : Γραφική αναπαράσταση του δενδρογράμματος των συμβατικών βιοχημικών δεικτών.



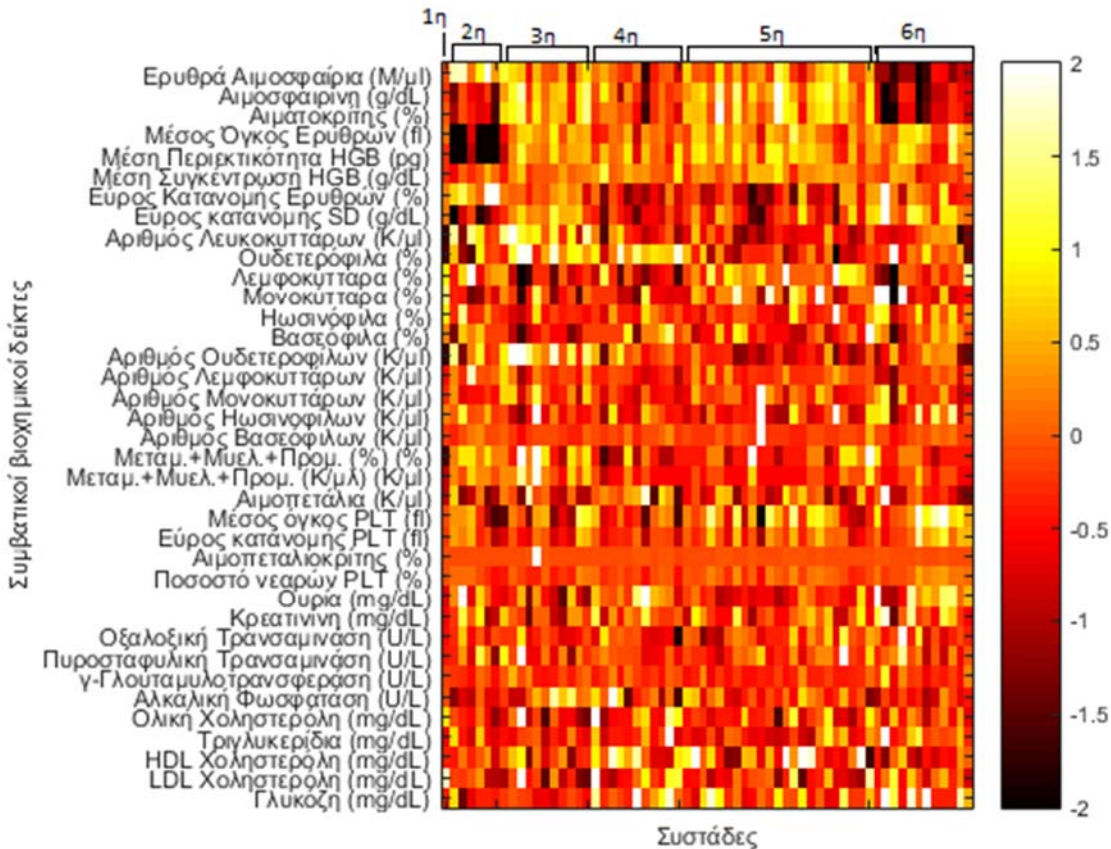
Εικόνα 3.2 : Γραφική αναπαράσταση μεγεθυμένου τμήματος του δένδρογράμματος των συμβατικών βιοχημικών δεικτών.

Στη συγκεκριμένη περίπτωση των συμβατικών βιοχημικών δεικτών, ο βέλτιστος αριθμός των συστάδων είναι ο αριθμός έξι. Ο αριθμός αυτός, επιλέχθηκε υπολογίζοντας το σημείο της γραφικής παράστασης στο οποίο παρατηρείται μέγιστη πτώση. Αυτό σημαίνει ότι η ομαδοποίηση των ασθενών βάσει των συμβατικών βιοχημικών δεικτών θα γίνει σε έξι φαινότυπους.



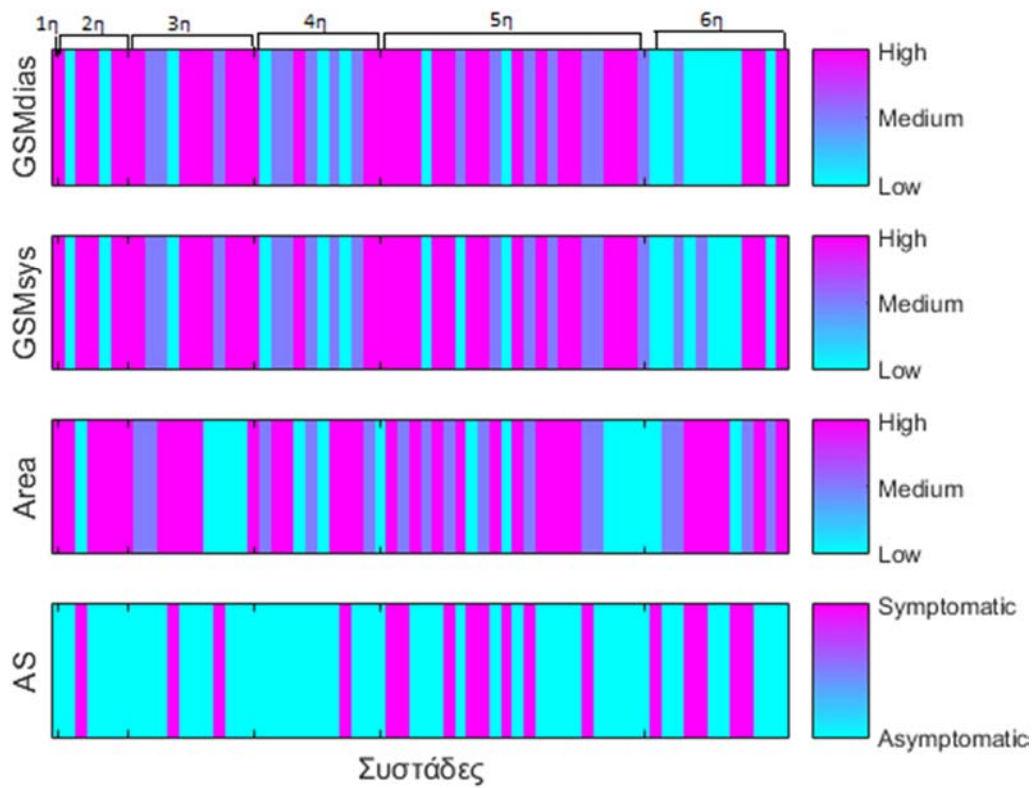
Εικόνα 3.3 : Γραφική παράσταση μέσης απόστασης των στοιχείων μεταξύ των ομάδων συναρτήσει του πλήθους των ομάδων για την εύρεση του βέλτιστου αριθμού συστάδων.

Στη συνέχεια χρησιμοποιήθηκαν τα αποτελέσματα της ιεραρχικής συσταδοποίησης για την κατασκευή ενός heatmap όπου οι στήλες αναπαριστούν τους ασθενείς ταξινομημένους ανά συστάδα και οι γραμμές τους συμβατικούς βιοχημικούς δείκτες των εξετάσεων. Το heatmap αποτελεί στην ουσία οπτικοποίηση των ιδιοτήτων των φαινοτύπων που προέκυψαν από την ιεραρχική συσταδοποίηση.

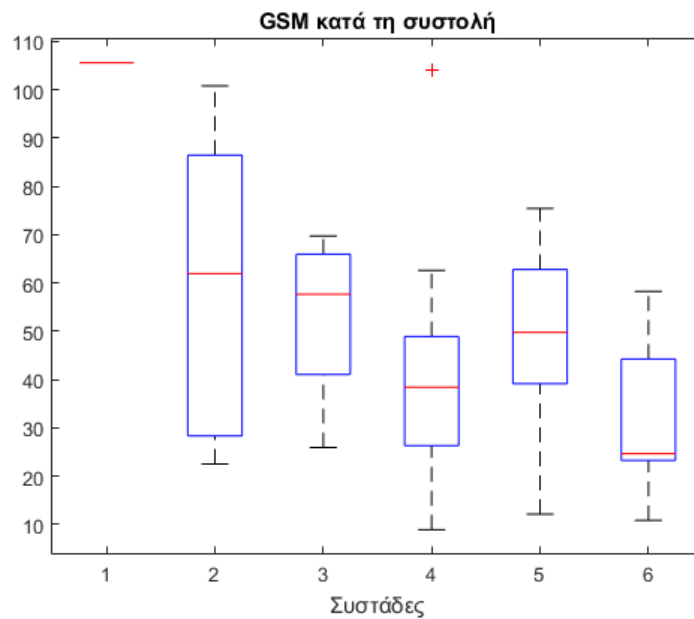


Εικόνα 3.4 : Το heatmap που αντιστοιχεί στην συσταδοποίηση βάσει των συμβατικών βιοχημικών δεικτών.

Επιπλέον, χρησιμοποιώντας δεδομένα για τα χαρακτηριστικά των ασθενών όπως η τιμή του δείκτη GSM, ο βαθμός στένωσης, το εμβαδό της πλάκας καθώς και η συμπτωματολογία των αγγείων κατασκευάστηκαν γραφήματα τα οποία οπτικοποιούν τη σχέση των δεικτών αυτών με τις ομάδες (clusters) των ασθενών.

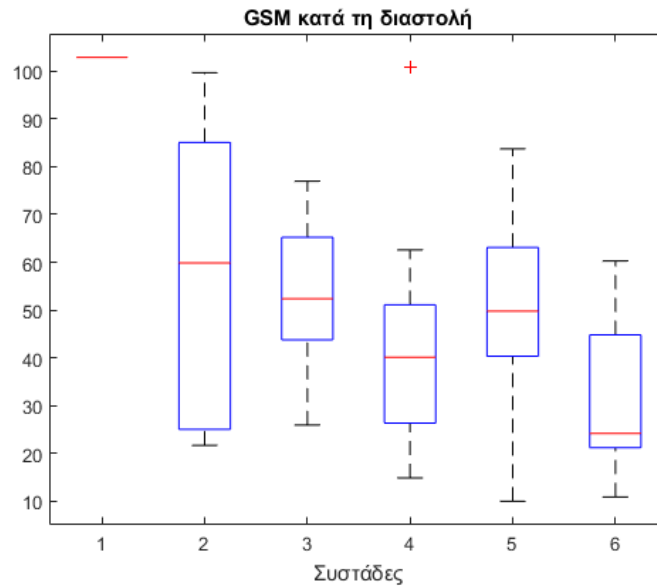


Εικόνα 3.5 : Απεικονιστικά χαρακτηριστικά των συστάδων.

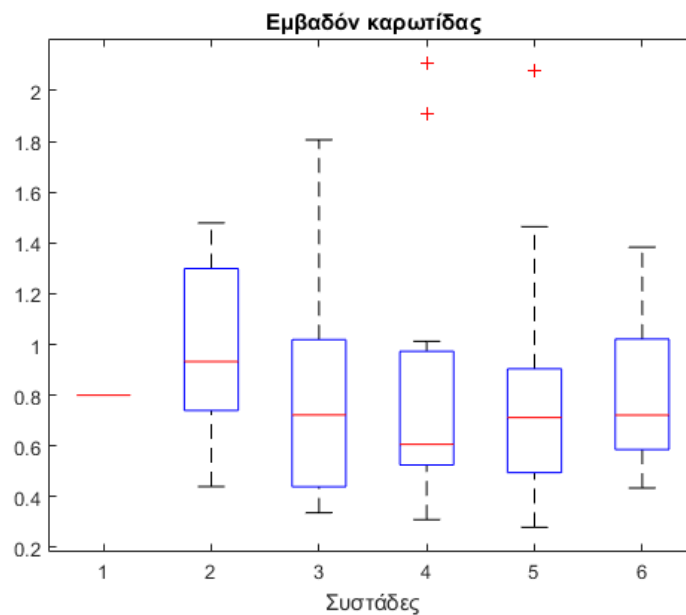


Εικόνα 3.6 : Βoxplot για το συστολικό GSM κατά τη συσταδοποίηση με βάση τους συμβατικούς βιοχημικούς δείκτες.





Εικόνα 3.7 : Βoxplot για το διαστολικό GSM κατά τη συσταδοποίηση με βάση τους συμβατικούς βιοχημικούς δείκτες.



Εικόνα 3.8 : Βoxplot για το εμβαδόν της καρωτίδας κατά τη συσταδοποίηση με βάση τους συμβατικούς βιοχημικούς δείκτες.

Στη συνέχεια οι συνεχείς τιμές των δεικτών GSM της διαστολής, της συστολής και του εμβαδού της καρωτίδας, ταξινομήθηκαν σε τρεις κατηγορίες (χαμηλό, μεσαίο, υψηλό). Ακολουθεί ο πίνακας των ποσοστών των ασθενών κάθε φαινοτύπου που εμπίπτει στις κατηγορίες υψηλή, μεσαία και χαμηλή.

Πίνακας 3.1 : Πίνακας των ποσοστών των ασθενών κάθε φαινοτύπου στις κατηγορίες υψηλή, μεσαία και χαμηλή.

	<b>GSM διαστολής</b>			
	Συνολικό πλήθος ασθενών ανά συστάδα	Κατηγορία  Χαμηλή (%)	Κατηγορία  Μεσαία (%)	Κατηγορία  Υψηλή (%)
<b>Συστάδα 1</b>	11	0	0	100
<b>Συστάδα 2</b>	21	33.3	0	66.67
<b>Συστάδα 3</b>	40	9.09	27.27	63.64
<b>Συστάδα 4</b>	22	27.27	45.46	27.27
<b>Συστάδα 5</b>	68	8.69	30.44	60.87
<b>Συστάδα 6</b>	43	66.6	8.33	25

Πίνακας 3.2 : Πίνακας των ποσοστών των ασθενών κάθε φαινοτύπου στις κατηγορίες υψηλή, μεσαία και χαμηλή.

	<b>GSM συστολής</b>			
	Συνολικό πλήθος ασθενών ανά συστάδα	Κατηγορία  Χαμηλή (%)	Κατηγορία  Μεσαία (%)	Κατηγορία  Υψηλή (%)
<b>Συστάδα 1</b>	11	0	0	100
<b>Συστάδα 2</b>	21	33.3	0	66.67
<b>Συστάδα 3</b>	40	9.09	27.27	63.64
<b>Συστάδα 4</b>	22	27.27	45.46	27.27
<b>Συστάδα 5</b>	68	13.04	26.08	60.87
<b>Συστάδα 6</b>	43	58.33	16.67	25

Πίνακας 3.3 : Πίνακας των ποσοστών των ασθενών κάθε φαινοτύπου στις κατηγορίες υψηλή, μεσαία και χαμηλή.

	<b>Εμβασό καρωτίδας</b>			
	Συνολικό πλήθος ασθενών ανά συστάδα	Κατηγορία Χαμηλή (%)	Κατηγορία Μεσαία (%)	Κατηγορία Υψηλή (%)
<b>Συστάδα 1</b>	11	0	0	100
<b>Συστάδα 2</b>	21	16.67	0	83.33
<b>Συστάδα 3</b>	40	36.36	18.18	45.46
<b>Συστάδα 4</b>	22	27.27	27.27	45.46
<b>Συστάδα 5</b>	68	26.09	30.44	43.48
<b>Συστάδα 6</b>	43	16.67	33.33	50

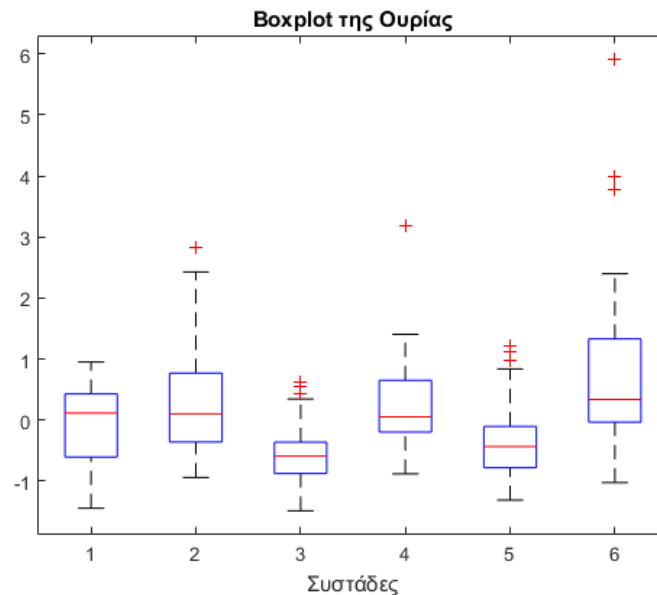
	<b>Ασυμπτωματικοί/Συμπτωματικοί</b>		
	Συνολικό πλήθος ασθενών ανά συστάδα	Κατηγορία Ασυμπτωματικοί (%)	Κατηγορία Συμπτωματικοί (%)
<b>Συστάδα 1</b>	11	100	0
<b>Συστάδα 2</b>	21	83.33	16.66
<b>Συστάδα 3</b>	40	81.81	18.18
<b>Συστάδα 4</b>	22	90.90	9.09
<b>Συστάδα 5</b>	68	65.21	34.78
<b>Συστάδα 6</b>	43	58.33	41.66

Προκειμένου να ερμηνευτεί το heatmap (Εικόνα 3.4), έγινε χρήση της μεθόδου Kruskal–Wallis, η οποία επιστρέφει ως αποτέλεσμα τους δείκτες οι οποίοι παρουσιάζουν την πιο μεγάλη διαφοροποίηση μεταξύ των φαινοτύπων για κανονικά δεδομένα. Τα αποτελέσματα της μεθόδου ήταν τα εξής με βάση τη σειρά υπερίσχυσης των δεικτών:

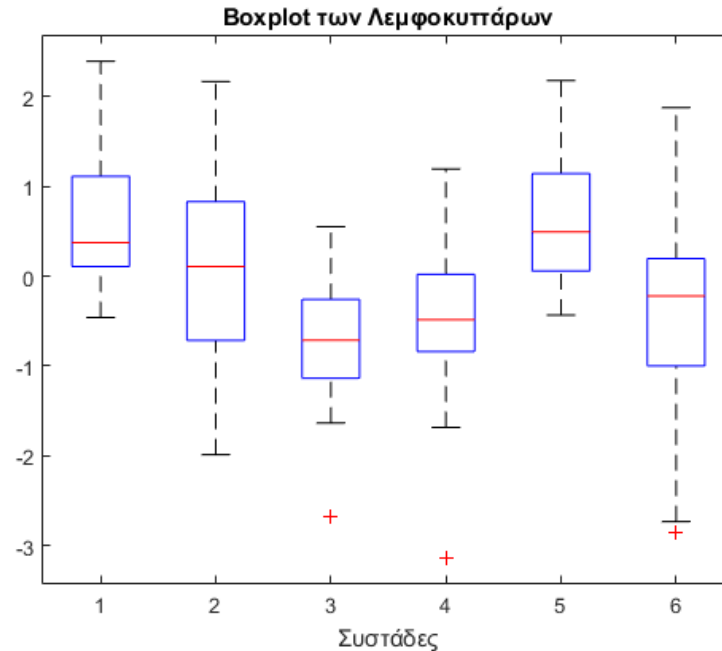
- 1) Αιμοσφαιρίνη (g/dL) με αντιπροσώπευση από τη γραμμή 2.
- 2) Αιματοκρίτης (%) με αντιπροσώπευση από τη γραμμή 3.

- 3) Αριθμός Ουδετεροφίλων (Κ/μl) με αντιπροσώπευση από τη γραμμή 15.
- 4) Ερυθρά Αιμοσφαίρια (Μ/μl) με αντιπροσώπευση από τη γραμμή 1.
- 5) Λεμφοκύτταρα (%) με αντιπροσώπευση από τη γραμμή 11.
- 6) Ουδετερόφιλα (%) με αντιπροσώπευση από τη γραμμή 10.
- 7) Ουρία (mg/dL) με αντιπροσώπευση από τη γραμμή 27.
- 8) Μέσος Όγκος Ερυθρών (fl) με αντιπροσώπευση από τη γραμμή 4.
- 9) Μέση Περιεκτικότητα ΗGB (pg) με αντιπροσώπευση από τη γραμμή 5.
- 10) Εύρος Κατανομής Ερυθρών (%) με αντιπροσώπευση από τη γραμμή 7.

Αυτοί επομένως είναι οι κύριοι δείκτες που λόγω της μεγάλης απόκλισης μεταξύ των μέσων τιμών τους, διαχωρίζουν τους ασθενείς σε τρεις ομάδες. Αυτό γίνεται αντιληπτό από την κατασκευή των boxplots, τα οποία παρουσιάζουν τη μέση τιμή που λαμβάνουν τα στοιχεία μιας συστάδας για ένα συγκεκριμένο δείκτη. Τέλος, παρατίθενται τα box plot της ουρίας και των λεμφοκυττάρων.



Εικόνα 3.9 : Το boxplot του συμβατικού δείκτη της Ουρίας.



Εικόνα 3.10 : Το boxplot του συμβατικού δείκτη των λεμφοκυττάρων.

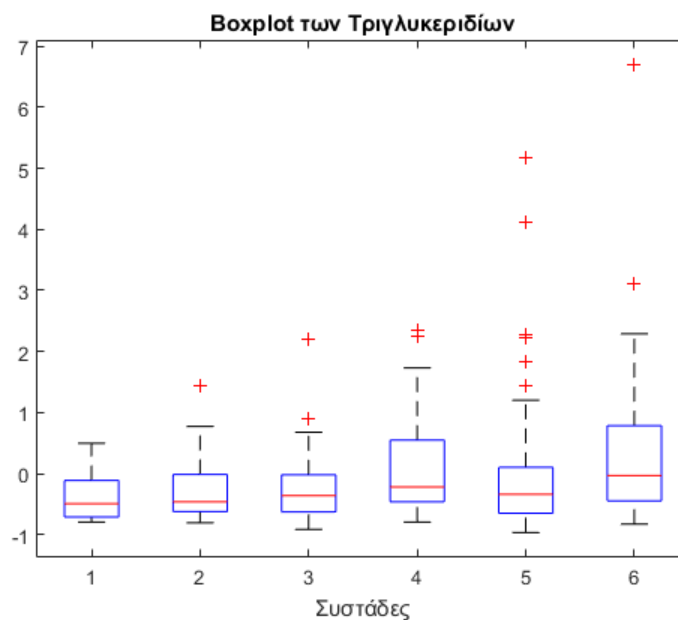
Λόγω των περιορισμένων δεδομένων για τις τιμές GSM και εμβαδού καρωτίδας σε κάθε συστάδα συγκριτικά με το πλήθος των ασθενών που περιλαμβάνουν, δεν μπορεί να προκύψει κάποιο συμπέρασμα για τη συσχέτιση των απεικονιστικών δεικτών με τους συμβατικούς βιοχημικούς δείκτες. Προκειμένου να εξαχθεί κάποιο συμπέρασμα γίνεται συσχέτιση των απεικονιστικών δεικτών (GSMdias, GSMsys και cararea) με τους συμβατικούς βιοχημικούς δείκτες των ασθενών για τους οποίους έχουμε δεδομένα για τις απεικονιστικές τιμές.

Με την εφαρμογή της μεθόδου της συσχέτισης Spearman για μη κανονικά δεδομένα, έγινε συσχέτιση μεταξύ του GSM της διαστολής, της συστολής καθώς και του μέσου εμβαδού της καρωτίδας με τους συμβατικούς βιοχημικούς δείκτες των ασθενών για τους οποίους γνωρίζουμε τους απεικονιστικούς δείκτες (με  $p_{value} < 0.05$ ). Στη συνέχεια παρατίθεται ο πίνακας που περιέχει τις συσχετίσεις που βρέθηκαν να υπάρχουν μεταξύ GSM κατά τη διαστολή, κατά τη συστολή και όσον αφορά το εμβαδό της επιφάνειας και συμβατικών βιοχημικών δεικτών.

Πίνακας 3.4 : Οι συμβατικοί βιοχημικοί δείκτες που παρουσιάζουν τη μεγαλύτερη συσχέτιση με το GSM διαστολής.

GSM κατά τη διαστολή		
	Συσχέτιση	p value
<b>Αιματοκρίτης</b>	0.31	0.011
<b>Ερυθρά Αιμοσφαίρια</b>	0.29	0.019
<b>Αιμοσφαιρίνη</b>	0.27	0.032
<b>Μεταμ.+Μυελ.+Προμ</b>	0.27	0.035
<b>Γλυκόζη</b>	-0.32	0.008
<b>Τριγλυκερίδια</b>	-0.4	0.001

Βάσει της βιβλιογραφίας που προηγήθηκε στο κεφάλαιο 2, τα τριγλυκερίδια αποτελούν έναν παράγοντα επικινδυνότητας για τη νόσο της αθηρωμάτωσης. Από τον Πίνακα 3.4 διαφαίνεται αρνητική συσχέτιση μεταξύ του δείκτη των τριγλυκεριδίων και του δείκτη GSM κατά τη διαστολή. Από την Εικόνα 3.7 φαίνεται ότι η μεγαλύτερη μέση τιμή του δείκτη GSM κατά τη διαστολή παρατηρείται στη δεύτερη συστάδα. Επομένως εάν η συσχέτιση είναι σωστή, θα πρέπει στη δεύτερη συστάδα να εμφανίζεται η χαμηλότερη μέση τιμή του δείκτη των τριγλυκεριδίων, γεγονός που επαληθεύεται από το παρακάτω boxplot για το δείκτη των τριγλυκεριδίων. Η πρώτη συστάδα αποτελείται από ένα μόνο ασθενή για τον οποίο γνωρίζουμε τα απεικονιστικούς δείκτες και παραλείπεται καθώς δεν αποτελεί αντιπροσωπευτικό δείγμα.

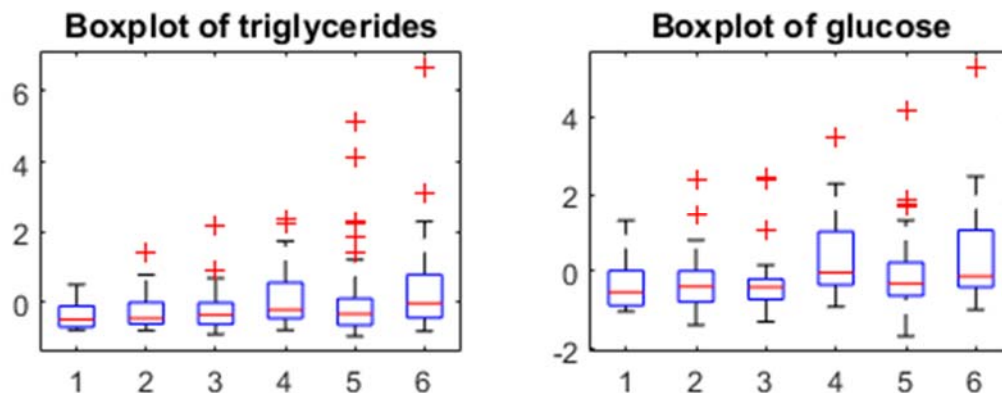


Εικόνα 3.11 : Boxplot των τριγλυκεριδίων.

Πίνακας 3.5 : Οι συμβατικοί βιοχημικοί δείκτες που παρουσιάζουν τη μεγαλύτερη συσχέτιση με το GSM διαστολής.

<b>GSM κατά τη συστολή</b>		
	Συσχέτιση	p value
<b>Γλυκόζη</b>	-0.34	0.006
<b>Τριγλυκερίδια</b>	-0.38	0.001

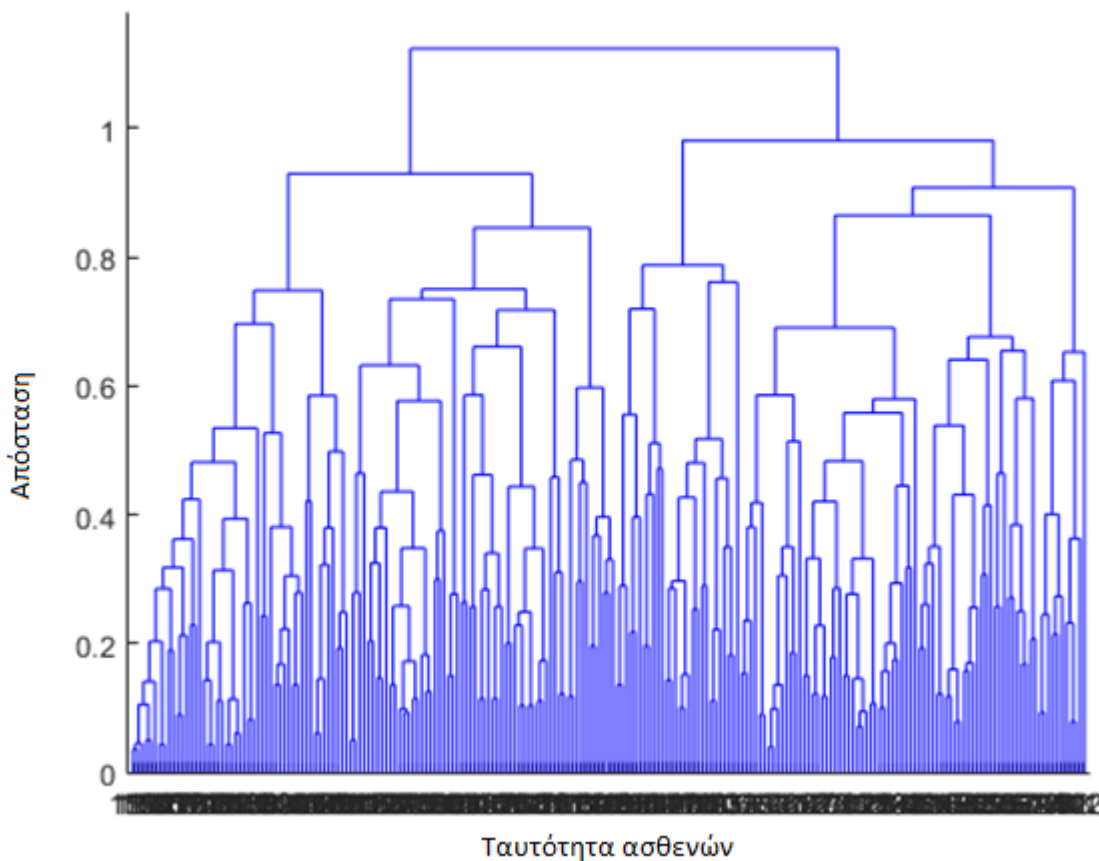
Από τον παραπάνω πίνακα προκύπτει ότι ο δείκτης GSM κατά τη συστολή, παρουσιάζει αρνητική συσχέτιση με τη γλυκόζη και τα τριγλυκερίδια. Επομένως, στην έκτη συστάδα όπου η μέση τιμή του δείκτη GSM κατά τη συστολή λαμβάνει την μικρότερη τιμή (Εικόνα 3.6), αν είναι σωστή η συσχέτιση, θα πρέπει η γλυκόζη και τα τριγλυκερίδια να εμφανίζουν μέγιστη μέση τιμή στη συγκεκριμένη συστάδα. Το γεγονός αυτό επιβεβαιώνεται από τα παρακάτω boxplots (Εικόνα 3.12). Συσχέτιση μεταξύ του εμβαδού της καρωτίδας και κάποιου συμβατικού βιοχημικού δείκτη δεν βρέθηκε.



Εικόνα 3.12 : Boxplots των τριγλυκεριδίων και της γλυκόζης.

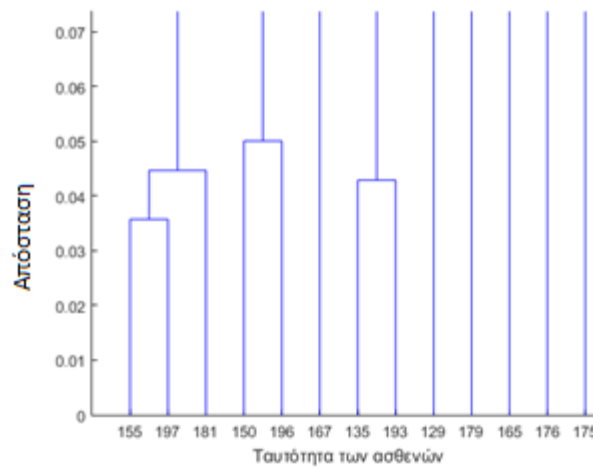
### 3.2 Χρήση εξειδικευμένων βιοχημικών δεικτών

Βασιζόμενοι στους εξειδικευμένους βιοχημικούς δείκτες, κατασκευάστηκε το δενδρογράμμα στο οποίο κάθε φύλλο αναπαριστά αρχικά κάθε ασθενή και κάθε κόμβος αναπαριστά τη συγχώνευση των ασθενών και των συστάδων (Εικόνα 3.13). Επιπλέον λόγω πολλαπλής επικάλυψης των ταυτοτήτων των ασθενών στον οριζόντιο άξονα, παρουσιάζεται ένα μεγεθυμένο τμήμα του δενδρογράμματος, στο οποίο ο διαχωρισμός των ταυτοτήτων αυτών είναι εμφανής (Εικόνα 3.14).



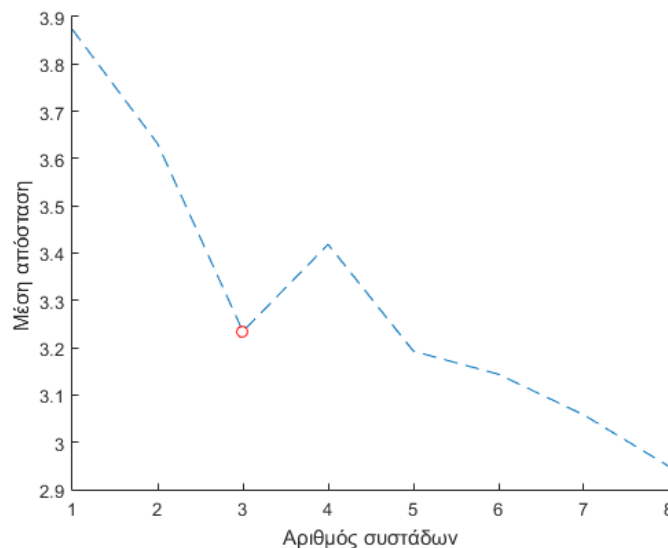
Εικόνα 3.13 : Γραφική αναπαράσταση του δενδρογράμματος των εξειδικευμένων βιοχημικών δεικτών.





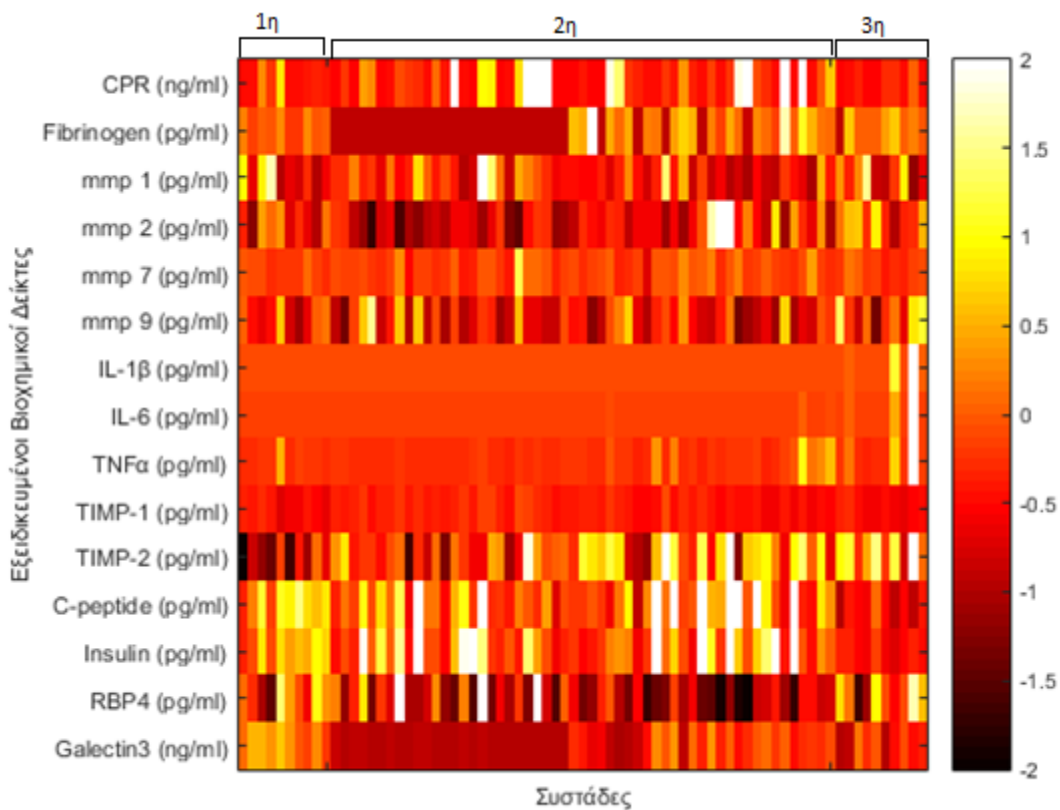
Εικόνα 3.14 : Μεγεθυμένο τμήμα του δενδρογράμματος των εξειδικευμένων βιοχημικών δεικτών.

Στη συγκεκριμένη περίπτωση των εξειδικευμένων βιοχημικών δεικτών, ο βέλτιστος αριθμός των συστάδων είναι ο αριθμός τρία. Ο αριθμός αυτός, επιλέχθηκε υπολογίζοντας το σημείο της γραφικής παράστασης στο οποίο παρατηρείτε μέγιστη πτώση. Αυτό σημαίνει ότι η ομαδοποίηση των ασθενών βάσει των εξειδικευμένων βιοχημικών δεικτών θα γίνει σε τρεις φαινότυπους.



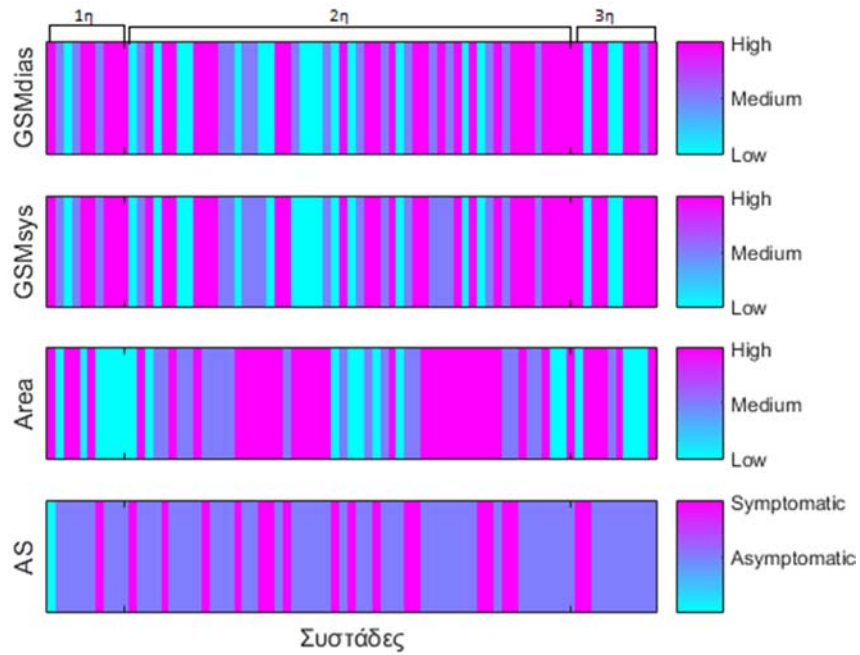
Εικόνα 3.15 : Γραφική παράσταση μέσης απόστασης των στοιχείων μεταξύ των ομάδων συναρτήσει του πλήθους των ομάδων για την εύρεση του βέλτιστου αριθμού συστάδων.

Στη συνέχεια χρησιμοποιήθηκαν τα αποτελέσματα της ιεραρχικής συσταδοποίησης για την κατασκευή ενός heatmap όπου οι στήλες αναπαριστούν τους ασθενείς και οι γραμμές τους εξειδικευμένους βιοχημικούς δείκτες των εξετάσεων. Το heatmap αποτελεί στην ουσία οπτικοποίηση των ιδιοτήτων των φαινοτύπων που προέκυψαν από την ιεραρχική συσταδοποίηση.

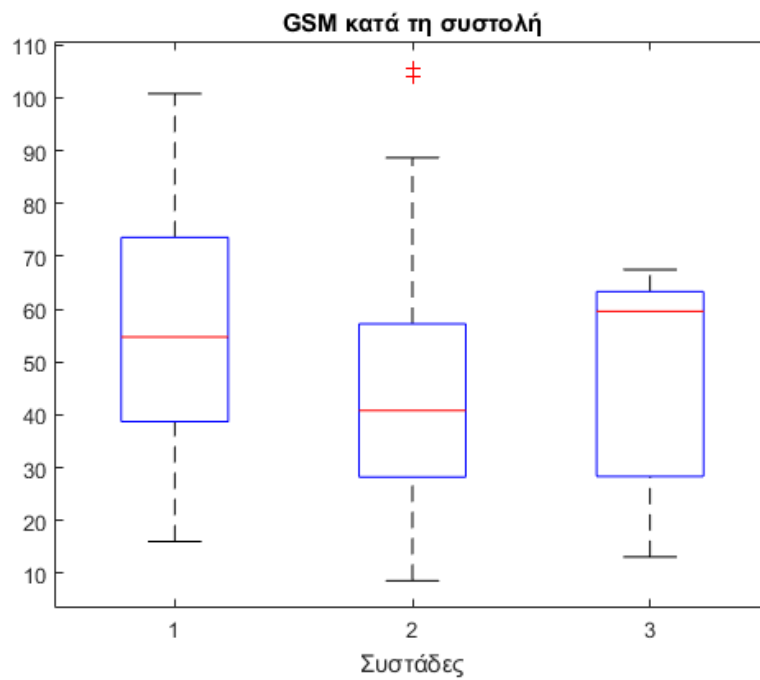


Εικόνα 3.16 : Το heatmap που αντιστοιχεί στην συσταδοποίηση βάσει των εξειδικευμένων βιοχημικών δεικτών.

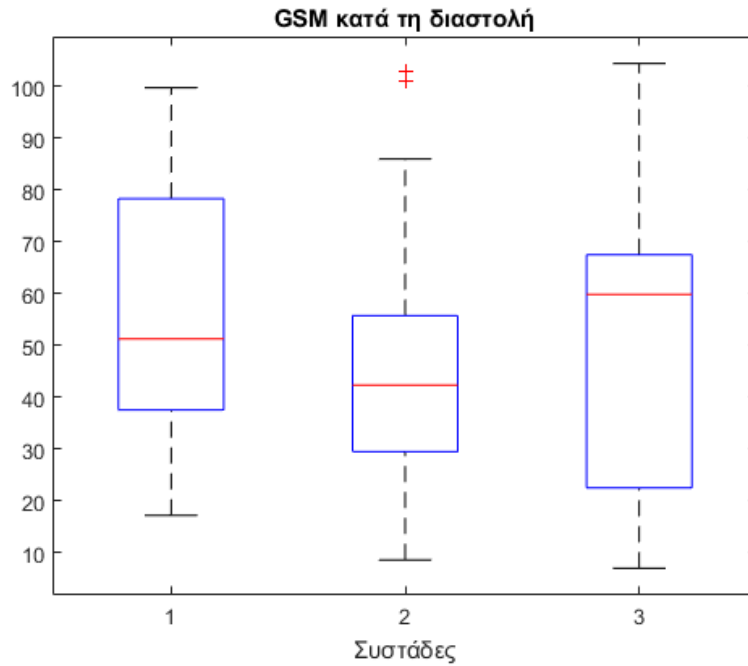
Επιπλέον, χρησιμοποιώντας δεδομένα για τα χαρακτηριστικά των ασθενών όπως η τιμή του δείκτη GSM, ο βαθμός στένωσης, το εμβαδό της πλάκας καθώς και η συμπτωματολογία των αγγείων και κατασκευάστηκαν γραφήματα τα οποία οπτικοποιούν τη σχέση των δεικτών αυτών με τις ομάδες (clusters) των ασθενών.



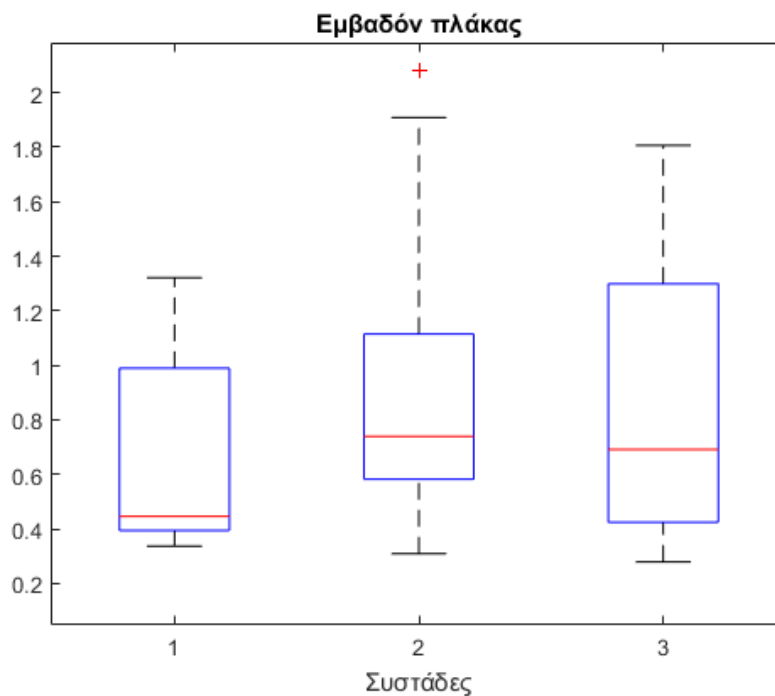
Εικόνα 3.17 : Απεικονιστικά χαρακτηριστικά των συστάδων.



Εικόνα 3.18 : Βοχplot για το συστολικό GSM κατά τη συσταδοποίηση με βάση τους εξειδικευμένους βιοχημικούς δείκτες.



Εικόνα 3.19 : Βοχplot για το διαστολικό GSM κατά τη συσταδοποίηση με βάση τους εξειδικευμένους βιοχημικούς δείκτες.



Εικόνα 3.20 : Βοχplot για το διαστολικό GSM κατά τη συσταδοποίηση με βάση τους εξειδικευμένους βιοχημικούς δείκτες.

Στη συνέχεια οι συνεχείς τιμές των δεικτών GSM της διαστολής, της συστολής και του εμβαδού της καρωτίδας, ταξινομήθηκαν σε τρεις κατηγορίες (χαμηλό, μεσαίο, υψηλό). Ακολουθεί ο πίνακας των ποσοστών των ασθενών κάθε φαινοτύπου που εμπίπτει στις κατηγορίες υψηλή, μεσαία και χαμηλή.

Πίνακας 3.6 : Πίνακας των ποσοστών των ασθενών κάθε φαινοτύπου στις κατηγορίες υψηλή, μεσαία και χαμηλή.

	<b>GSM διαστολής</b>			
	Συνολικό πλήθος ασθενών ανά συστάδα	Κατηγορία Χαμηλή (%)	Κατηγορία Μεσαία (%)	Κατηγορία Υψηλή (%)
<b>Συστάδα 1</b>	28	10	30	60
<b>Συστάδα 2</b>	78	27.27	27.27	45.45
<b>Συστάδα 3</b>	109	30	10	60

Πίνακας 3.7 : Πίνακας των ποσοστών των ασθενών κάθε φαινοτύπου στις κατηγορίες υψηλή, μεσαία και χαμηλή.

	<b>GSM συστολής</b>			
	Συνολικό πλήθος ασθενών ανά συστάδα	Κατηγορία Χαμηλή (%)	Κατηγορία Μεσαία (%)	Κατηγορία Υψηλή (%)
<b>Συστάδα 1</b>	28	10	30	60
<b>Συστάδα 2</b>	78	27.27	29.1	43.64
<b>Συστάδα 3</b>	109	30	0	70

Πίνακας 3.8 : Πίνακας των ποσοστών των ασθενών κάθε φαινοτύπου στις κατηγορίες υψηλή, μεσαία και χαμηλή.

<b>Εμβαδόν καρωτίδας</b>				
	Συνολικό πλήθος ασθενών ανά συστάδα	Κατηγορία  Χαμηλή (%)	Κατηγορία  Μεσαία (%)	Κατηγορία  Υψηλή (%)
<b>Συστάδα 1</b>	28	60	0	40
<b>Συστάδα 2</b>	78	16.36	32.7	50.9
<b>Συστάδα 3</b>	109	40	10	50

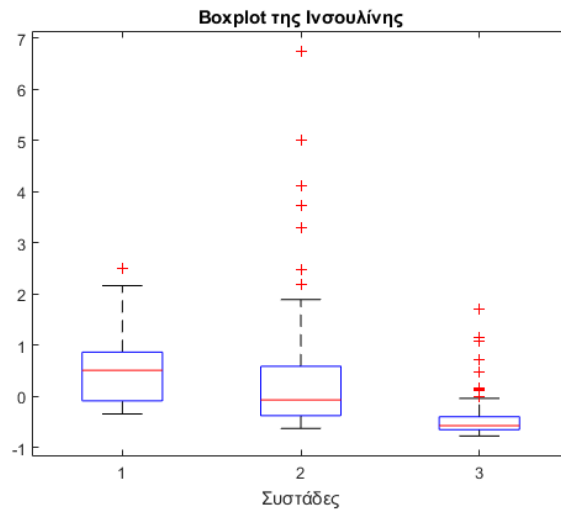
<b>Ασυμπτωματικοί/Συμπτωματικοί</b>			
	Συνολικό πλήθος ασθενών ανά συστάδα	Κατηγορία  Ασυμπτωματικοί (%)	Κατηγορία  Συμπτωματικοί (%)
<b>Συστάδα 1</b>	28	80	20
<b>Συστάδα 2</b>	78	70.90	29.10
<b>Συστάδα 3</b>	109	80	20

Προκειμένου να ερμηνευτεί το παραπάνω heatmap, έγινε χρήση της μεθόδου Kruskal–Wallis, η οποία επιστρέφει ως αποτέλεσμα τους δείκτες οι οποίοι παρουσιάζουν την πιο μεγάλη διαφοροποίηση μεταξύ των φαινοτύπων για κανονικά δεδομένα. Τα αποτελέσματα της μεθόδου ήταν τα εξής :

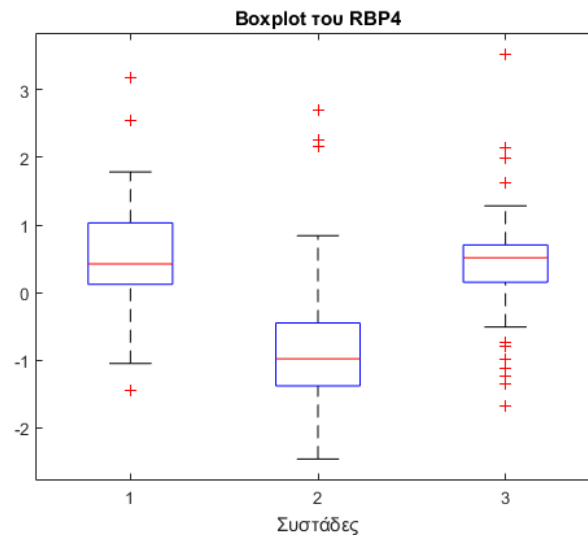
- 1) Insulin (pg/ml) με αντιπροσώπευση από τη γραμμή 13.
- 2) RBP4 (pg/ml) με αντιπροσώπευση από τη γραμμή 14.
- 3) C-peptide (pg/ml) με αντιπροσώπευση από τη γραμμή 12.
- 4) Galectin3 (ng/ml) με αντιπροσώπευση από τη γραμμή 15.
- 5) mmp2 (pg/ml) με αντιπροσώπευση από τη γραμμή 4.
- 6) IL-1β (pg/ml) με αντιπροσώπευση από τη γραμμή 7.
- 7) mmp9 (pg/ml) με αντιπροσώπευση από τη γραμμή 6.
- 8) Fibrinogen (pg/ml) με αντιπροσώπευση από τη γραμμή 2.
- 9) TIMP-2 (pg/ml) με αντιπροσώπευση από τη γραμμή 11.

## 10) TNFα (pg/ml) με αντιπροσώπευση από τη γραμμή 9.

Αυτοί επομένως είναι οι κύριοι δείκτες που λόγω της μεγάλης απόκλισης μεταξύ των μέσων τιμών τους, διαχωρίζουν τους ασθενείς σε τρεις ομάδες. Αυτό γίνεται αντιληπτό από την κατασκευή των boxplots, τα οποία παρουσιάζονται στην συνέχεια. Τέλος, παρατίθεται το box plot της ινσουλίνης και του RBP4.



Εικόνα 3.21 : Το boxplot του εξειδικευμένου δείκτη της ινσουλίνης.



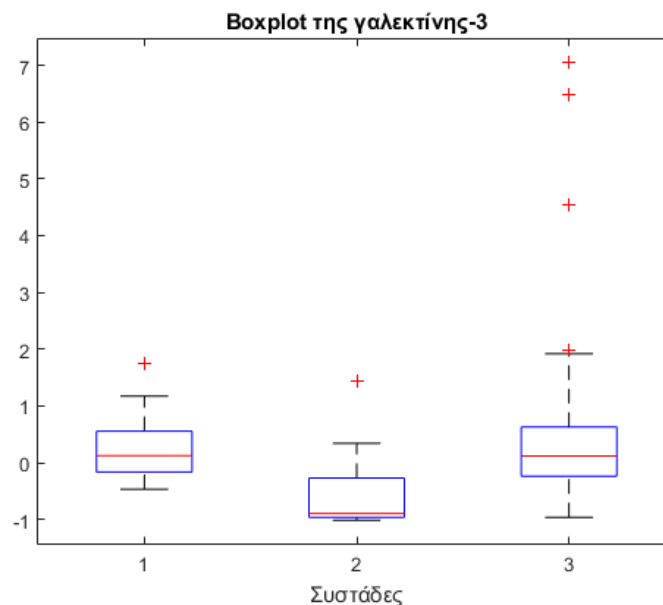
Εικόνα 3.22 : Το boxplot του εξειδικευμένου δείκτη RBP4.

Με την εφαρμογή της μεθόδου της συσχέτισης Spearman για μη κανονικά δεδομένα, έγινε συσχέτιση μεταξύ του GSM της διαστολής, της συστολής καθώς και του μέσου εμβαδού της καρωτίδας με τους εξειδικευμένους βιοχημικούς δείκτες των ασθενών για τους οποίους γνωρίζουμε τους απεικονιστικούς δείκτες (με  $p_{\text{value}} < 0.05$ ). Στη συνέχεια παρατίθεται ο πίνακας που περιέχει τις συσχετίσεις που βρέθηκαν να υπάρχουν μεταξύ GSM κατά τη διαστολή, κατά τη συστολή και όσον αφορά το εμβαδό της επιφάνειας και εξειδικευμένων βιοχημικών δεικτών.

Πίνακας 3.9 : Οι εξειδικευμένοι βιοχημικοί δείκτες που παρουσιάζουν τη μεγαλύτερη συσχέτιση με το GSM διαστολής.

GSM κατά τη διαστολή		
	Συσχέτιση	p value
<b>Galectin3</b>	0.24	0.003
<b>mmp7</b>	-0.23	0.045
<b>TIMP-1</b>	-0.4	0.000

Συνεπώς αναμένεται θετική συσχέτιση μεταξύ του δείκτη GSM της διαστολής με την Γαλεκτίνη-3, όπως φαίνεται ενδεικτικά στο παρακάτω boxplot, και αρνητική συσχέτιση με τον εξειδικευμένο δείκτη TIMP1. Μέγιστη μέση τιμή του δείκτη GSM της διαστολής παρατηρείται στη τρίτη συστάδα όπως φαίνεται στην (Εικόνα 3.19).



Εικόνα 3.23 : Boxplot της Γαλεκτίνης-3.



Πίνακας 3.10 : Οι εξειδικευμένοι βιοχημικοί δείκτες που παρουσιάζουν τη μεγαλύτερη συσχέτιση με το GSM συστολής.

<b>GSM κατά τη συστολή</b>		
	Συσχέτιση	p value
<b>TIMP-1</b>	-0.4	0.001

Συνεπώς αναμένεται αρνητική συσχέτιση μεταξύ του δείκτη GSM της συστολής με τον δείκτη TIMP-1.

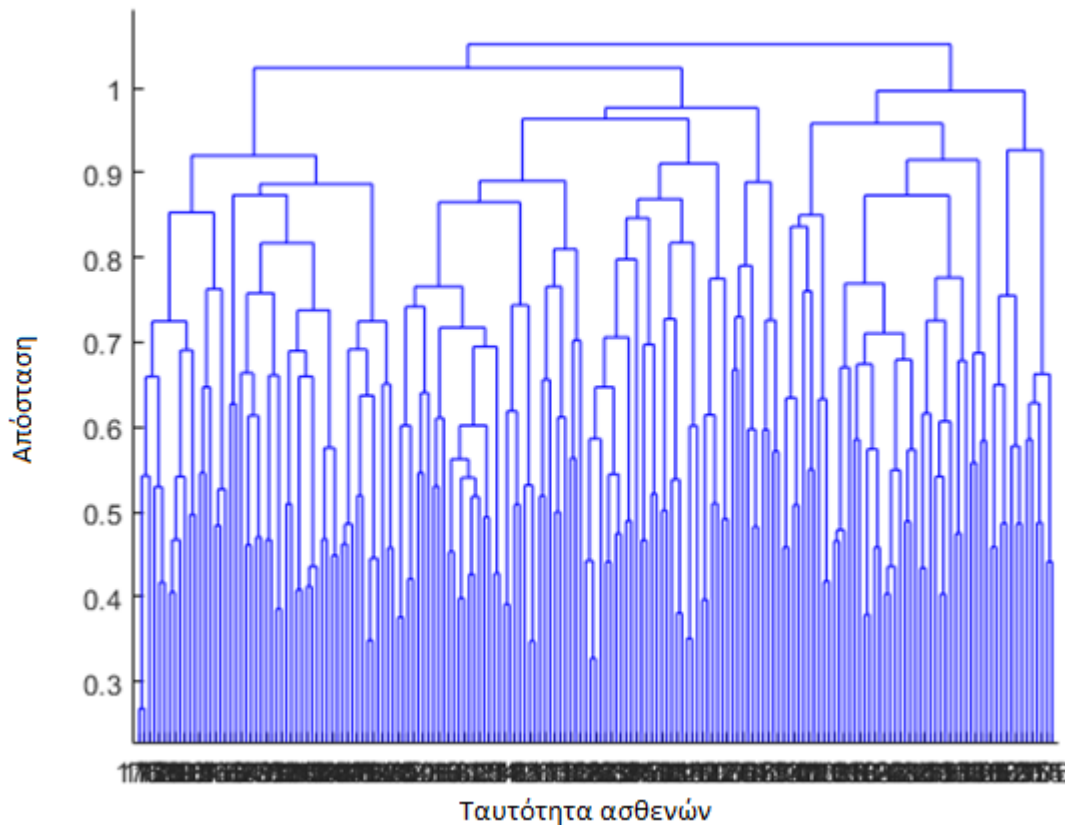
Πίνακας 3.11 : Οι εξειδικευμένοι βιοχημικοί δείκτες που παρουσιάζουν τη μεγαλύτερη συσχέτιση με το εμβαδόν της καρωτίδας.

<b>Εμβαδόν καρωτίδας</b>		
	Συσχέτιση	p value
<b>RBP4</b>	-0.3	0.009

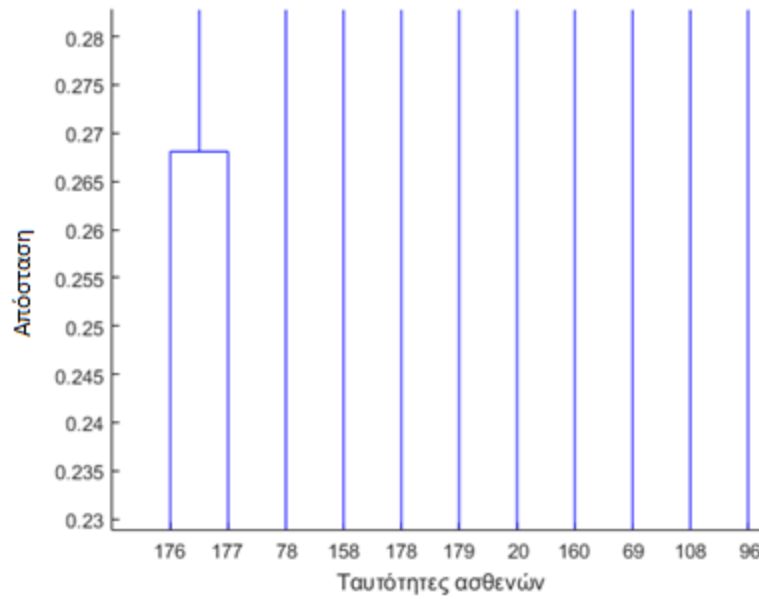
Από τον παραπάνω πίνακα προκύπτει ότι αναμένεται αρνητική συσχέτιση μεταξύ του δείκτη RBP4 και του εμβαδού της καρωτίδας. Η υπόθεση αυτή πιστοποιείται από την Εικόνα 3.20 και την Εικόνα 3.22 αφού το εμβαδόν της καρωτίδας εμφανίζει την ελάχιστη μέση τιμή στην πρώτη συστάδα και ο δείκτης RBP4 εμφανίζει μέγιστη τιμή στη συστάδα αυτή.

### 3.3 Συνδυασμός συμβατικών και εξειδικευμένων δεικτών

Βασιζόμενοι στους συμβατικούς και εξειδικευμένους βιοχημικούς δείκτες, κατασκευάστηκε το δενδρόγραμμα στο οποίο κάθε φύλλο αναπαριστά αρχικά κάθε ασθενή και κάθε κόμβος αναπαριστά τη συγχώνευση των ασθενών και των συστάδων (Εικόνα 3.24). Επιπλέον λόγω πολλαπλής επικάλυψης των ταυτοτήτων των ασθενών στον οριζόντιο άξονα, παρουσιάζεται ένα μεγεθυμένο τμήμα του δενδρογράμματος, στο οποίο ο διαχωρισμός των ταυτοτήτων αυτών είναι εμφανής (Εικόνα 3.25).

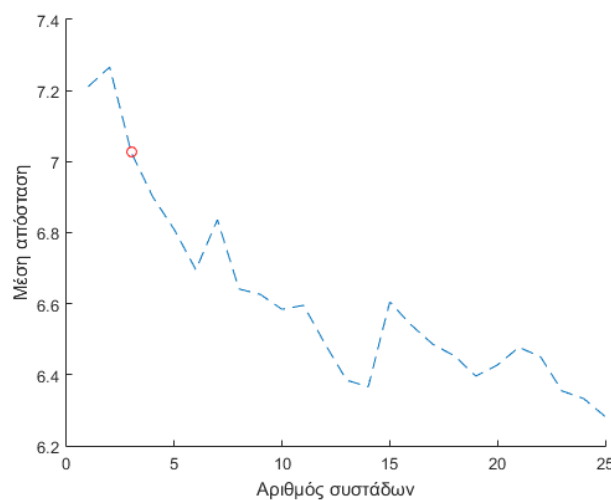


Εικόνα 3.24 : Γραφική αναπαράσταση του δενδρογράμματος των συμβατικών και εξειδικευμένων βιοχημικών δεικτών.



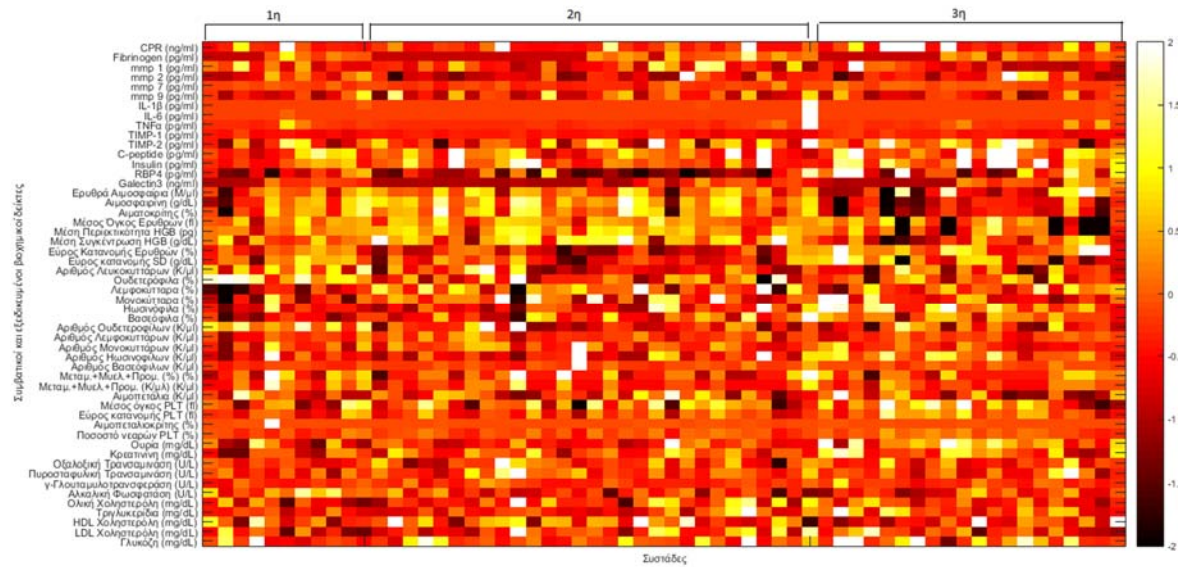
Εικόνα 3.25 : Μεγεθυμένο τμήμα του δενδρογράμματος των εξειδικευμένων και συμβατικών βιοχημικών δεικτών.

Στη συγκεκριμένη περίπτωση των εξειδικευμένων βιοχημικών δεικτών, ο βέλτιστος αριθμός των συστάδων είναι ο αριθμός τρία. Ο βέλτιστος αυτός αριθμός, επιλέχθηκε υπολογίζοντας το σημείο της γραφικής παράστασης στο οποίο παρατηρείτε μέγιστη πτώση. Αυτό σημαίνει ότι η ομαδοποίηση των ασθενών βάσει των εξειδικευμένων βιοχημικών δεικτών θα γίνει σε τρεις φαινότυπους.



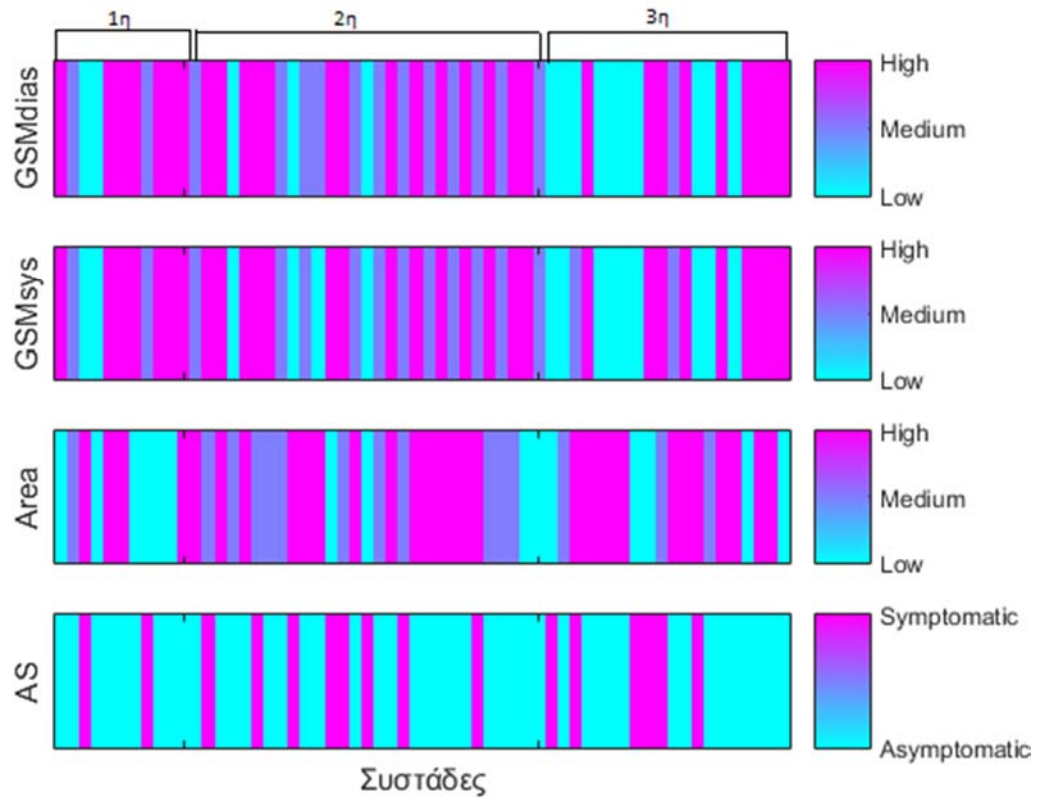
Εικόνα 3.26 : Γραφική παράσταση μέσης απόστασης των στοιχείων μεταξύ των ομάδων συναρτήσει του πλήθους των ομάδων για την εύρεση του βέλτιστου αριθμού συστάδων.

Στη συνέχεια χρησιμοποιήθηκαν τα αποτελέσματα της ιεραρχικής συσταδοποίησης για την κατασκευή ενός heatmap όπου οι στήλες αναπαριστούν τους ασθενείς και οι γραμμές τους εξειδικευμένους και συμβατικούς βιοχημικούς δείκτες των εξετάσεων. Το heatmap αποτελεί στην ουσία οπτικοποίηση των ιδιοτήτων των φαινοτύπων που προέκυψαν από την ιεραρχική συσταδοποίηση.

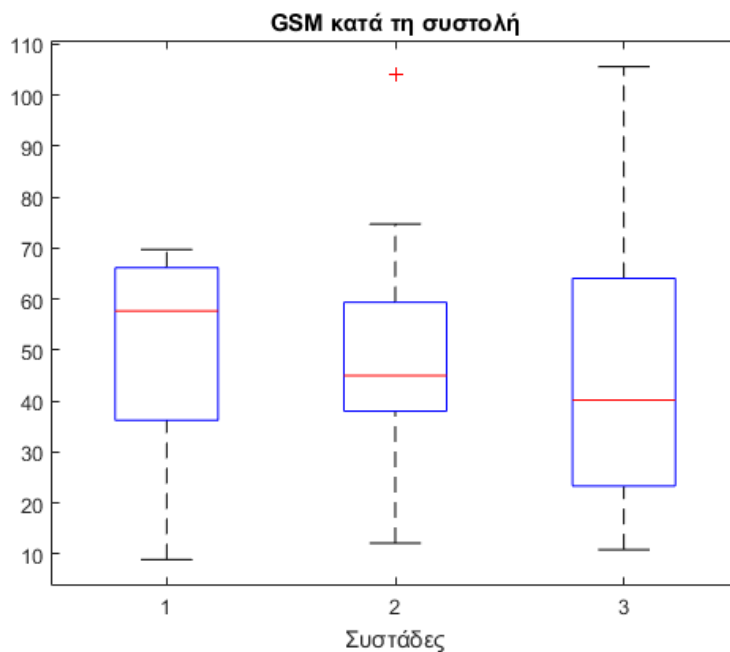


Εικόνα 3.27 : Το heatmap που αντιστοιχεί στην συσταδοποίηση βάσει των συμβατικών και των εξειδικευμένων βιοχημικών δεικτών.

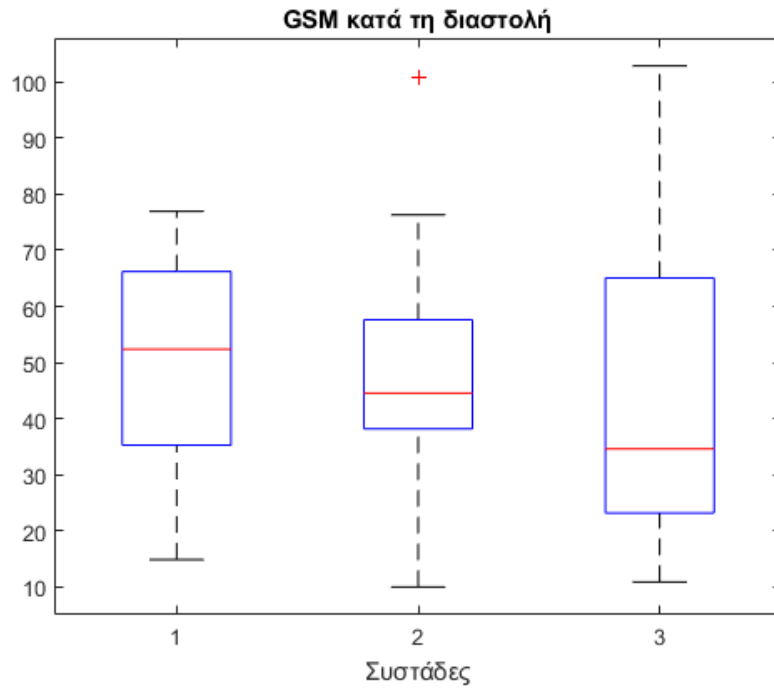
Επιπλέον, χρησιμοποιώντας δεδομένα για τα χαρακτηριστικά των ασθενών όπως η τιμή του δείκτη GSM, ο βαθμός στένωσης, το εμβαδό της πλάκας καθώς και η συμπτωματολογία των αγγείων και κατασκευάστηκαν γραφήματα τα οποία οπτικοποιούν τη σχέση των δεικτών αυτών με τις ομάδες (clusters) των ασθενών.



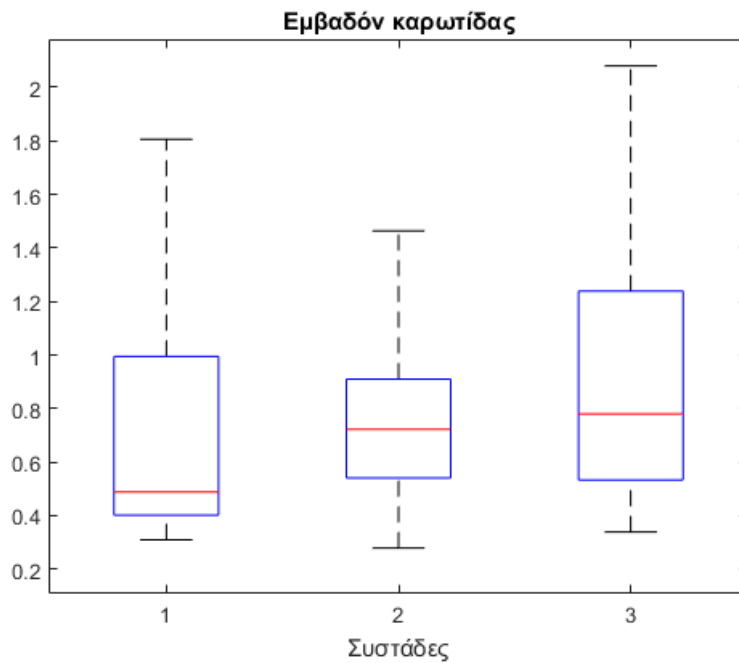
Εικόνα 3.28 : Απεικονιστικά χαρακτηριστικά των συστάδων



Εικόνα 3.29 : Βοχplot για το συστολικό GSM κατά τη συσταδοποίηση με βάση τους εξειδικευμένους και συμβατικούς βιοχημικούς δείκτες.



Εικόνα 3.30 : Βοχplot για το διαστολικό GSM κατά τη συσταδοποίηση με βάση τους εξειδικευμένους και συμβατικούς βιοχημικούς δείκτες.



Εικόνα 3.31 : Βοχplot για το εμβαδόν της καρωτίδας κατά τη συσταδοποίηση με βάση τους εξειδικευμένους και συμβατικούς βιοχημικούς δείκτες.

Στη συνέχεια οι συνεχείς τιμές των δεικτών GSM της διαστολής, της συστολής και του εμβαδού της καρωτίδας, ταξινομήθηκαν σε τρεις κατηγορίες (χαμηλό, μεσαίο, υψηλό). Ακολουθεί ο πίνακας των ποσοστών των ασθενών κάθε φαινοτύπου που εμπίπτει στις κατηγορίες υψηλή, μεσαία και χαμηλή.

Πίνακας 3.12 : Πίνακας των ποσοστών των ασθενών κάθε φαινοτύπου στις κατηγορίες υψηλή, μεσαία και χαμηλή.

	<b>GSM διαστολής</b>			
	Συνολικό πλήθος ασθενών ανά συστάδα	Κατηγορία Χαμηλή (%)	Κατηγορία Μεσαία (%)	Κατηγορία Υψηλή (%)
<b>Συστάδα 1</b>	51	18.18	18.18	63.63
<b>Συστάδα 2</b>	76	10.34	41.37	48.27
<b>Συστάδα 3</b>	54	50	5	45

Πίνακας 3.13 : Πίνακας των ποσοστών των ασθενών κάθε φαινοτύπου στις κατηγορίες υψηλή, μεσαία και χαμηλή.

	<b>GSM συστολής</b>			
	Συνολικό πλήθος ασθενών ανά συστάδα	Κατηγορία Χαμηλή (%)	Κατηγορία Μεσαία (%)	Κατηγορία Υψηλή (%)
<b>Συστάδα 1</b>	51	18.18	18.18	63.63
<b>Συστάδα 2</b>	76	13.79	37.93	48.27
<b>Συστάδα 3</b>	54	45	10	45

Πίνακας 3.14 : Πίνακας των ποσοστών των ασθενών κάθε φαινοτύπου στις κατηγορίες υψηλή, μεσαία, χαμηλή και του είδους συμπτωμάτων.

	<b>Εμβαδόν καρωτίδας</b>			
	Συνολικό πλήθος ασθενών ανά συστάδα	Κατηγορία Χαμηλή (%)	Κατηγορία Μεσαία (%)	Κατηγορία Υψηλή (%)
<b>Συστάδα 1</b>	51	54.54	9.09	36.36
<b>Συστάδα 2</b>	76	13.79	37.93	48.27
<b>Συστάδα 3</b>	54	25	15	60

	<b>Ασυμπτωματικοί/Συμπτωματικοί</b>		
	Συνολικό πλήθος ασθενών ανά συστάδα	Κατηγορία Ασυμπτωματικοί (%)	Κατηγορία Συμπτωματικοί (%)
<b>Συστάδα 1</b>	51	81.81	18.18
<b>Συστάδα 2</b>	76	72.24	27.58
<b>Συστάδα 3</b>	54	70	30

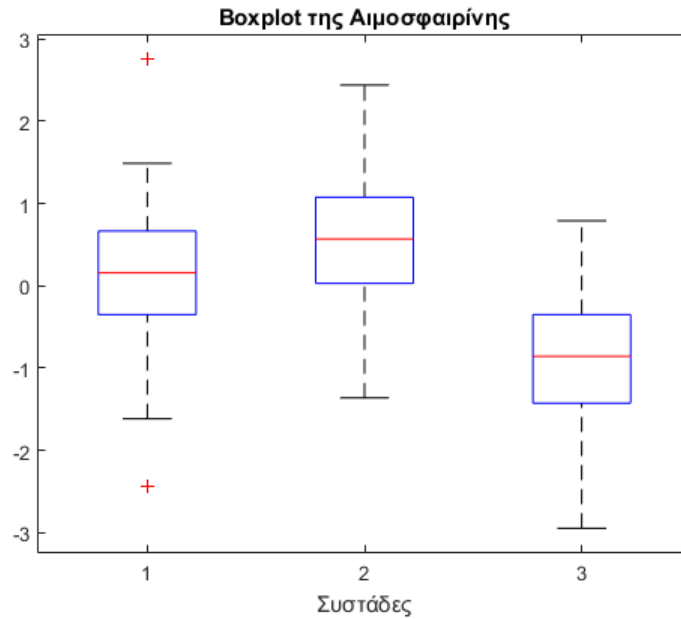
Προκειμένου να ερμηνευτεί το παραπάνω heatmap, έγινε χρήση της μεθόδου Kruskal–Wallis, η οποία επιστρέφει ως αποτέλεσμα τους δείκτες οι οποίοι παρουσιάζουν την πιο μεγάλη διαφοροποίηση μεταξύ των φαινοτύπων για κανονικά δεδομένα. Τα αποτελέσματα της μεθόδου ήταν τα εξής :

- 1) Αιμοσφαιρίνη (g/dL) με αντιπροσώπευση από τη γραμμή 17.
- 2) Αιματοκρίτης (%) με αντιπροσώπευση από τη γραμμή 18.
- 3) Ουδετερόφιλα (%) με αντιπροσώπευση από τη γραμμή 25.
- 4) Αριθμός Ουδετεροφίλων (Κ/μl) με αντιπροσώπευση από τη γραμμή 30.
- 5) Λεμφοκύτταρα (%) με αντιπροσώπευση από τη γραμμή 26.
- 6) Ηωσινόφιλα (%) με αντιπροσώπευση από τη γραμμή 28.
- 7) Αριθμός Λευκοκυττάρων (Κ/μl) με αντιπροσώπευση από τη γραμμή 24.
- 8) Βασεόφιλα (%) με αντιπροσώπευση από τη γραμμή 29.
- 9) Ερυθρά Αιμοσφαίρια (Μ/μl) με αντιπροσώπευση από τη γραμμή 16.

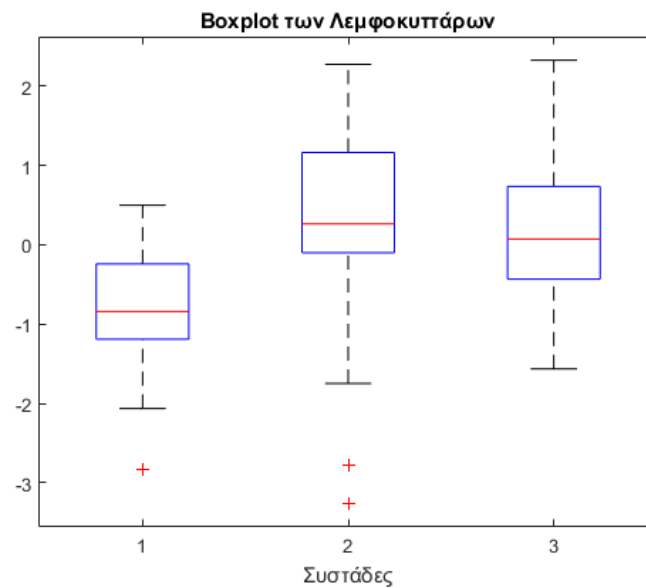


- 10) C-peptide (pg/ml) με αντιπροσώπευση από τη γραμμή 12.
- 11) Εύρος Κατανομής Ερυθρών (%) με αντιπροσώπευση από τη γραμμή 22.
- 12) Ουρία (mg/dL) με αντιπροσώπευση από τη γραμμή 42.
- 13) Αλκαλική Φωσφατάση (U/L) με αντιπροσώπευση από τη γραμμή 47.
- 14) Μέση Περιεκτικότητα HGB (pg) με αντιπροσώπευση από τη γραμμή 20.
- 15) Μέση Συγκέντρωση HGB (g/dL) με αντιπροσώπευση από τη γραμμή 21.
- 16) mmp 7 (pg/ml) με αντιπροσώπευση από τη γραμμή 5.
- 17) Insulin (pg/ml) με αντιπροσώπευση από τη γραμμή 13.
- 18) mmp 2 (pg/ml) με αντιπροσώπευση από τη γραμμή 4.
- 19) RBP4 (pg/ml) με αντιπροσώπευση από τη γραμμή 14.
- 20) Αριθμός Ηωσινοφίλων (K/μl) με αντιπροσώπευση από τη γραμμή 33.
- 21) Ολική Χοληστερόλη (mg/dL) με αντιπροσώπευση από τη γραμμή 48.
- 22) Galectin3 (ng/ml) με αντιπροσώπευση από τη γραμμή 15.
- 23) LDL Χοληστερόλη (mg/dL) με αντιπροσώπευση από τη γραμμή 51.
- 24) mmp9 (pg/ml) με αντιπροσώπευση από τη γραμμή 6.
- 25) Fibrinogen (pg/ml) με αντιπροσώπευση από τη γραμμή 2.
- 26) mmp 1 (pg/ml) με αντιπροσώπευση από τη γραμμή 3.
- 27) IL-1β (pg/ml) με αντιπροσώπευση από τη γραμμή 7.
- 28) Γλυκόζη (mg/dL) με αντιπροσώπευση από τη γραμμή 52.

Αυτοί επομένως είναι οι κύριοι δείκτες που λόγω της μεγάλης απόκλισης μεταξύ των μέσων τιμών τους, διαχωρίζουν τους ασθενείς σε τρεις ομάδες. Αυτό γίνεται αντιληπτό από την κατασκευή των boxplots, τα οποία παρουσιάζονται στην συνέχεια. Τέλος, παρατίθεται το box plot της αιμοσφαιρίνης και των λεμφοκυττάρων.



Εικόνα 3.32 : Boxplot της Αιμοσφαιρίνης.



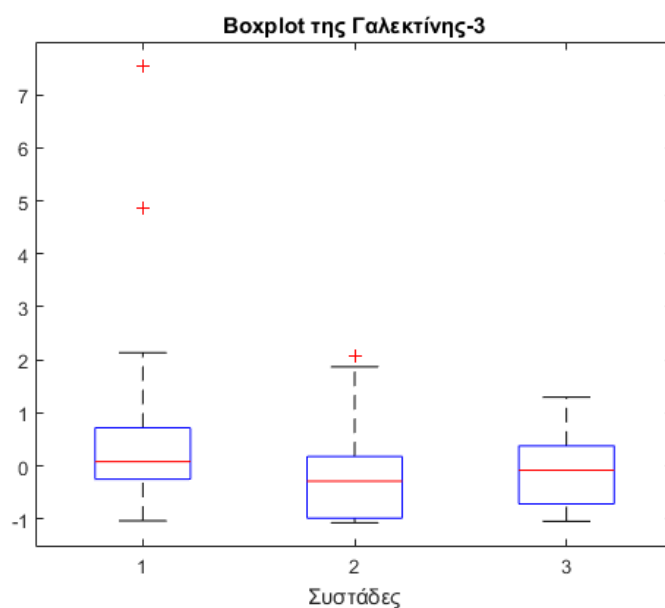
Εικόνα 3.33 : Boxplot των Λεμφοκυττάρων.

Στη συνέχεια παρατίθεται ο πίνακας που περιέχει τις συσχετίσεις που βρέθηκαν να υπάρχουν μεταξύ των εξειδικευμένων βιοχημικών δεικτών και του δείκτη GSM κατά τη διαστολή, κατά τη συστολή και όσον αφορά το εμβαδό της επιφάνειας.

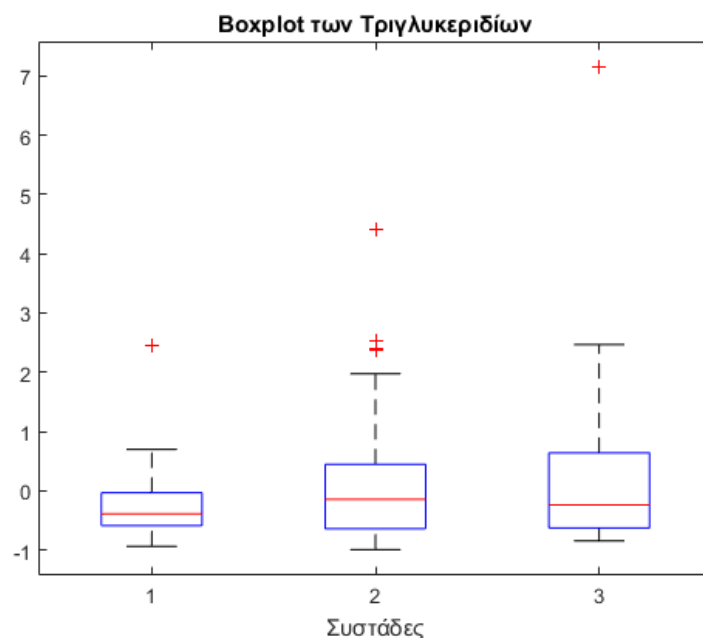
Πίνακας 3.15 : Οι συμβατικοί και εξειδικευμένοι βιοχημικοί δείκτες που παρουσιάζουν τη μεγαλύτερη συσχέτιση με το GSM διαστολής.

<b>GSM κατά τη διαστολή</b>		
	Συσχέτιση	p value
<b>Galectin3</b>	0.33	0.009
<b>Αιματοκρίτης</b>	0.33	0.009
<b>Ερυθρά Αιμοσφαίρια</b>	0.30	0.015
<b>Αιμοσφαιρίνη</b>	0.30	0.018
<b>Γλυκόζη</b>	-0.27	0.030
<b>TIMP-1</b>	-0.34	0.008
<b>Τριγλυκερίδια</b>	-0.35	0.005

Συνεπώς αναμένεται μέγιστη θετική συσχέτιση μεταξύ του δείκτη GSM της διαστολής με την Γαλεκτίνη-3, όπως φαίνεται ενδεικτικά στο παρακάτω boxplot, και μέγιστη αρνητική συσχέτιση με τον δείκτη των τριγλυκεριδίων.



Εικόνα 3.34 : Boxplot της Γαλεκτίνης-3.



Εικόνα 3.35 : Boxplot των Τριγλυκεριδίων.

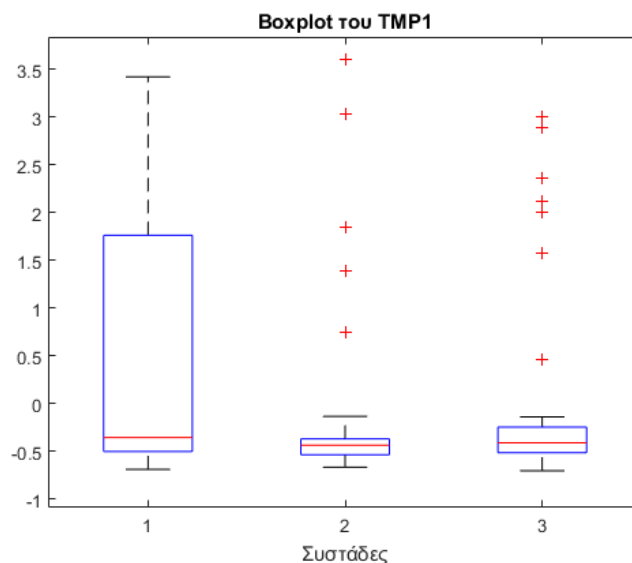
Παρατηρούμε ότι η μέγιστη μέση τιμή του δείκτη GSM κατά τη διαστολή παρουσιάζεται στην πρώτη συστάδα, όπως φαίνεται από την Εικόνα 3.30. Κοιτώντας τώρα το boxplot της Γαλεκτίνης-3, παρατηρούμε ότι η μέγιστη μέση τιμή της γαλεκτίνης εμφανίζεται στην πρώτη συστάδα γεγονός που επιβεβαιώνει την θετική συσχέτιση των δύο δεικτών.

Πίνακας 3.16 : Οι συμβατικοί και εξειδικευμένοι βιοχημικοί δείκτες που παρουσιάζουν τη μεγαλύτερη θετική και αρνητική συσχέτιση με το GSM συστολής.

<b>GSM κατά τη συστολή</b>		
	Συσχέτιση	p value
<b>Αιματοκρίτης</b>	0.33	0.009
<b>Τριγλυκερίδια</b>	-0.33	0.009
<b>TIMP-1</b>	-0.34	0.007

Παρατηρείται ότι ο δείκτης GSM κατά τη συστολή συσχετίζεται θετικά με το δείκτη του αιματοκρίτη και αρνητικά με το δείκτη TIMP-1. Από τη βιβλιογραφία που εξετάσαμε στο δεύτερο κεφάλαιο ο δείκτης TIMP-1 σχετίζεται αρνητικά με το GSM κατά τη συστολή γεγονός που πιστοποιείται από τη σύγκριση του παρακάτω boxplot με το boxplot του δείκτη GSM κατά τη συστολή. Η μέγιστη μέση τιμή του δείκτη GSM κατά τη συστολή παρατηρείται στην πρώτη

συστάδα επομένως θα πρέπει η ελάχιστη μέση τιμή του δείκτη TIMP-1, στο boxplot της μεταβλητής TIMP-1, να παρατηρηθεί πάλι στην πρώτη συστάδα.



Εικόνα 3.36 : Boxplot του TIMP1.

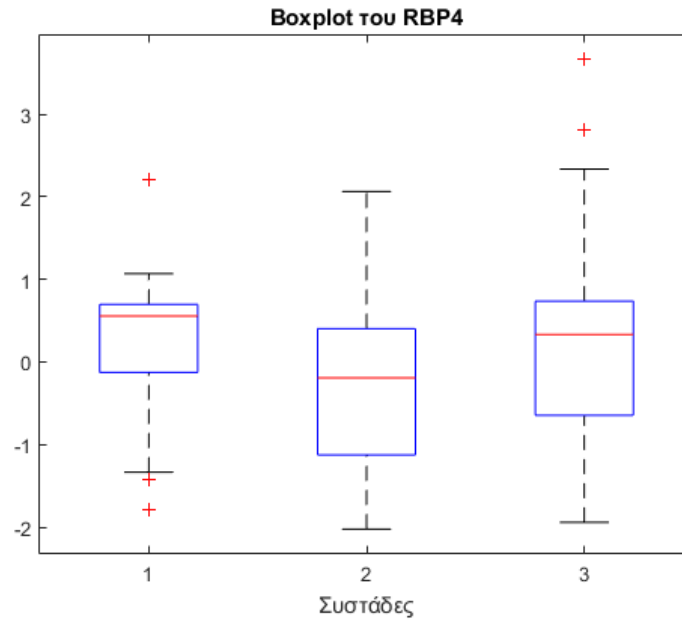
Όπως φαίνεται από την παραπάνω εικόνα, η ελάχιστη μέση τιμή του δείκτη TIMP1 εμφανίζεται στην πρώτη συστάδα αλλά και στις υπόλοιπες δυο. Θα πρέπει να τονιστεί ωστόσο ότι ο ισχυρισμός αυτός είναι βάσιμος και πιστοποιείται και από τη βιβλιογραφία που αναλύθηκε στο δεύτερο κεφάλαιο.

Πίνακας 3.17 : Οι συμβατικοί και εξειδικευμένοι βιοχημικοί δείκτες που παρουσιάζουν τη μεγαλύτερη θετική και αρνητική συσχέτιση με το εμβαδόν της καρωτίδας.

Εμβαδόν καρωτίδας		
	Συσχέτιση	p value
<b>TIMP-1</b>	0.26	0.030
<b>RBP4</b>	-0.33	0.009

Παρατηρούμε τώρα πως ο δείκτης του εμβαδού της καρωτίδας συσχετίζεται θετικά με το βιοχημικό δείκτη TIMP-1 και αρνητικά με το δείκτη RBP4. Το boxplot του δείκτη TIMP-1 φαίνεται στην Εικόνα 3.36. Παρακάτω παρουσιάζεται το boxplot του δείκτη RBP4 στο οποίο η μέγιστη μέση τιμή εμφανίζεται στην πρώτη συστάδα ενώ στην πρώτη συστάδα αντιστοιχεί η

ελάχιστη τιμή του δείκτη του εμβαδού της καρωτίδας, αποδεικνύοντας έτσι την αρνητική τους συσχέτιση.



Εικόνα 3.37 : Boxplot του RBP4.



## Κεφάλαιο 4 : Συζήτηση

### Περίληψη

Το παρόν κεφάλαιο αποτελεί τον επίλογο της παρούσας διπλωματικής εργασίας. Αρχικά γίνεται σχολιασμός επί των αποτελεσμάτων του προηγούμενου κεφαλαίου. Στη συνέχεια πραγματοποιείται σύγκριση των συμπερασμάτων με συναφείς εργασίες που αναφέρθηκαν στο. Τέλος, γίνεται μια επισκόπηση της νόσου και πιθανών μελλοντικών προεκτάσεων της εργασίας και του αλγορίθμου.



## 4.1 Συμπεράσματα

### Αναγνώριση φαινοτύπων με βάση τους βιοχημικούς δείκτες

Ο αλγόριθμος που αναπτύχθηκε τα πλαίσια της παρούσας εργασίας, εντόπισε διαφορετικούς φαινοτύπους με βάση τους βιοχημικούς δείκτες. Σε κάθε περίπτωση εντοπίστηκε ο αριθμός αυτών των φαινοτύπων καθώς και οι βιοχημικοί δείκτες που οδήγησαν, κατά κύριο λόγο, στο διαχωρισμό αυτό.

Στην πρώτη περίπτωση δηλαδή με βάση τους συμβατικούς βιοχημικούς δείκτες, εντοπίστηκαν έξι διαφορετικοί φαινότυποι. Ο διαχωρισμός αυτός οφείλεται στις διακυμάνσεις που εμφανίζουν ορισμένοι συμβατικοί δείκτες μέσα στο δείγμα μας, όπως αναλύθηκε και στο προηγούμενο κεφάλαιο με τη μέθοδο του Kruskal-Wallis. Αυτές οι διακυμάνσεις αναπαρίστανται στο heatmap (Εικόνα 3.4) και οφείλονται σε συμβατικούς βιοχημικούς δείκτες όπως : η αιμοσφαιρίνη, ο αιματοκρίτης, ο αριθμός ουδετεροφίλων, τα ερυθρά αιμοσφαίρια , τα λεμφοκύτταρα, τα ουδετερόφιλα, η ουρία, ο μέσος όγκος ερυθρών, η μέση περιεκτικότητα αιμοσφαιρίνης και το εύρος κατανομής ερυθρών.

Στη δεύτερη περίπτωση, με βάση τους εξειδικευμένους βιοχημικούς δείκτες, εντοπίστηκαν τρεις διαφορετικοί φαινότυποι. Ο διαχωρισμός αυτός οφείλεται στις διακυμάνσεις που εμφανίζουν ορισμένοι εξειδικευμένοι δείκτες μέσα στο δείγμα μας. Αυτές οι διακυμάνσεις αναπαρίστανται στο heatmap (Εικόνα 3.16) και οφείλονται, όπως έδειξε το κριτήριο Kruskal-Wallis, σε εξειδικευμένους βιοχημικούς δείκτες όπως η ινσουλίνη, η RBP4, η C-peptide, η γαλεκτίνη-3, η mmp2, η IL-1β, mmp9, η Fibrinogen, η μεταλλοπρωτεϊνάση-2 και η TNFα.

Τέλος στην τρίτη περίπτωση που χρησιμοποιήθηκε ο συνδυασμός των βιοχημικών δεικτών εντοπίστηκαν τρεις διαφορετικοί φαινότυποι. Ο διαχωρισμός αυτός οφείλεται στις διακυμάνσεις που εμφανίζουν συγκεκριμένοι βιοχημικοί δείκτες μέσα στο δείγμα μας. Αυτές οι διακυμάνσεις αναπαρίστανται στο heatmap (Εικόνα 3.27) και οφείλονται στους δείκτες όπως η αιμοσφαιρίνη, ο αιματοκρίτης, ο αριθμός ουδετεροφίλων, τα λεμφοκύτταρα, τα ηωσινόφιλα, ο αριθμός λευκοκυττάρων, τα βασεόφιλα, τα ερυθρά αιμοσφαίρια, C-peptide, το εύρος κατανομής ερυθρών, η ουρία, η αλκαλική φωσφατάση, μέση περιεκτικότητα αιμοσφαιρίνης,

μέση συγκέντρωση αιμοσφαιρίνης, η mmp 7, η ινσουλίνη, η mmp 2, η RBP4, ο αριθμός ηωσινοφύλων, η ολική χοληστερόλη, η γαλεκτίνη-3, η LDL χοληστερόλη, η mmp9, η Fibrinogen, η mmp 1, η IL-1β και η γλυκόζη.

Πολλοί από τους παραπάνω δείκτες όπως η LDL χοληστερόλη, η ουρία, τα λεμφοκύτταρα, η γαλεκτίνη, η RBP4, η IL-6 και η μεταλλοπρωτεΐναση-1, έχουν συνδεθεί με τη νόσο της αθηρωμάτωσης όπως αναφέρθηκε στο κεφάλαιο 1 και παρουσιάστηκε συνοπτικά στον Πίνακα 1.2. Στη συνέχεια έγινε προσπάθεια για την ερμηνεία των φαινοτύπων μέσω των απεικονιστικών δεικτών και των συσχετίσεων που βρεθήκαν.

### **Ερμηνεία φαινοτύπων με χρήση απεικονιστικών δεικτών**

Στην πρώτη περίπτωση, η συστάδα 1,2,3 και 5 παρουσίασαν μεγάλη τιμή του δείκτη GSM κατά τη συστολή και τη διαστολή ενώ η συστάδα 4 και 6 παρουσίασαν μεσαία και χαμηλή τιμή του GSM αντίστοιχα. Όσο αφορά το εμβαδό της πλάκας, και οι έξι συστάδες κατηγοριοποιήθηκαν στην υψηλή κατηγορία όπως φαίνεται στον Πίνακα 3.3 και όσο αφορά τη συμπτωματολογία και οι έξι συστάδες κατατάχθηκαν στην ασυμπτωματική κατηγορία. Επιπλέον λαμβάνοντας υπόψιν τους πίνακες συσχετίσεων του δείκτη GSM με τους βιοχημικούς δείκτες, παρατηρούμε ότι θετικότερη συσχέτιση με το δείκτη GSM έχει ο αιματοκρίτης και αρνητικότερη έχουν τα τριγλυκερίδια (Πίνακας 3.4 και Πίνακας 3.5). Αυτό σημαίνει ότι οι συστάδες 1,2,3 και 5 έχουν αυξημένο αιματοκρίτη και χαμηλά επίπεδα τριγλυκεριδίων ενώ οι συστάδες 4 και 6 έχουν λίγο υψηλότερα επίπεδα όπως φαίνεται και από την Εικόνα 3.11. Τέλος, για το εμβαδό της πλάκας δεν βρέθηκε κάποια ισχυρή συσχέτιση σε αυτή την περίπτωση.

Στην δεύτερη περίπτωση, η συστάδα 1,2 και 3 παρουσίασαν μεγάλη τιμή του δείκτη GSM κατά τη συστολή και τη διαστολή. Όσο αφορά το εμβαδό της πλάκας, και οι συστάδες 2 και 3 κατηγοριοποιήθηκαν στην υψηλή κατηγορία ενώ η συστάδα 1 στη χαμηλή όπως φαίνεται στον Πίνακα 3.8 και όσο αφορά τη συμπτωματολογία και οι τρεις συστάδες κατατάχθηκαν στην ασυμπτωματική κατηγορία. Επιπλέον λαμβάνοντας υπόψιν τους πίνακες συσχετίσεων του δείκτη GSM με τους βιοχημικούς δείκτες, παρατηρούμε ότι θετικότερη συσχέτιση με το δείκτη GSM έχει η γαλεκτίνη-3 και αρνητικότερη έχουν η μεταλλοπρωτεΐναση-1. Αυτό σημαίνει ότι οι

συστάδες 1,2 και 3 χαρακτηρίζονται από υψηλά επίπεδα γαλεκτίνης και χαμηλά επίπεδα μεταλλοπρωτεΐνωσης-1. Για το εμβασό της πλάκας, σε αυτή την περίπτωση βρέθηκε αρνητική συσχέτιση με την RBP4 (Πίνακας 3.11) γεγονός που σημαίνει ότι οι συστάδες 2 και 3 χαρακτηρίζονται από χαμηλά επίπεδα RBP4 ενώ η συστάδα 1 που είχε χαμηλό δείκτη εμβασού πλάκας χαρακτηρίζεται από υψηλά επίπεδα RBP4 (αρνητική συσχέτιση).

Τέλος, στην τρίτη περίπτωση, η συστάδα 1 και 2 παρουσίασαν μεγάλη τιμή του δείκτη GSM κατά τη συστολή και τη διαστολή ενώ η συστάδα 3 χαμηλή τιμή του δείκτη. Όσο αφορά το εμβασό της πλάκας, και οι συστάδες 2 και 3 κατηγοριοποιήθηκαν στην υψηλή κατηγορία ενώ η συστάδα 1 στη χαμηλή όπως φαίνεται στον Πίνακα 3.14 και όσο αφορά τη συμπτωματολογία και οι τρεις συστάδες κατατάχθηκαν στην ασυμπτωματική κατηγορία. Επιπλέον λαμβάνοντας υπόψιν τους πίνακες συσχετίσεων του δείκτη GSM με τους βιοχημικούς δείκτες, παρατηρούμε ότι θετικότερη συσχέτιση με το δείκτη GSM έχει η γαλεκτίνη-3 και ο αιματοκρίτης και αρνητικότερη έχουν η μεταλλοπρωτεΐνωση-1 και τα τριγλυκερίδια. Αυτό σημαίνει ότι οι συστάδες 1 και 2 χαρακτηρίζονται από υψηλά επίπεδα γαλεκτίνης και αιματοκρίτη και χαμηλά επίπεδα μεταλλοπρωτεΐνωσης-1 και τριγλυκεριδίων. Ακριβώς το αντίστροφο ισχύει για τη συστάδα 3 που εμφάνισε χαμηλή τιμή του GSM. Για το εμβασό της πλάκας, σε αυτή την περίπτωση βρέθηκε θετική συσχέτιση με την μεταλλοπρωτεΐνωση-1 και αρνητική συσχέτιση με την RBP4 (Πίνακας 3.17) γεγονός που σημαίνει ότι οι συστάδες 2 και 3 χαρακτηρίζονται από χαμηλά επίπεδα RBP4 και υψηλά επίπεδα μεταλλοπρωτεΐνωσης-1, ενώ η συστάδα 1 που είχε χαμηλό δείκτη εμβασού πλάκας χαρακτηρίζεται από υψηλά επίπεδα RBP4 και χαμηλά επίπεδα μεταλλοπρωτεΐνωσης-1 (αρνητική συσχέτιση).

## 4.2 Επίλογος και μελλοντικές προεκτάσεις

Ο σχηματισμός αθηρωματικής πλάκας στην καρωτίδα μπορεί να προκαλέσει ποικίλες καρδιαγγειακές διαταραχές όπως εγκεφαλικό επεισόδιο ή παροδικό ισχαιμικό επεισόδιο (TIA-transient ischemic attack) [10]. Στα πλαίσια της παρούσας διπλωματικής εργασίας έγινε προσπάθεια διασύνδεσης των συμβατικών βιοχημικών δεικτών με τη νόσο της αθηρωμάτωσης και συσχέτισης με απεικονιστικούς δείκτες όπως ο δείκτης GSM, το εμβαδό της πλάκας και η συμπτωματολογία, χρησιμοποιώντας δεδομένα ασθενών που έπασχαν από καρωτιδική αθηρωμάτωση. Τα δεδομένα προήλθαν από έρευνες του Εργαστηρίου Βιοιατρικής Τεχνολογίας της σχολής Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών του Εθνικού Μετσόβιου Πολυτεχνείου. Αυτή η προσπάθεια πραγματοποιήθηκε με τη χρήση της ιεραρχικής μεθόδου συσταδοποίησης (Hierarchical clustering), η οποία αναλύθηκε εκτενώς, με στόχο την ομαδοποίηση των ασθενών με κοινά χαρακτηριστικά σε κατηγορίες.

Η έννοια της συσταδοποίησης είναι σημαντική σε πολλές επιστήμες όπως η στατιστική, η βιολογία και η βιοιατρική καθώς επίσης και σε πολλούς τομείς όπως η τεχνητή νοημοσύνη, η μηχανική μάθηση και η αναγνώριση προτύπων. Επιπλέον όσο αφορά την επεξεργασία δεδομένων, η συσταδοποίηση μπορεί να συμβάλει στην αποδοτικότερη διαχείριση των δεδομένων, στη μείωση των δεδομένων, στην παραγωγή υποθέσεων καθώς επίσης και στην πραγματοποίηση προβλέψεων βασιζόμενη στα χαρακτηριστικά των συστάδων. Τα τελευταία χρόνια, η χρήση μεθόδων συσταδοποίησης γίνεται ευρέως και αποτελεί αντικείμενο εντατικής έρευνας.

Όσο αφορά τη μείωση των δεδομένων, η συσταδοποίηση χρησιμοποιείται σαν ένα μέσο συμπίεσης δεδομένων σε συστάδες. Έτσι μπορεί η επεξεργασία των δεδομένων να γίνει με βάση τους αντιπροσώπους των συστάδων και όχι όλων των δεδομένων. Επιπλέον, η συσταδοποίηση μπορεί να οδηγήσει στην παραγωγή μιας υπόθεσης μετά από την επεξεργασία ενός συνόλου δεδομένων. Αυτή η υπόθεση θα μπορούσε να αφορά για παράδειγμα κάποια προτίμηση ενός συνόλου ανθρώπων.

Σημαντικότερο χαρακτηριστικό ωστόσο αποτελεί η δυνατότητα πραγματοποίησης μιας πρόβλεψης, η οποία βασίζεται στα χαρακτηριστικά των εξαγμένων συστάδων. Οι συστάδες χαρακτηρίζονται από τα χαρακτηριστικά των δεδομένων που περιέχουν και έτσι είναι δυνατόν

να ταξινομηθεί μια νέα παρατήρηση με βάση την ομοιότητα που παρουσιάζει με τα ήδη εξαγμένα χαρακτηριστικά των συστάδων.

Μια μελλοντική προέκταση της παρούσας εργασίας στο χώρο της βιοιατρικής θα μπορούσε να αποτελέσει η χρήση μεθόδου συσταδοποίησης σε ένα σύνολο δεδομένων που αφορούν ασθενείς που πάσχουν από την ίδια πάθηση, γνωρίζοντας δεδομένα όπως οι αγωγές που ακολούθησαν. Με αυτό τον τρόπο θα προκύπταν συστάδες βασισμένες στις αντιδράσεις των ασθενών στις αντίστοιχες φαρμακευτικές αγωγές και θα ήταν δυνατόν να μελετηθεί και να αναλυθεί η εξέλιξη των δεικτών των ασθενών στον άξονα του χρόνου για τις αντίστοιχες αγωγές και η αποδοτικότητα των αγωγών αυτών. Επιπλέον, ο συνδυασμός των βιοχημικών δεικτών με τεχνικές συσταδοποίησης μπορούν να αποτελέσουν ένα εργαλείο εκτίμησης της επικινδυνότητας διαφόρων νόσων.

Στην παρούσα διπλωματική εργασία διερευνήθηκε η συσχέτιση των βιοχημικών δεικτών με τη νόσο της καρωτιδικής αθηρωμάτωσης. Στα πλαίσια της μεθοδολογίας που αναπτύχθηκε στην παρούσα εργασία, θα μπορούσε να αποτελέσει μελλοντική προοπτική, η διερεύνηση της νόσου με άλλους, πιο εξελιγμένους απεικονιστικούς δείκτες όπως οι δείκτες υφής που προέρχονται από ανάλυση βασιζόμενη σε wavelet αλγορίθμους ή από δεδομένα προερχόμενα από Support Vector Machines [77]. Με την τεχνική της υφής τα τελικά χαρακτηριστικά προκύπτουν εξάγοντας το συχνοτικό περιεχόμενο των εικόνων. Με την τεχνική αυτή καταγράφηκε μεγάλη ακρίβεια ως προς την ταξινόμηση των δεδομένων, προοπτική που καθιστά τη διάγνωση και την αντιμετώπιση της ασταθούς πλάκας στην καρωτίδα πιο αποδοτική. Σε σύγκριση με τις παραδοσιακές στατιστικές μεθόδους, λαμβάνονται πληροφορίες σχετικά και με την συχνότητα και το χωρικό περιεχόμενο της εικόνας, λύνοντας με αυτόν τον τρόπο ένα σημαντικό πρόβλημα του χώρου της υγείας με την εκτίμηση της κατάστασης της καρωτίδας όχι μόνο από τον βαθμό στένωσης και τα συμπτώματα αλλά και από την γκρι διαβάθμιση της εικόνας (gray scale) [77]. Αντίστοιχα, θα μπορούσε με την μεθοδολογία της παρούσας εργασίας να διερευνηθούν εξελιγμένοι δείκτες κίνησης. Αυτό θα μπορούσε να επιτευχθεί μέσω της υποβοηθούμενων από υπολογιστή διαγνωστικών εργαλείων (CAD) [78]. Έρευνες έχουν δείξει πως ένα τέτοιο εργαλείο σε συνδυασμό με τους κατάλληλους δείκτες κίνησης είναι ικανό να διαχωρίσει συμπτωματικούς και ασυμπτωματικούς ασθενείς με ακρίβεια μεγαλύτερη από 88 %

γεγονός που θα συντελέσει ουσιαστικά στην έγκαιρη διάγνωση . Ο ορθός χαρακτηρισμός και η σωστή ερμηνεία των εξαγόμενων δεδομένων αποτελεί μεγάλη πρόκληση και θα επιφέρει σημαντική πρόοδο στο τομέα της Υγείας και της Βιοιατρικής Τεχνολογίας.

# Βιβλιογραφία

- [1] P. Castelnovo, L. Dallan, and M. Tschabitscher, "Surgical anatomy of the internal carotid artery", *An Atlas for Skull Base Surgeons*, Book, 2013.
- [2] F. H. Martini, "Blood vessels and circulation in fundamentals of anatomy and physiology", *Prentice Hall Inc*, pp. 717-777, 1995.
- [3] Blausen gallery 2014, *Wikiversity journal of medicine*, available on [www.en.wikiversity.org](http://www.en.wikiversity.org).
- [4] J. Bang, T. Dahl, A. Bruinsma, J. H. Kaspersen, T. A. Nagelhus Hernes, and H. O. Myhre, "A new method for analysis of motion of carotid plaques from RF ultrasound images", *Ultrasound in Medicine & Biology*, Vol.29, pp. 967-976, July 2003.
- [5] [www.hopkinsmedicine.com](http://www.hopkinsmedicine.com), accessed on 02/07/15.
- [6] A. J. Lusis, "Atherosclerosis", *Nature*, pp. 233-241, 14 September 2000.
- [7] H. C. Stary, A. B. Chandler, R. E. Dinsmore, V. Fuster, S. Glagov, W. Insull Jr, M. E. Rosenfeld, C. J. Schwartz, W. D. Wagner, and R. W. Wissler, "A definition of advanced types of atherosclerotic lesions and a histological classification of atherosclerosis", *Arteriosclerosis, Thrombosis, and Vascular Biology*, pp. 1512-1531, 1995.
- [8] H. Handa, M. Matsumoto, H. Maeda, H. Hougaku, and T. Kamada, "Ischemic stroke events and carotid atherosclerosis", *Stroke*, pp. 1781- 1786, 1995.
- [9] P. Libby, "Current concepts of the pathogenesis of the acute coronary syndromes", *Circulation*, vol. 104, pp. 365-372, 2001.
- [10] P. Sobieszczyk, and J. Beckman, "Carotid Artery Disease", *Circulation*, vol.114, pp.244-247.
- [11] [www.strokecenter.org/patients/about-stroke/ischemic-stroke/](http://www.strokecenter.org/patients/about-stroke/ischemic-stroke/), accessed on 03/07/2015.
- [12] <http://www.nhlbi.nih.gov/health/health-topics/topics/bdt/types>, accessed on 03/07/2015.

- [13] N. P. E. Kadoglou, T. Gerasimidis, A. Moumtzouoglou, A. Kapelouzou, N. Sailer, G. Fotiadis, I. Vitta, A. Katinios, P. Kougias, S. Bandios, K. Voliotis, P. E. Karayannacos, and C. D. Liapis, "Intensive lipid-lowering therapy ameliorates novel calcification markers and GSM score in patients with carotid stenosis", *European Journal of Vascular and Endovascular Surgery*, vol. 35, pp. 661-668, 2008.
- [14] A. Mizumaa, C. Kijimaa, K. Iijimab, Y. Gotob, K. Honmaa, T. Yasudab, K. Tokuokab, Y. Kitagawab, and S. Takizawaa, "Relationship between atherosclerotic risk factors and aortic plaques in patients with first-ever ischemic stroke", *Heart, Lung and Circulation*, vol. 23, pp.930-935, 2014.
- [15] M. Griffin, A. Frazer, A. Johnson, P. Collins, D. Owens, and GH. Tomkin, "Cellular cholesterol synthesis--the relationship to post-prandial glucose and insulin following weight loss", *Atherosclerosis*, vol. 138, pp.313-318, 1998.
- [16] I. M. Graham, L. E. Daly, H. M. Refsum, K. Robinson, L. E. Brattström, P. M Ueland, R. J. Palma-Reis, G. H. Boers, R. G. Sheahan, B. Israelsson, C. S. Uiterwaal, R. Meleady, D. McMaster, P. Verhoef, J. Witteman, P. Rubba, H. Bellet, J. C. Wautrecht, H. W. de Valk, L. Sales, F. M. Parrot-Rouland, K. S. Tan, I. Higgins, D. Garcon, and G. Andria, " Plasma homocysteine as a risk factor for vascular disease. The European Concerted Action Project", *Journal of the American Medical Association*, vol. 277, 1997.
- [17] M. J. Pletcher, K. Bibbins-Domingo, K. Liu, S. Sidney, F. Lin, E. Vittinghoff, and S. B. Hulley, "Nonoptimal lipids commonly present in young adults and coronary calcium later in life: The Cardia (Coronary Artery Risk Development in Young Adults) study", *Annals of Internal Medicine*, vol. 153, pp. 137-146, 2010.
- [18] A. Kini, J. D. Marmur, S. Kini, G. Dangas, T. P. Cocke, S. Wallenstein, E. Brown, J. A. Ambrose, and S. K. Sharma, "Creatine kinase-MB elevation after coronary intervention correlates with diffuse atherosclerosis, and low-to-medium level elevation has a benign clinical course", *Journal of the American College of Cardiology*, vol. 34, pp. 663-671, 1999.
- [19] C. Albert, P. R. Mertens, and P. Bartsch, "Urea and atherosclerosis—evidence for a direct link involving apolipoprotein B protein modifications", *International Urology and Nephrology*, vol. 43, pp. 933–936, 2011.
- [20] D. Bray, M. Clerici, and D. L. Trabattoni, "Markers for atherosclerosis 2009", Patent.



- [21] M. L. Grønholdt, H. Sillesen, B.M. Wiebe, H. Laursen, and B. G. Nordestgaard, "Increased acute phase reactants are associated with levels of lipoproteins and increased carotid plaque volume", *European Journal of Vascular and Endovascular Surgery*, vol. 21, pp. 227-234, 2001.
- [22] N. P. E. Kadoglou, G.S. Sfyroeras, A. Spathis, C. Gkekas, A. Gastounioti, G. Mantas, K. S. Nikita, P. Karakitsos, and C.D. Liapis, "Galectin-3, carotid plaque vulnerability, and potential effects of statin therapy", *European Journal of Vascular and Endovascular Surgery*, vol. 49, pp.4-9, 2015.
- [23] N. P. E. Kadoglou, V. Lambadiari, A. Gastounioti, C. Gkekas, T. G. Giannakopoulos, K. Koulia, E. Maratou, M. Alepaki, J. Kakisis, P. Karakitsos, K. S. Nikita, G. Dimitriadis, and C. D. Liapi, "The relationship of novel adipokines, RBP4 and omentin-1, with carotid atherosclerosis severity and vulnerability", *Atherosclerosis*, vol. 235, pp. 606-612, 2014.
- [24] H. Yamagami, K. Kitagawa, Y. Nagai, H. Hougaku, M. Sakaguchi, K. Kuwabara, K. Kondo, T. Masuyama, M. Matsumoto, and M. Hori, "Higher levels of interleukin-6 are associated with lower echogenicity of carotid artery plaques", *American Heart Association*, vol. 35, pp. 677-681, 2004.
- [25] J. Sundström, J. C. Evans, E. J. Benjamin, D. Levy, M. G. Larson, D. B. Sawyer, D. A. Siwik, W. S. Colucci, P. W.F. Wilson, and R. S. Vasan, "Relations of plasma total TIMP-1 levels to cardiovascular risk factors and echocardiographic measures: the framingham heart study", pp. 1509-1516, 2004.
- [26] N. Rodondi, P. Marques-Vidal, J. Butler, K. Sutton-Tyrrell, J. Cornuz, S. Satterfield, T. Harris, D. C. Bauer, L. Ferrucci, E. Vittinghoff, and A. B. Newman, "Markers of atherosclerosis and inflammation for prediction of coronary heart disease in older adults", *American Journal of Epidemiology*, vol. 171, pp. 540-549, 2010.
- [27] E.C. Kyriacou, C. Pattichis, M. Pattichis, C. Loizou, C. Christodoulou, S.K. Kakkos, A. Ncolaidis, "A review of noninvasive ultrasound image processing methods in the analysis of carotid plaque morphology for the assessment of stroke risk", *IEEE Transactions On Information Technology In Biomedicine.*, vol. 14 (4), pp. 1027-1038, July 2010.
- [28] S. Aly, and C. C. Bishop, "An objective characterization of atherosclerotic lesion-an alternative method to identify unstable plaque", pp. 1921-1924, *Stroke*, August 2000.
- [29] B. Ibanez, J.J. Badimon, J.M. Garcia, "Diagnosis of atherosclerosis by imaging", *American Journal of Medicine*, S15-S25, 2009.

[30] S. Golemati, A. Gastounioti, and K.S. Nikita, "Toward novel noninvasive and low-cost markers for predicting strokes in asymptomatic carotid atherosclerosis: the role of ultrasound image analysis", *IEEE Transactions on Biomedical Engineering*, vol.60, pp. 652-658, March 2013.

[31] [www.hopkinsmedicine.org](http://www.hopkinsmedicine.org), accessed on 02/07/2015.

[32] D. Hykes, W. R. Hedrick, and D.E. Starchman, "Ultrasound physics and instrumentation", *Churchill Livingstone Inc*, 1985.

[33] S. Hairong, T. Varghese, R.S. Dempsey, M.S. Salamat, and J.A. Zagzebski, "Relationship between ultrasonic attenuation, size and axial strain parameters for ex-vivo atherosclerotic carotid plaque", *Ultrasound in Medicine & Biology*, Vol.34, pp. 1666-1677, 2008.

[32] F. Rakebrandt, D.C. Crawford, D. Havard, D.Coleman, and S.D. Woodcock, "Relationship between ultrasound texture classification images and histology of atherosclerotic plaque", *Ultrasound in Medicine & Biology*, Vol.26, pp. 1393-1402, 2000.

[35] C. L. Korte, H. H. G. Hansen, and A. F.W. Van der Steen, "Vascular ultrasound for atherosclerosis imaging", *Interface Focus*, vol. 1, pp. 565-575, August 2011.

[36] Δ. Κουτσούρης, Κ. Νικήτα, Σ. Παυλόπουλος, *Ιατρικά Απεικονιστικά Συστήματα*, 2004.

[37] [www.ultrasoundpaedia.com/normal-carotids/](http://www.ultrasoundpaedia.com/normal-carotids/), accessed on 05/07/2015.

[38] S. Golemati, J. Stoitsis, and K.S. Nikita, "Motion analysis of the carotid artery wall and plaque using B-mode ultrasound", *Vascular Disease Prevention*, vol.4, pp. 296-302, November 2007.

[39] L. Saba, G. Mallarini, R. Sanfilippo, G. Zeng, R. Montisci, and J. Suri, " Intima media thickness variability (IMTV) and its association with cerebrovascular events: a novel marker of carotid atherosclerosis?", *Cardiovascular Diagnosis and Therapy*, vol.2, pp. 10-18, March 2012.

[40] S. C. Kofoed, H.H. Wittrup, H. Sillesen, and B. G. Nordestgaard, "Fibrinogen predicts ischaemic stroke and advanced atherosclerosis but not echolucent, rupture-prone carotid plaques", *European Heart Journal*, vol. 24, pp. 567-576, 2003.

[41] G. Zahnd, D. Vray, A. Sérusclat, D. Alibay, M. Bartold, A. Brown, M. Durand , L. M. Jamieson, K. Kapellas, L. J. Maple-Brown, K. O'Dea, P. Moulin, D. S. Celermajer, and M. R. Skilton, "Longitudinal displacement of the carotid wall and cardiovascular risk factors: Associations with

aging, adiposity, blood pressure and periodontal disease independent of cross-sectional distensibility and intima-media thickness", *Ultrasound in Medicine & Biology*, vol. 38, pp. 1705-1715, 2012.

[42] A. Gastouniotti, S. Golemati, J. S. Stoitsis, and K. S. Nikita, "Carotid artery wall motion analysis from B-mode ultrasound using adaptive block matching: in silico evaluation and in vivo application", *Physics in medicine and biology*, vol. 58, pp. 8647-8661, 2013.

[43] P. Maragos, "*Image Analysis and Computer Vision*", Lecture Notes, National Technical University of Athens, 2005.

[44] N. El-Barghouty, G. Geroulakos, A. Nicolaidis, A. Androulakis, V. Bahal, "Computer-assisted carotid plaque characterisation", *European Journal Of Vascular and Endovascular Surgery*, vol. 9, pp. 389-393, May 1995.

[45] P. Asvestas, S. Golemati, G. K. Matsopoulos, K. S. Nikita, A. Nicolaidis, "Fractal dimension estimation of carotid atherosclerotic plaques from B-mode ultrasound: A pilot study," *Ultrasound In Medicine and Biology*, vol. 28, pp. 1129–1136, September 2002.

[46] A. M. Raid, W. M. Khedr, M. A. El-dosuky, and M. Aoud, " Image restoration based on morphological operations", *International Journal of Computer Science, Engineering and Information Technology (IJCEIT)*, Vol. 4 (3), June 2014.

[47] A.B. Ashraf, D. Daye, S. Gavenosis, C. Mies, M. Feldman, M. Rosen, and D. Kontos, "Identification of Intrinsic Imaging Phenotypes for Breast Cancer Tumors: Preliminary associations with gene expression profiles", *Radiological Society of North America*, vol.272, pp. 374-384, August 2014.

[48] H. C. Romesburg, "Cluster Analysis for Researchers", 2004.

[49] B. S. Duran, and P. L. Odell, "Cluster analysis: A survey", Lecture Notes In Economics and Mathematical Systems, 2013.

[50] D. H. Ye, K. M. Pohl, and C. Davatzikos, "Semi-supervised pattern classification: Application to structural MRI of Alzheimer's disease", *IEEE International Workshop on Pattern Recognition in NeuroImaging*, PP. 1-4, 2011.

- [51] T. Hastie, R. Tibshirani, and J. Frieman, "Unsupervised Learning", *The elements of statistical learning*, Springer Series in Statistics, pp. 485-585, 2009.
- [52] W. Voyt, and D. Nagel, "Cluster analysis in diagnosis", *Clinical Chemistry*, vol. 38, pp. 182-198, 1992.
- [53] C. D. Manning , P. Raghavan, H. Schütz, "An introduction to information retrieval" *Cambridge University Press*.
- [54] A. V. Lukashin, and R. Fuchs, " Analysis of temporal gene expression profiles: clustering by simulated annealing and determining the optimal number of clusters", *Bioinformatics*, vol.17, pp. 405-414, 2001.
- [55] R. Tibshirani, G. Walther, and T. Hastie, "Estimating the number of clusters in a data set via the gap statistic", *Journal of the Royal Statistical Society*, vol.63, pp. 411-423, 2001.
- [54] G. Carlsson, and F. Memoli, " Characterization, stability and convergence of hierarchical clustering methods", *Journal of Machine Learning Research*, vol.11, pp. 1425-1470, 2010.
- [57] M. A. Khamsi, and W. A. Kirk, " §1.4 The triangle inequality in  $\mathbb{R}$ ", *An introduction to metric spaces and fixed point theory*, 2001.
- [58] "The distance procedure: proximity measures", *Users Guide, SAS Institute*, version 9.2, last updated April 2010.
- [59] C. Cheadle, M. P. Vawter, W. J. Freed, and Kevin G. Becker, "Analysis of microarray data using z-score transformation", *Journal of Molecular Diagnostics*, vol.5, pp. 73-81, 2003.
- [60] J. Hauke, and T. Kossowski, "Comparison of values of pearson's and spearman's correlation coefficients on the same sets of data", *Quaestiones Geographicae*, vol.30, pp. 87-93, 2011.
- [61] R. D. Maesschalck, D. J Rimbaud, and D.L. Massart, "The Mahalanobis distance", *Chemometrics and Intelligent Laboratory Systems*, vol.10, pp. 1-18, 2000.
- [62] E. F. Krause, "Taxicab geometry: An adventure in non-euclidean geometry", 1975.
- [63] J. M. Merigó, and M. Casanovas, "A new minkowski distance based on induced aggregation operators", *International Journal of Computational Intelligence Systems*, vol.4, 2011.

[64] F. Heijden, R. Duin, D. Ridder, and D. M. J. Tax, "Classification, parameter estimation and state estimation: an engineering", 2005.

[65] C. Spearman, "The proof and measurement of association between two things", *The American Journal of Psychology*, vol. 15, pp. 72-101, 1904.

[66] [www.mines.humanoriented.com/classes/2010/fall/csci568/portfolio\\_exports/sdaugherty/similarity.html](http://www.mines.humanoriented.com/classes/2010/fall/csci568/portfolio_exports/sdaugherty/similarity.html), accessed on 07/07/2015.

[67] A. Meyer, A. A. F. Garcia, A. P. de Souza, and C. L. de Souza, "Comparison of similarity coefficients used for cluster analysis with dominant markers in maize (*Zea mays* L)", *Genetics and Molecular Biology*, vol.27, pp. 83-91, 2004.

[68] J. M. Duarte, J. B. Santos, and L. C. Melo, "Comparison of similarity coefficients based on rapid markers in the common bean", *Genetics and Molecular Biology*, vol.22, pp. 427-432, 1999.

[69] NCSS Statistical Software, "Hierarchical clustering / Dendrograms", Chapter 445, available on [www.NCSS.com](http://www.NCSS.com).

[70] B. Mirkin, "Mathematical classification and clustering: from how to what and why", *Classification, Data Analysis, and Data Highways*, pp. 172-181, 1998.

[71] J. C. Gower, and G. J. S. Ross, "Minimum spanning trees and single linkage cluster analysis", *Journal of the Royal Statistical Society*, vol.18, pp. 54-64, 1969.

[72] J. Quackenbush, "Computational analysis of microarray data," *Nature Reviews Genetics*, vol. 2, pp. 418– 427, 2001.

[73] <http://nlp.stanford.edu>, accessed on 18/07/2015.

[74] J. H. Ward, "Hierarchical Grouping to Optimize an Objective Function", *Journal of the American Statistical Association*, vol.58, pp. 236-244, 1963.

[75] A. Gastouniotti, V. Koliass, S. Golemati, N.N. Tsiaparas, A. Matsakou, J.S. Stoitsis, N.P.E. Kadoglou, C. Gkekas, J.D. Kakisis, C.D. Liapis, P. Karakitsos, I. Sarafis, P. Angelidis, and K.S. Nikita, "Carotid – a web-based platform for optimal personalized management of atherosclerotic patients," *Comput Meth Programs Biomed*, vol. 114, pp. 183-193, 2014.

[76] N. Tsiaparas, S. Golemati, I. Andreadis, J. S. Stoitsis, I. Valavanis, and K. S. Nikita, "Comparison of multiresolution features for texture classification of carotid atherosclerosis from B-Mode ultrasound", *IEEE Transactions On Information Technology In Biomedicine*, vol. 15, pp. 130-137, 2011.

[77] N. Tsiaparas, S. Golemati, I. Andreadis, J. Stoitsis, I. Valavanis, and K. S. Nikita, "Assessment of carotid atherosclerosis from B-mode ultrasound images using directional multiscale texture features", *Measurement Science and Technology*, vol. 23, pp. 1-10, 2012.

[78] A. Gastounioti, S. Makrodimitris, S. Golemati, N. Kadoglou, C. Liapis, and K. S. Nikita, "A novel computerized tool to stratify risk in carotid atherosclerosis using kinematic features of the arterial wall", *IEEE Journal of Biomedical and Health Informatics*, vol. 19, 2015.