



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

**Ανάπτυξη Ευφρούς Συστήματος Υποστήριξης Κλινικής
Απόφασης για την Έγκαιρη Διάγνωση και την
Εξατομικευμένη Διαχείριση Γυναικών που βρίσκονται σε
κίνδυνο ανάπτυξης Καρκίνου του Μαστού**

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

Της

Μαριλένας Ταρούση,

Διπλωματούχου Ηλεκτρολόγου Μηχανικού και Μηχανικού Υπολογιστών Ε.Μ.Π.

Αθήνα, Ιούλιος 2022



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

Ανάπτυξη Ευφυούς Συστήματος Υποστήριξης Κλινικής Απόφασης για την Έγκαιρη Διάγνωση και την Εξατομικευμένη Διαχείριση Γυναικών που βρίσκονται σε κίνδυνο ανάπτυξης Καρκίνου του Μαστού

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

Της

Ταρούση Μαριλένας

Διπλωματούχου Ηλεκτρολόγου Μηχανικού και Μηχανικού Υπολογιστών, Ε.Μ.Π

Συμβουλευτική επιτροπή: Δημήτριος-Διονύσιος Κουτσούρης (Επιβλέπων)

Κωνσταντίνα Νικήτα

Γεώργιος Ματσόπουλος

Εγκρίθηκε από την επταμελή εξεταστική επιτροπή την 12^η Ιουλίου 2022.

Δημήτριος-Διονύσιος Κουτσούρης
Καθηγητής Ε.Μ.Π.

Κωνσταντίνα Νικήτα
Καθηγήτρια Ε.Μ.Π.

Γεώργιος Ματσόπουλος
Καθηγητής Ε.Μ.Π.

Γεώργιος Ζωγράφος
Καθηγητής Ε.Κ.Π.Α

Μανούσος Κωνσταντουλάκης
Καθηγητής Ε.Κ.Π.Α

Μαρία Χαρίτου
Ερευνήτρια Α', Ε.Μ.Π.

Θεμιστοκλής Έξαρχος
Επ. Καθηγητής Ιόνιο Πανεπιστήμιο



.....
Ταρούση Μαριλένα

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

Copyright © Μαριλένα Ταρούση, 2022

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

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

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

Περίληψη

Ο πληθυσμιακός έλεγχος με τη χρήση μαστογραφίας έχει μειώσει δραματικά τα ποσοστά καρκίνου του μαστού παγκοσμίως. Σε πολλές χώρες η πρόληψη του καρκίνου του μαστού βασίζεται σε συχνές και επαναλαμβανόμενες μαστογραφίες, ακολουθούμενες από υπερηχογράφημα μαστών και, εάν είναι απαραίτητο, από ιστολογική εξέταση στο βιολογικό υλικό της βιοψίας. Όπως έχει διαπιστωθεί από τη μελέτη της διεθνούς βιβλιογραφίας, η μαστογραφία είναι μια εξέταση με σχετικά χαμηλή διαγνωστική απόδοση, τόσο σε επίπεδα ευαισθησίας όσο και σε επίπεδα ειδικότητας. Είναι γεγονός ότι η ερμηνεία των μαστογραφικών ευρημάτων επηρεάζεται από υποκειμενικούς παράγοντες και είναι επιρρεπής σε διαγνωστικά σφάλματα, όπως αποδεικνύεται στην παρούσα μελέτη. Η μαγνητική τομογραφία μαστού, είναι μια διαγνωστική πρακτική που, παρ' ότι θεωρείται υψηλής διαγνωστικής ακρίβειας, ενδείκνυται σε περιπτώσεις υψηλής πυκνότητας μαστού καθώς πρόκειται για μια εξέταση υψηλού κόστους για τα συστήματα υγειονομικής περίθαλψης. Επιπλέον, ο γενετικός έλεγχος χρησιμοποιείται για τη διάγνωση του κληρονομικού καρκίνου του μαστού, ο οποίος όχι μόνο αντιπροσωπεύει ένα μικρό ποσοστό των περιπτώσεων (5 - 10% του συνόλου των καρκίνων μαστού), αλλά αποτελεί και μια εξειδικευμένη εξέταση υψηλού κόστους. Ακόμα και η ευρέως χρησιμοποιούμενη βιοψία με λεπτή βελόνα αναρρόφησης (Fine Needle Aspiration - FNA), σε ορισμένες περιπτώσεις ενέχει κινδύνους και παράγει ψευδώς αρνητικά αποτελέσματα. Λαμβάνοντας υπόψη τα προαναφερθέντα στοιχεία, η παρούσα διδακτορική διατριβή παρουσιάζει την ανάπτυξη μοντέλων βασισμένων στην Τεχνητή Νοημοσύνη και τη δημιουργία ενός Συστήματος υποστήριξης κλινικής απόφασης (Clinical Decision Support System - CDSS), με σκοπό την έγκαιρη διάγνωση και την εξατομικευμένη διαχείριση των γυναικών που βρίσκονται σε κίνδυνο ανάπτυξης καρκίνου του μαστού. Η ανάπτυξη των μοντέλων αυτών βασίζεται σε μη επεμβατικά ιατρικά δεδομένα που συλλέγονται κατά την καθημερινή κλινική πρακτική στο Κέντρο Μαστού του Γενικού Νοσοκομείου «Ελενα Βενιζέλου», καθώς και σε δεδομένα του προσωπικού ιατρικού φακέλου των γυναικών του Κέντρου Μαστού. Ως εκ τούτου, παρουσιάζουμε ένα εργαλείο, το οποίο μέσω του συνδυασμού των αποτελεσμάτων μη επεμβατικών και χαμηλού κόστους εξετάσεων και του ιατρικού φακέλου υγείας των γυναικών, θα είναι εύκολο στην καθημερινή χρήση του από τους κλινικούς ιατρούς και ταυτόχρονα θα είναι οικονομικά αποδοτικό για τα συστήματα υγείας. Χρησιμοποιώντας τις τεχνικές επιλογής χαρακτηριστικών και τα εργαλεία της μηχανικής μάθησης, διαπιστώσαμε ότι ο ταξινομητής Random Forest (RF) έδειξε τις υψηλότερες τιμές στα μέτρα της ακρίβειας (95,9%), ευαισθησίας (96,2%), NPV (95%), Youden's index (0,92) και F1 score (0,97), αποδεικνύοντας την καταλληλότητά του για τη δημιουργία του καινοτόμου CDSS καθώς προσφέρει εξαιρετικά χαμηλά ποσοστά ψευδώς αρνητικών και ψευδώς θετικών αποτελεσμάτων και εντοπίζει με πολύ μεγάλη επιτυχία τα πραγματικά θετικά περιστατικά (ασθενείς). Το προτεινόμενο CDSS στοχεύει να αποτελέσει μια βιώσιμη και οικονομικά αποδοτική λύση για τα εθνικά συστήματα υγειονομικής περίθαλψης αφού μπορεί να συμβάλει σε σημαντική μείωση του κόστους όσον

αφορά στη διαχείριση των γυναικών που κινδυνεύουν να αναπτύξουν καρκίνο του μαστού, επιτυγχάνοντας παράλληλα αξιοσημείωτη αύξηση της εγκυρότητας της διάγνωσης.

Λέξεις Κλειδιά: Καρκίνος του μαστού, Μαστογραφία, Υπέρηχος μαστών, Ατομικός ιατρικός φάκελος υγείας, Εκτίμηση κινδύνου, Συστήματα υποστήριξης κλινικής απόφασης, Εξατομικευμένη ιατρική, Μηχανική μάθηση, Μπεϋζιανά δίκτυα, Τυχαία δάση

Abstract

Women population screening using mammography has dramatically reduced breast cancer rates worldwide. Nowadays, in many countries, the prevention of breast cancer policy is based on frequent and repeated mammographies, followed by breast ultrasound and, if necessary, by histological examination in the biological material of the biopsy. However, evaluating mammography findings is considered as a difficult process which can be properly performed only by a highly experienced and well-trained medical staff. Subsequently, the interpretation of the findings could be easily influenced by subjective factors and therefore can be prone to diagnostic errors as evidenced in the present study. Breast MRI (Magnetic Resonance Imaging), is a diagnostic practice indicated in cases of high breast density and at the same time, a high-cost examination for healthcare systems. Furthermore, genetic testing that is used to diagnose hereditary breast cancer, represents a small proportion of breast cancers and at the same time it is a high-cost specialized test. Even the most widely used biopsy, Fine Needle Aspirate (FNA), in some cases involves risks and produces false negative results. Considering the aforementioned facts, this study presents the development of AI (Artificial Intelligence) - based models and the creation of a Clinical Decision Support System (CDSS) for the early diagnosis and personalized management of women at risk of developing breast cancer. These models are based on non-invasive medical data collected during daily clinical practice at the Breast Center of the General Hospital "Elena Venizelou", as well as on data from women's personal medical health records. Therefore, we demonstrate a tool, which, by combining the results of non-invasive and low-cost diagnostic tests and the medical health record, will be easy to use by clinicians on a daily basis and at the same time cost-effective for the health systems. Using feature selection techniques and machine learning tools, we found that the Random Forest (RF) classifier showed the highest values of accuracy (95.9%), sensitivity (96.5 %), NPV (95%), Youden's index (0.92) και F1 score (0.97), proving to be an effective algorithm in the development of our innovative CDSS. The proposed CDSS aims to be a sustainable and cost-effective solution for national healthcare systems as it can contribute to a significant cost reduction in terms of personalized management of women at risk of developing breast cancer, while achieving a remarkable increase in the validity of diagnosis.

Keywords: Breast cancer, Mammography, Breast ultrasound, Medical health record, Risk assessment, Clinical decision support systems, Personalized medicine, Machine learning, Bayesian networks, Random forests

Η παρούσα Διδακτορική Διατριβή είναι αφιερωμένη στις γυναίκες της
οικογένειάς μου

Ευχαριστίες

Πρωτίστως, θα ήθελα να εκφράσω τις θερμές ευχαριστίες μου στον επιβλέποντα μου, Καθηγητή Ε.Μ.Π. Δημήτρη Κουτσούρη. Τον ευχαριστώ για την τόσο σημαντική ευκαιρία που μου έδωσε να πραγματοποιήσω το όνειρο μου, για τις πάντα εύστοχες συμβουλές και τη διαρκή υποστήριξή του καθ' όλη τη διάρκεια εκπόνησης της διατριβής μου. Ο κ.Κουτσούρης είναι ένας μεγάλος δάσκαλος για όλους όσους έχουμε τη χαρά και τιμή να τον γνωρίζουμε.

Θα ήθελα, επίσης, να εκφράσω τη βαθιά ευγνωμοσύνη μου για τον άνθρωπο που δε σταμάτησε να είναι δίπλα μου σε όλο αυτό το δύσκολο ταξίδι, τον μέντορά μου, Δρ. Παναγιώτη Μπούντρη. Θα θυμάμαι πάντα την αγωνία με την οποία του εξέθετα τους προβληματισμούς μου και την επιστημονική και ψυχολογική του υποστήριξη που με έκαναν κάθε φορά να συνεχίζω και τελικά να φτάσω στην ολοκλήρωση του έργου μου.

Επιπλέον, θα ήθελα να ευχαριστήσω όλο το προσωπικό του Κέντρου Μαστού του Γενικού Νοσοκομείου Αττικής «Έλενα Βενιζέλου» οι οποίοι με τη συνεργασία τους συνέβαλαν τα μέγιστα στην επιτυχή ολοκλήρωση της παρούσας διατριβής. Ιδιαίτερα ευχαριστώ τον κ. Παναγιώτη Δασκαλάκη, Διευθυντή ΕΣΥ και Επιστημονικό Υπεύθυνο της Διατομεακής Μονάδας Μαστού του Χειρουργικού τμήματος για τη συνεργασία και καθοδήγηση του.

Θερμές ευχαριστίες οφείλω επίσης στους Καθηγητές Ε.Μ.Π., Γιώργο Ματσόπουλο και Κωνσταντίνα Νικήτα για τη βοήθεια που μου πρόσφεραν σε ότι χρειάστηκα καθώς και για τη συμμετοχή τους στην Τριμελή Συμβουλευτική Επιτροπή της διατριβής μου. Θα ήθελα επίσης να ευχαριστήσω θερμά τους Καθηγητές του Ε.Κ.Π.Α. Γεώργιο Ζωγράφο και Μανούσο Κωνσταντουλάκη και τον Επίκουρο Καθηγητή του Ιονίου Πανεπιστημίου Θεμιστοκλή Έξαρχο, για τη συμμετοχή τους στην Επταμελή Συμβουλευτική Επιτροπή της διατριβής μου. Ιδιαίτερα ευχαριστώ την Ερευνήτρια Α΄ Δρ. Μαρία Χαρίτου, μέλος της Επταμελούς Συμβουλευτικής Επιτροπής, για την πολύτιμη βοήθεια και τη διαρκή υποστήριξη, επιστημονική και ψυχολογική, που έλαβα από αυτήν.

Ευχαριστώ πολύ επίσης όλους τους συναδέλφους και μέλη του Εργαστηρίου Βιοϊατρικής Τεχνολογίας του Ε.Μ.Π. για τη συνεργασία μας όλα αυτά τα χρόνια και τη συνεχή τους υποστήριξη και συμπαράσταση. Επιπλέον, θα ήθελα να ευχαριστήσω από την καρδιά μου τη Δρ. Ράνια Πετροπούλου που ήταν κοντά μου και με υποστήριζε διαρκώς σε όλη την πορεία μου, όπως επίσης και τον Δρ. Γιάννη Κουρή που ήταν για μένα τόσο επιστημονικό όσο και ψυχολογικό στήριγμα.

Θα ήθελα επίσης να ευχαριστήσω τους ανθρώπους που όντας συνοδοιπόροι μου όλο αυτό τον καιρό, μου έδωσαν όλη τους την αγάπη και την στήριξη για να φτάσω στο τέλος αυτής της διαδρομής. Ελπίζω να

καταφέρω να τους σταθώ όπως εκείνες και εκείνοι για μένα, γερό στήριγμα και ζεστή αγκαλιά μαζί. Ιδιαίτερα ευχαριστώ την Ολυμπία και τη Θέλμα που άκουγαν πάντα με υπομονή και αγάπη τους πιο μικρούς και πιο μεγάλους προβληματισμούς μου. Πολλά ευχαριστώ οφείλω επίσης στον άνθρωπο που στάθηκε στο πλάι μου σε κάθε βήμα, τον Πάνο, που με το μυαλό και την καρδιά του με συντρόφευσε σε όλη την πορεία μου.

Στη συνέχεια, θα ήθελα να εκφράσω τη βαθιά ευγνωμοσύνη μου στην οικογένεια μου. Η αγάπη μου για αυτήν μου έδωσε το έναυσμα να ασχοληθώ με όλα αυτά που πραγματεύεται η παρούσα διατριβή. Οι γυναίκες της οικογένειάς μου, στις οποίες και αφιερώνω όλη αυτή την προσπάθεια, μου έχουν μάθει πως η αγάπη για τη ζωή δε νικιέται με κανένα τρόπο.

Τέλος, χωρίς την Άννα μου, τη Δήμητρα και το Θόδωρα, δε θα είχα καταφέρει να ξεπεράσω όλα τα εμπόδια και τις δυσκολίες που συνάντησα, αυτοί είναι η ασπίδα και το καταφύγιο μου, και θα τους ευχαριστώ για πάντα.

Πίνακας Περιεχομένων

Περίληψη	5
Abstract	7
Ευχαριστίες	11
Κατάλογος Εικόνων	17
Κατάλογος Πινάκων	20
Συντομογραφίες	22
Κεφάλαιο 1°	25
Εισαγωγή	25
Κεφάλαιο 2°	28
Ο Καρκίνος του Μαστού	28
2.1 Επιδημιολογικά Δεδομένα για τον Καρκίνο του Μαστού	28
2.1.1 Στοιχεία σε παγκόσμιο επίπεδο.....	29
2.1.2 Ευρώπη	32
2.1.3 Ελλάδα	33
2.2 Ανατομία και Φυσιολογία του Μαστού.....	35
2.2.1 Η Ανατομία του Μαστού	36
2.2.2 Η Φυσιολογία και η Ανάπτυξη του Μαστού	38
2.3 Η Παθολογία του Μαστού	39
2.3.1 Καλοήθεις Παθήσεις Μαστού.....	39
2.3.2 Καλοήθεις και Κακοήθεις Αποτιτανώσεις.....	43
2.3.3 Ο Καρκίνος του Μαστού	44
2.4 Κατηγοριοποίηση και Τύποι του Καρκίνου του Μαστού.....	44
2.4.1 Ιστολογική Ταξινόμηση του Καρκίνου του Μαστού.....	44
2.4.2 Βαθμός Κακοήθειας ή Διαφοροποίησης – Tumor Grade	46
2.4.3 Υπότυποι και Μοριακή Ταξινόμηση του Καρκίνου του Μαστού	47
2.4.4 Σταδιοποίηση του Καρκίνου του Μαστού.....	48
2.5 Συμπτώματα του Καρκίνου του Μαστού.....	50
2.6 Αιτιολογία του Καρκίνου του Μαστού και Παράγοντες Κινδύνου	52
Ηλικία και αναπαραγωγικοί παράγοντες	52
Φυλετικός Παράγοντας.....	59

Δείκτης Μάζας Σώματος και Παχυσαρκία	60
Πυκνότητα μαστών	61
Μέγεθος Στήθους.....	62
Οικογενειακό ιστορικό.....	63
Γενετικοί παράγοντες.....	66
Κάπνισμα	67
Υπέρταση.....	68
Υψηλή χοληστερίνη.....	68
Διαταραχές στο θυρεοειδή αδένα	69
2.7 Θεραπεία του Καρκίνου του Μαστού.....	71
2.7.1 Τρόποι Αντιμετώπισης του Καρκίνου του Μαστού	72
2.7.2 Θεραπευτικό Σχήμα ανάλογα με το στάδιο της νόσου	74
Κεφάλαιο 3 ^ο	76
Διαγνωστικές μέθοδοι για την Ανίχνευση του Καρκίνου του Μαστού – Ανασκόπηση των Τεχνικών και της Απόδοσής τους	76
3.1 Διαγνωστικές Μέθοδοι για την Ανίχνευση του Καρκίνου του Μαστού.....	76
3.1.1 Απεικονιστικά Διαγνωστικά Μέσα.....	78
3.1.2 Βιοψία Μαστού.....	87
3.1.3 Γενετική Εξέταση	91
3.1.4 Κλινική Εξέταση και Αυτοεξέταση Μαστού.....	92
3.2 Ανασκόπηση και Σύγκριση της Απόδοσης των Διαγνωστικών Εξετάσεων για τον Καρκίνου του Μαστού	93
3.2.1 Ποιοτική Εκτίμηση της Διαγνωστικής Απόδοσης των BSE και CBE.....	94
3.2.2 Η Διαγνωστική Απόδοση της Μαστογραφίας	95
3.2.3 Η Διαγνωστική Απόδοση του Υπερήχου Μαστών	97
3.2.4 Η Διαγνωστική Απόδοση της Μαγνητικής Τομογραφίας Μαστών.....	98
3.2.5 Η Διαγνωστική Απόδοση της Βιοψίας Μαστών.....	99
3.2.6 Η Διαγνωστική Απόδοση του Συνδυασμού των Εξετάσεων	100
3.2.7 Πίνακες Ανασκόπησης της Απόδοσης των Διαγνωστικών Δοκιμασιών και των Συνδυασμών τους	102
3.3 Συζήτηση - Συμπεράσματα	105
Κεφάλαιο 4 ^ο	108
Στατιστικά Μοντέλα Εκτίμησης Κινδύνου του Καρκίνου του Μαστού και Ευφυή Συστήματα Υποστήριξης Κλινικής Απόφασης.....	108

4.1 Στατιστικά Μοντέλα Εκτίμησης Κινδύνου.....	108
Μοντέλα που προβλέπουν τον Κίνδυνο Ανάπτυξης Καρκίνου του Μαστού	108
Μοντέλα πρόβλεψης Γενετικών Μεταλλάξεων.....	114
Μοντέλα που προβλέπουν τον Κίνδυνο Ανάπτυξης Καρκίνου του Μαστού και τις Γενετικές Μεταλλάξεις.....	117
4.2 Ευφυή Συστήματα Υποστήριξης Κλινικής Απόφασης για τη Διάγνωση και την Κλινική Διαχείριση του Καρκίνου του Μαστού	120
4.2.1 Ιατρικά Δεδομένα.....	120
4.2.2 Ευφυή Συστήματα Υποστήριξης Κλινικής Απόφασης για τη Διάγνωση και την Κλινική Διαχείριση του Καρκίνου του Μαστού.....	122
4.3 Συζήτηση – Συμπεράσματα	125
Κεφάλαιο 5°	132
Παρουσίαση του Ιατρικού Προβλήματος.....	132
5.1 Ιατρικά Δεδομένα: Περιγραφή και Οργάνωση.....	132
5.2 Προεπεξεργασία των Ιατρικών Δεδομένων.....	136
5.3 Παρουσίαση του Ιατρικού Προβλήματος.....	143
5.4 Διαχείριση Περιστατικών με Ελλιπή Δεδομένα	148
Κεφάλαιο 6°	150
Αξιολόγηση και Κατάταξη των Χαρακτηριστικών του Διανύσματος Εισόδου με χρήση Μεθόδων Επιλογής Χαρακτηριστικών και Ανάπτυξη των Μοντέλων Πρόβλεψης/Ταξινόμησης	150
6.1 Επιλογή Χαρακτηριστικών στη Μηχανική Μάθηση	150
6.1.1 Επιβλεπόμενη Επιλογή Χαρακτηριστικών (Supervised Feature Selection)	155
6.2 Τεχνικές Επιλογής Χαρακτηριστικών από το σύνολο των 35 Χαρακτηριστικών του Διανύσματος Εισόδου	160
6.2.1 Information Gain.....	160
6.2.2. mRMR.....	162
6.2.3 Random Forest	164
6.2.4 Λίστες Κατάταξης των Χαρακτηριστικών Εισόδου	166
6.3 Μοντέλα Πρόβλεψης/Ταξινόμησης.....	170
6.3.1 Μπεϋζιανός Ταξινομητής Naive Bayes	171
6.3.2 Τυχαία Δάση	174
6.4 Αποτελέσματα Μοντέλων Ταξινόμησης	177
6.4.1 Αποτελέσματα Ταξινόμησης Αλγορίθμου Naive Bayes.....	178
6.4.2 Αποτελέσματα Ταξινόμησης Αλγορίθμου Random Forest	183

6.5 Συζήτηση - Συμπεράσματα	188
Κεφάλαιο 7°	192
Το πρόβλημα BI-RADS 0 και η Αντιμετώπισή του μέσω των εργαλείων Μηχανικής Μάθησης.....	192
7.1 Το Πρόβλημα BI-RADS 0	192
7.2 Διαγνωστική Απόδοση του Ταξινομητή Random Forest για τα Περιστατικά γυναικών με αποτέλεσμα BI-RADS 0 στη Μαστογραφία τους	196
7.3 Συζήτηση - Συμπεράσματα	198
Κεφάλαιο 8°	199
Συμπεράσματα και Μελλοντική Έρευνα	199
8.1 Συζήτηση και Συμπεράσματα	199
8.2 Μελλοντική Έρευνα.....	203
Παράρτημα Α.....	206
Βασικά Διαγνωστικά Μέτρα Απόδοσης	206
Παράρτημα Β.....	210
Αποτελέσματα των Ταξινομητών με τη χρήση Διαφορετικών Λιστών Κατάταξης Χαρακτηριστικών ...	210
Βιβλιογραφία	219
Δημοσιεύσεις σε Διεθνή Περιοδικά και Συνέδρια (με χρονολογική σειρά)	238

Κατάλογος Εικόνων

Εικόνα 1: Ο καρκίνος του μαστού, ο πιο συχνός τύπος καρκίνου στις γυναίκες σε 157 χώρες το 2020 [2]	30
Εικόνα 2: Εκτιμώμενος αριθμός νέων περιπτώσεων καρκίνου του μαστού παγκοσμίως το 2020 στον γενικό πληθυσμό [2].....	31
Εικόνα 3: Εκτιμώμενος αριθμός περιστατικών καρκίνου του μαστού σε γυναίκες όλων των ηλικιών παγκοσμίως στα επόμενα 5 χρόνια από το 2020 [2].....	31
Εικόνα 4: Εκτιμώμενος αριθμός νέων περιπτώσεων καρκίνου του μαστού το 2020 στις γυναίκες στην Ευρώπη [2].....	32
Εικόνα 5: Εκτιμώμενος αριθμός θανάτων από καρκίνο για τις γυναίκες όλων των ηλικιών στην Ευρώπη [2].....	33
Εικόνα 6: Εκτιμώμενος αριθμός περιστατικών καρκίνου του μαστού το 2020 στην Ελλάδα για γυναίκες όλων των ηλικιών [2].....	34
Εικόνα 7: Εκτιμώμενος αριθμός νέων περιστατικών καρκίνου του μαστού και θανάτων σε γυναίκες όλων των ηλικιών στην Ελλάδα έως το 2040 [2].....	34
Εικόνα 8: Η ανατομία και φυσιολογία του μαστού [16].....	38
Εικόνα 9: Μορφή κυττάρων σε τύπο καρκίνου DCIS [21]	45
Εικόνα 10: Συμπτώματα του καρκίνου του μαστού [31].....	51
Εικόνα 11: Οι 4 κατηγορίες πυκνότητας των μαστών όπως απεικονίζονται στην εικόνα της Μαστογραφίας [97].....	79
Εικόνα 12: Ανίχνευση κακοήθειας - Μια αναπτυσσόμενη πυκνότητα ως το μόνο σημάδι παθολογίας, ορατό μόνο κατά τη σύγκριση της τρέχουσας εικόνας b με την προηγούμενη εικόνα a	80
Εικόνα 13: Τυπική εμφάνιση κακοήθους βλάβης με δυσδιάκριτα όρια και μεγαλύτερη πυκνότητα από το φυσιολογικό παρέγχυμα [97]	82
Εικόνα 14: Βλάβες μαστού στο υπερηχογράφημα – a) κακοήθης όγκος με δυσδιάκριτα όρια, έντονα υποηχογενής, με φωτεινές κουκκίδες που αντιπροσωπεύουν μικροασβεστώσεις- b) καλοήθης αλλοίωση - ινοαδένωμα - με ομαλά περιγράμματα, παράλληλο προσανατολισμό και ομογενή δομή [97].....	85
Εικόνα 15: Καρκίνος του μαστού στη μαγνητική τομογραφία. Κακοήθεις βλάβες (βέλη) που εμφανίζονται ως μαζική βλάβη (a) και εκτεταμένη μη μαζική ενίσχυση (Non-Mass-Like Enhancement - NMLE) (b) [97]	86
Εικόνα 16: Γραφική απεικόνιση περιστατικών ανά κατηγορία BI-RADS	145
Εικόνα 17: Απεικόνιση της διαφοράς των διαδικασιών Εξαγωγής χαρακτηριστικών (Feature extraction) και Επιλογής χαρακτηριστικών (Feature selection) [208].	151

Εικόνα 18: Πλαίσιο για την επιλογή χαρακτηριστικών [201].	152
Εικόνα 19: Λογικό διάγραμμα για τη σχέση μεταξύ των μοντέλων φίλτρου, ενσωματωμένων μοντέλων και περιτυλίγματος [212]	153
Εικόνα 20: Η διαδικασία επιλογής χαρακτηριστικών με βάση την ομαδοποίηση (clustering) [201].	158
Εικόνα 21: Οι τιμές του κέρδους πληροφορίας (IG) για τα 35 χαρακτηριστικά του διανύσματος εισόδου	162
Εικόνα 22: Κατάταξη των 35 χαρακτηριστικών εισόδου με χρήση του αλγορίθμου Random Forest και με βάση το μέτρο Gini importance.	166
Εικόνα 23: Ταξινόμηση με τον αλγόριθμο Τυχαίων Δασών [199].	175
Εικόνα 24: Η λειτουργία του αλγορίθμου Random Forest [256]	176
Εικόνα 25: Διάγραμμα ροής της κατασκευής των μοντέλων ταξινόμησης	178
Εικόνα 26: Διαγνωστική ακρίβεια του μοντέλου ταξινόμησης Naive Bayes σύμφωνα με τη λίστα κατάταξης των χαρακτηριστικών εισόδου με χρήση της λίστας Information Gain	179
Εικόνα 27: Ευαισθησία του μοντέλου ταξινόμησης Naive Bayes σύμφωνα με τη λίστα κατάταξης των χαρακτηριστικών εισόδου με χρήση της λίστας Information Gain	181
Εικόνα 28: Ειδικότητα του μοντέλου ταξινόμησης Naive Bayes σύμφωνα με τη λίστα κατάταξης των χαρακτηριστικών εισόδου με χρήση της λίστας Information Gain	182
Εικόνα 29: Διαγνωστική ακρίβεια του μοντέλου Random Forest σύμφωνα με τη λίστα κατάταξης κατάταξης των χαρακτηριστικών εισόδου με χρήση της λίστας mRMR	184
Εικόνα 30: Ευαισθησία του μοντέλου ταξινόμησης Random Forest σύμφωνα με τη λίστα κατάταξης των χαρακτηριστικών εισόδου με χρήση του mRMR	186
Εικόνα 31: Ειδικότητα του μοντέλου ταξινόμησης Random Forest σύμφωνα με τη λίστα κατάταξης των χαρακτηριστικών εισόδου με χρήση του mRMR	187
Εικόνα 32: Διαγνωστική ακρίβεια του μοντέλου ταξινόμησης Naive Bayes σύμφωνα με τη λίστα κατάταξης των χαρακτηριστικών εισόδου με χρήση της μεθόδου mRMR	210
Εικόνα 33: Ευαισθησία του μοντέλου ταξινόμησης Naive Bayes σύμφωνα με τη λίστα κατάταξης των χαρακτηριστικών εισόδου με χρήση της μεθόδου mRMR	211
Εικόνα 34: Ειδικότητα του μοντέλου ταξινόμησης Naive Bayes σύμφωνα με τη λίστα κατάταξης των χαρακτηριστικών εισόδου με χρήση της μεθόδου mRMR	212
Εικόνα 35: Διαγνωστική ακρίβεια του μοντέλου ταξινόμησης Random Forest σύμφωνα με τη λίστα κατάταξης των χαρακτηριστικών εισόδου με χρήση της μεθόδου Information Gain	213
Εικόνα 36: Ευαισθησία του μοντέλου ταξινόμησης Random Forest σύμφωνα με τη λίστα κατάταξης των χαρακτηριστικών εισόδου με χρήση της μεθόδου Information Gain	214

Εικόνα 37: Ειδικότητα του μοντέλου ταξινόμησης Random Forest σύμφωνα με τη λίστα κατάταξης των χαρακτηριστικών εισόδου με χρήση της μεθόδου Information Gain	215
Εικόνα 38: Διαγνωστική ακρίβεια του μοντέλου ταξινόμησης Random Forest σύμφωνα με τη λίστα κατάταξης των χαρακτηριστικών εισόδου με χρήση του gini importance	216
Εικόνα 39: Ευαισθησία του μοντέλου ταξινόμησης Random Forest σύμφωνα με τη λίστα κατάταξης των χαρακτηριστικών εισόδου με χρήση του gini importance	217
Εικόνα 40: Ειδικότητα του μοντέλου ταξινόμησης Random Forest σύμφωνα με τη λίστα κατάταξης των χαρακτηριστικών εισόδου με χρήση του gini importance	218

Κατάλογος Πινάκων

Πίνακας 1: Τα στάδια του καρκίνου του μαστού	48
Πίνακας 2: Ευαισθησία και Ειδικότητα μαστογραφίας ανάλογα με την ηλικία και την πυκνότητα μαστού [97].....	102
Πίνακας 3: Σύγκριση διαγνωστικής απόδοσης μεταξύ υπερηχογραφήματος μαστών και δύο μεθόδων βιοψίας (FNA & CNB) για τη διάγνωση των διηθημένων λεμφαδένων [97].....	103
Πίνακας 4: Απόδοση απεικονιστικών μεθόδων όταν συνδυάζονται με μαστογραφία σε πυκνούς μαστούς (γυναίκες με αρνητική μαστογραφία) [97]	103
Πίνακας 5: Συνολική ευαισθησία και ειδικότητα για τη διαγνωστική ικανότητα των απεικονιστικών μεθόδων μέσω των μελετών που περιλαμβάνονται στην ανασκόπηση των N. Aristokli et al [4].	104
Πίνακας 6: Διαγνωστική απόδοση των στατιστικών μοντέλων εκτίμησης κινδύνου του καρκίνου του μαστού και των ευφών συστημάτων υποβοηθούμενης διάγνωσης	125
Πίνακας 7: Μεταβλητές που προκύπτουν από τη φόρμα χαρακτηριστικών και τις διαγνωστικές εξετάσεις και χαρακτηρίζουν κάθε ασθενή (Μεταβλητές Διανύσματος Χαρακτηριστικών)	138
Πίνακας 8: Αριθμός περιστατικών ανά κατηγορία σύμφωνα με τη βαθμονόμηση BI-RADS	144
Πίνακας 9: Πίνακας Σύγκρισης 1 - Αποτελέσματα της Μαστογραφίας με κατώφλι θετικότητας (BI-RADS 4a+)	146
Πίνακας 10: Πίνακας Σύγκρισης 2 - Αποτελέσματα της Μαστογραφίας με κατώφλι θετικότητας τη βαθμολογία (BI-RADS 4b+).....	146
Πίνακας 11: Πίνακας Σύγκρισης 3 - Αποτελέσματα της Μαστογραφίας με κατώφλι θετικότητας (BI-RADS 4c+).....	146
Πίνακας 12: Διαγνωστική απόδοση της μαστογραφίας για την ανίχνευση του καρκίνου του μαστού στον υπό μελέτη πληθυσμό	147
Πίνακας 13: Επιμέρους λίστες κατάταξης των χαρακτηριστικών εισόδου ως προς την αποτελεσματικότητα/σημαντικότητά τους για κάθε μια από τις τεχνικές κατάταξης	168
Πίνακας 14: Βέλτιστο υποσύνολο χαρακτηριστικών για το μοντέλο ταξινόμησης Naive Bayes	180
Πίνακας 15: Πίνακας σύγκρισης του ταξινομητή Naive Bayes για τα 8 καλύτερα χαρακτηριστικά με βάση το Information Gain.....	180
Πίνακας 16: Διαγνωστική απόδοση του μοντέλου ταξινόμησης Naive Bayes με χρήση του Information Gain και με το βέλτιστο υποσύνολο χαρακτηριστικών	182
Πίνακας 17: Βέλτιστο υποσύνολο χαρακτηριστικών για το μοντέλο ταξινόμησης Random Forest.....	184
Πίνακας 18: Πίνακας σύγκρισης του ταξινομητή Random Forest για τα καλύτερα 7 χαρακτηριστικά....	185

Πίνακας 19: Διαγνωστική απόδοση του ταξινομητή Random Forest με χρήση του mRMR και με το βέλτιστο υποσύνολο χαρακτηριστικών.....	187
Πίνακας 20: Συγκεντρωτικά αποτελέσματα της διαγνωστικής απόδοσης των επιλεγμένων ταξινομητών	188
Πίνακας 21: Συγκριτικός πίνακας για τη διαγνωστική απόδοση του ταξινομητή RF σε σχέση με τη μαστογραφία*	190
Πίνακας 22: Πίνακας σύγκρισης του ταξινομητή Random Forest σε σύνολο δοκιμής αποκλειστικά από την κατηγορία BI-RADS 0.....	197
Πίνακας 23: Διαγνωστική απόδοση του αλγορίθμου RF σε σύνολο δοκιμής αποκλειστικά από την κατηγορία BI-RADS 0.....	197
Πίνακας 24: Πίνακας σύγκρισης για δυαδική ταξινόμηση.	206

Συντομογραφίες

ABUS	Automated Breast Ultrasound
ADH	Atypical Ductal Hyperplasia
AI	Artificial Intelligence
AUC	Area under the ROC curve
BCRAT	Breast Cancer Risk Assessment Tool
BCSC	Breast Cancer Surveillance Consortium
BI-RADS	Breast Imaging Reporting and Data System
BOADICEA	Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm
BPNN	Back Propagation Neural Network
BSE	Breast Self-Examination
CBE	Clinical Breast Examination
CDSS	Clinical Decision Support System
CE - MAMMO	Contrast - Enhanced Mammography
CE - MRI	Contrast - Enhanced Magnetic Resonance Imaging
CI	Confidence Interval
CNB	Core Needle Biopsy
CPGs	Clinical Practice Guidelines

DBT	Digital Breast Tomosynthesis
DCIS	Ductal Carcinoma In Situ
ER	Estrogen receptor
FNA	Fine Needle Aspiration
FS	Feature Selection
GL-DSS	Guidelines Based Decision Support System
GLOBOCAN	Global Cancer Observatory
HER2	Human Epidermal growth factor Receptor Hand-Held Ultrasound
HHUS	Hand-Held Ultrasound
IBC	Inflammatory Breast Cancer
IDC	Invasive Ductal Carcinoma
IG	Information Gain
ILC	Invasive Lobular Carcinoma
kNN	k-Nearest Neighbors algorithm
LCIS	Lobular Carcinoma In Situ
LR	Linear Regression
ML	Machine Learning
MLP	Multilayer Perceptron

MRI	Magnetic Resonance Imaging
mRMR	Max - Relevance Min - Redundancy
NB	Naive Bayes
NCCN	National Comprehensive Cancer Network
NCI	National Cancer Institute
NPV	Negative Predictive Values
PPV	Positive Predictive Values
PR	Progesterone receptor
PRS	Polygenic Risk Score
RF	Random Forest
SVM	Support Vector Machine
US	Ultrasound
VAB	Vacuum-Assisted Biopsy
VABD	Vacuum-Assisted Biopsy Device
WDBC	Wisconsin Breast Cancer Dataset (diagnostic)
WOBC	Wisconsin Breast Cancer Dataset (original)
ΕΠΕ	Ελληνική Παθολογική Εταιρεία
ΠΟΥ	Παγκόσμιος Οργανισμός Υγείας

Κεφάλαιο 1^ο

Εισαγωγή

Ο καρκίνος του μαστού, με βάση τα επιδημιολογικά δεδομένα, αποτελεί αναμφίβολα ένα μεγάλο ζήτημα υγείας με επιπτώσεις τόσο σε ατομικό όσο και σε κοινωνικό επίπεδο. Είναι μια εξαιρετικά ετερογενής νόσος και ποικίλλει ως προς την αιτιολογία και τα παθολογικά χαρακτηριστικά της. Ανεξάρτητα από το επίπεδο ανάπτυξης των χωρών, ο καρκίνος του μαστού είναι ο κυριότερος καρκίνος που προσβάλλει τον γυναικείο πληθυσμό. Σύμφωνα με τα πιο πρόσφατα δεδομένα του Παγκόσμιου Παρατηρητηρίου για τον Καρκίνο (Global Cancer Observatory - GLOBOCAN), για το 2020, ο γυναικείος καρκίνος του μαστού είναι ο πιο συχνός τύπος καρκίνου σε 159 από τις 185 χώρες [1]. Ο καρκίνος του μαστού είναι η πέμπτη κύρια αιτία θνησιμότητας από καρκίνο παγκοσμίως, με 685.000 θανάτους [1]. Όσον αφορά στα περιστατικά καρκίνου του μαστού στην Ελλάδα, πρόκειται για το συχνότερο καρκίνο και την κύρια αιτία θανάτου στις γυναίκες (τρίτη κύρια αιτία στο γενικό πληθυσμό) σύμφωνα με τα επιδημιολογικά δεδομένα του 2020 [2]. Λαμβάνοντας υπόψη τις τεράστιες επιπτώσεις της νόσου στην παγκόσμια κοινότητα, μία από τις μεγαλύτερες προκλήσεις της ιατρικής και βιοϊατρικής επιστήμης είναι η όσο το δυνατόν πιο έγκαιρη και ακριβής διάγνωση του καρκίνου του μαστού.

Η πρόληψη και η έγκαιρη διάγνωση είναι κρίσιμης και ζωτικής σημασίας για την αντιμετώπιση του καρκίνου του μαστού. Όσο νωρίτερα ανιχνεύεται η νόσος, τόσο καλύτερη είναι η πρόγνωση για την επιβίωση των ασθενών, με αποτέλεσμα την αποφυγή δύσκολων θεραπειών. Κάθε περίπτωση της νόσου θα πρέπει να αντιμετωπίζεται εξατομικευμένα, προκειμένου να επιτευχθεί η βέλτιστη διαχείριση και θεραπεία. Σύμφωνα με τις ευρωπαϊκές κατευθυντήριες οδηγίες και τα πρωτόκολλα για τη διαχείριση και τον προσυμπτωματικό έλεγχο γυναικών με πιθανό καρκίνο του μαστού, η μαστογραφία θεωρείται το διαγνωστικό εργαλείο αναφοράς [3]. Στη συνέχεια, ακολουθεί η χρήση του υπερηχογραφήματος του μαστού ως συμπληρωματική ιατρική απεικονιστική μέθοδος και σε πιο σπάνιες και ειδικές περιπτώσεις εξαιρετικά πυκνών μαστών, διενεργείται επίσης μαγνητική τομογραφία (Magnetic Resonance Imaging-MRI) [3]. Η μαστογραφία και ο υπέρηχος μαστών θεωρούνται οι βασικές διαγνωστικές εξετάσεις που χρησιμοποιούνται για την ανίχνευση του καρκίνου του μαστού στην καθημερινή κλινική πράξη, ωστόσο τα ποσοστά της διαγνωστικής του απόδοσης δεν είναι ικανοποιητικά σε πολλές περιπτώσεις [4]. Από την άλλη πλευρά, η MRI είναι μια εξέταση υψηλής διαγνωστικής ακρίβειας αλλά λόγω του υψηλού της

κόστους δεν θεωρείται οικονομικά αποδοτική για τα συστήματα υγειονομικής περίθαλψης [4]. Επιπλέον, ο γενετικός έλεγχος είναι επίσης μια εξειδικευμένη εξέταση υψηλού κόστους και χρησιμοποιείται αποκλειστικά για τη διάγνωση των σπάνιων περιπτώσεων του κληρονομικού καρκίνου του μαστού [5]. Τέλος, η βιοψία των μαστών αποτελεί μια επεμβατική πρακτική και έχει αποδειχθεί ότι ακόμα και η ελάχιστη επεμβατική και ευρέως χρησιμοποιούμενη βιοψία με λεπτή βελόνα αναρρόφησης, η FNA, σε ορισμένες περιπτώσεις ενέχει κινδύνους και παράγει ψευδώς αρνητικά αποτελέσματα [6].

Η παρούσα διδακτορική διατριβή έχει ως στόχο την ανάπτυξη μοντέλων βασισμένων στην Τεχνητή Νοημοσύνη και τη δημιουργία ενός Συστήματος υποστήριξης κλινικής απόφασης με σκοπό την έγκαιρη διάγνωση και την εξατομικευμένη διαχείριση των γυναικών που βρίσκονται σε κίνδυνο ανάπτυξης καρκίνου του μαστού. Η ανάπτυξη των μοντέλων αυτών βασίζεται σε μη επεμβατικά ιατρικά δεδομένα που συλλέγονται κατά την καθημερινή κλινική πρακτική στο Κέντρο Μαστού του Γενικού Νοσοκομείου «Έλενα Βενιζέλου», καθώς και σε δεδομένα του προσωπικού ιατρικού φακέλου των γυναικών του Κέντρου Μαστού. Ως εκ τούτου, στοχεύουμε στην κατασκευή ενός εργαλείου, το οποίο μέσω του συνδυασμού των αποτελεσμάτων μη επεμβατικών και χαμηλού κόστους εξετάσεων και του ιατρικού φακέλου, θα είναι εύκολο στην καθημερινή χρήση του από τους κλινικούς ιατρούς και ταυτόχρονα θα είναι οικονομικά αποδοτικό για τα συστήματα υγείας.

Στη συνέχεια παρουσιάζεται η διάρθρωση της διατριβής. Στο Κεφάλαιο 2 παρουσιάζονται τα επιδημιολογικά δεδομένα για τον καρκίνο του μαστού τόσο παγκόσμια όσο και στην Ελλάδα, αναδεικνύοντας την υψίστης σημασίας ανάγκη αντιμετώπισης της νόσου. Στη συνέχεια, αναλύεται η ανατομία και παθολογία του μαστού όπως επίσης και οι κατηγορίες των παθήσεων του μαστού. Δίνεται ιδιαίτερη έμφαση στον εντοπισμό των παραγόντων κινδύνου ανάπτυξης του καρκίνου του μαστού καθώς στην παρούσα διδακτορική διατριβή χρησιμοποιούμε σχεδόν όλους του παράγοντες κινδύνου που έχουν αναφερθεί στη διεθνή βιβλιογραφία. Στη συνέχεια, στο Κεφάλαιο 3 περιγράφονται όλες οι διαθέσιμες διαγνωστικές μέθοδοι που εφαρμόζονται για την ανίχνευση του καρκίνου του μαστού και αναλύονται τα πλεονεκτήματα και οι αδυναμίες τους. Αφού παρουσιάζεται η διαγνωστική απόδοση όλων των διαγνωστικών εξετάσεων για τον καρκίνο του μαστού, το κεφάλαιο κλείνει με τους συγκριτικούς πίνακες ανασκόπησης της απόδοσης των διαγνωστικών διαδικασιών και των συνδυασμών τους. Στο Κεφάλαιο 4 παρατίθενται αναλυτικά όλα τα δημοσιευμένα, μέχρι σήμερα, στη διεθνή βιβλιογραφία, στατιστικά μοντέλα εκτίμησης κινδύνου και τα ευφυή συστήματα υποστήριξης της κλινική απόφασης μαζί με το είδος των δεδομένων που χρησιμοποιούν αλλά και τη διαγνωστική τους απόδοση. Στο τέλος του κεφαλαίου, παρουσιάζεται ο συγκεντρωτικός πίνακας με τις, συναφέστερες στη μελέτη μας, έρευνες και τη διαγνωστική απόδοση των συστημάτων που παρουσίασαν. Έπειτα, στο Κεφάλαιο 5, παρουσιάζεται το ιατρικό πρόβλημα που καλείται να αντιμετωπίσει η παρούσα διδακτορική έρευνα και αναλύονται τα στάδια

προεπεξεργασίας και επεξεργασίας των πραγματικών δεδομένων που είναι αντικείμενο της έρευνάς μας, αλλά και ο τρόπος που αντιμετωπίστηκε το πρόβλημα των ελλιπών τιμών. Στο Κεφάλαιο 6 αναλύονται οι τεχνικές επιλογής χαρακτηριστικών στη μηχανική μάθηση και πραγματοποιείται η αξιολόγηση και κατάταξη των χαρακτηριστικών του διανύσματος εισόδου των προτεινόμενων μοντέλων με τη χρήση των μεθόδων επιλογής χαρακτηριστικών. Στη συνέχεια, παρουσιάζεται και αναλύεται η κατασκευή των μοντέλων ταξινόμησης με χρήση των αλγορίθμων Naive Bayes και Random Forest. Για την κατασκευή των μοντέλων ταξινόμησης χρησιμοποιήθηκαν τα διανύσματα εισόδου χαρακτηριστικών (λίστες κατάταξης) που προέκυψαν από την εφαρμογή των μεθόδων επιλογής χαρακτηριστικών. Χρησιμοποιώντας τις λίστες κατάταξης χαρακτηριστικών στοχεύσαμε στον εντοπισμό του βέλτιστου υποσυνόλου χαρακτηριστικών με τη μεγαλύτερη διαγνωστική πληροφορία, δηλαδή του διανύσματος εισόδου που επιτυγχάνει την υψηλότερη απόδοση των ταξινομητών με το μικρότερο δυνατό πλήθος χαρακτηριστικών. Στο τέλος του κεφαλαίου παρουσιάζονται αναλυτικά τα αποτελέσματα των ταξινομητών και καταλήγουμε στην επιλογή του μοντέλου με την υψηλότερη απόδοση σύμφωνα με τους εγκυρότερους δείκτες αξιολόγησης της διαγνωστικής απόδοσης. Στο Κεφάλαιο 7 που ακολουθεί, επικεντρωνόμαστε σε μια ειδική κατηγορία γυναικών, αυτές με την αξιολόγηση BI-RADS 0 στη μαστογραφία τους, προκειμένου να διερευνήσουμε την απόδοση του μοντέλου στο οποίο καταλήξαμε στο Κεφάλαιο 6. Στόχος είναι να αποδειχθεί η δυνατότητα του μοντέλου να κατατάσσει σωστά τις περιπτώσεις οι οποίες κατά τη διαγνωστική διαδικασία αποτελούν το «μαύρο κουτί» για τους κλινικούς, μην επιτρέποντάς τους να εξάγουν κάποιο συμπέρασμα για τον κίνδυνο ανάπτυξης καρκίνου του μαστού. Τέλος, στο Κεφάλαιο 8, συνοψίζονται τα συμπεράσματα που προέκυψαν καθ' όλη τη διάρκεια της πορείας της διδακτορικής διατριβής, σημειώνοντας την υψηλή διαγνωστική απόδοση του προτεινόμενου μοντέλου. Τονίζεται ιδιαίτερα ότι το προτεινόμενο CDSS είναι ένα οικονομικά αποδοτικό εργαλείο με καθολική εφαρμογή και σημαντική προστιθέμενη αξία στη διαδικασία διάγνωσης του καρκίνου του μαστού. Κλείνοντας, παρουσιάζουμε το στόχο μας να αναπτύξουμε μια ομοσπονδιακή έκδοση του προτεινόμενου CDSS, ώστε να μπορεί να εκτελεστεί εντός ομοσπονδιακών περιβαλλόντων για την υποστήριξη της εκπαίδευσης σε βάσεις δεδομένων με κλινικά δεδομένα από διαφορετικές χώρες, όπου η εγκατάσταση τοπικού εξοπλισμού ή η ενσωμάτωση των δεδομένων δεν είναι εφικτή.

Τέλος θα θέλαμε να τονίσουμε ότι η παρούσα διδακτορική διατριβή δεν θα μπορούσε να πραγματοποιηθεί χωρίς την ουσιαστική συμμετοχή και συνεργασία των χειρουργών μαστού, των ακτινολόγων και των γυναικολόγων, οι οποίες και οποίοι με την διάθεση των ιατρικών δεδομένων και την καθοδήγηση τους συνέβαλαν τα μέγιστα στην επιτυχή ολοκλήρωση της έρευνας και της διατριβής.

Κεφάλαιο 2^ο

Ο Καρκίνος του Μαστού

2.1 Επιδημιολογικά Δεδομένα για τον Καρκίνο του Μαστού

Ο καρκίνος αποτελεί μία από τις κύριες αιτίες θανάτου παγκοσμίως. Το 2008 καταγράφηκαν 8 εκατομμύρια θάνατοι ως αποτέλεσμα κακοήθων παθήσεων και ο αριθμός αυτός εκτιμάται ότι θα φθάσει τα 11 εκατομμύρια μέχρι το 2030 [7]. Ο καρκίνος του μαστού αποτελεί την πιο συχνή κακοήθη νόσο στις γυναίκες και είναι μία από τις σημαντικότερες αιτίες θανάτου. Πρόκειται για μια πολυπαραγοντική νόσο και διάφοροι παράγοντες συμβάλλουν στην εμφάνισή της. Παρόλο που η νόσος εμφανίζεται σε όλο τον κόσμο, τα ποσοστά εμφάνισης, θνησιμότητας και επιβίωσης διαφέρουν σημαντικά μεταξύ των διαφόρων περιοχών του κόσμου, γεγονός που μπορεί να οφείλεται σε πολλούς παράγοντες, όπως η δομή του πληθυσμού, ο τρόπος ζωής, οι γενετικοί παράγοντες και το περιβάλλον. Οι αλλαγές στους παράγοντες κινδύνου έχουν οδηγήσει σε αύξηση του επιπολασμού του καρκίνου του μαστού. Παρόλο που ο συστηματικός προσυμπτωματικός έλεγχος μπορεί να μειώσει την επιβάρυνση από τον καρκίνο του μαστού, η υπερδιάγνωση και το αυξημένο κόστος είναι τα μειονεκτήματα αυτής της μεθόδου. Η ταξινόμηση των γυναικών με βάση τους παράγοντες κινδύνου για καρκίνο του μαστού μπορεί να είναι αποτελεσματική στη βελτίωση των μεθόδων χωρίς κίνδυνο και στο σχεδιασμό στοχευμένων προγραμμάτων προσυμπτωματικού ελέγχου για τον καρκίνο του μαστού [7].

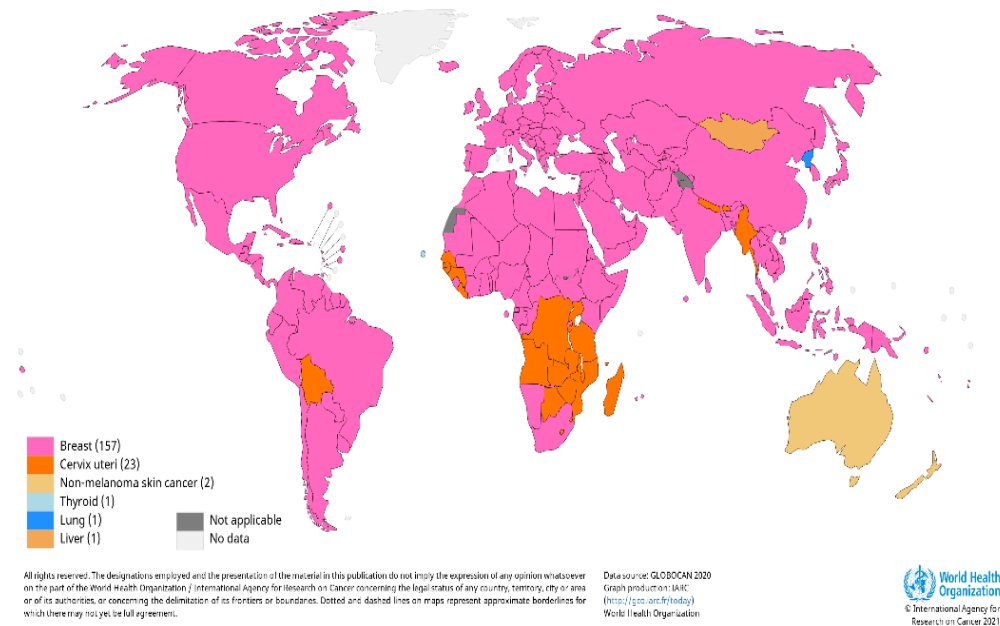
Μια πρόσφατη ενδελεχής επιδημιολογική μετά - ανάλυση για τον καρκίνο του μαστού διεξήχθη από τους K. Maajani et al. [8] και συμπεριελάμβανε 126 μελέτες. Σκοπός αυτής της μελέτης ήταν η μέτρηση των παγκόσμιων και περιφερειακών ποσοστών επιβίωσης των γυναικών με καρκίνο του μαστού. Η εν λόγω ερευνητική ομάδα αναζήτησε πληροφορίες στο Medline/PubMed, το Web of Science, το Scopus και το Google Scholar ώστε να εντοπίσει τις μελέτες σχετικά με το ποσοστό επιβίωσης των γυναικών με πρωτοπαθή διηθητικό καρκίνο του μαστού έως το τέλος Ιουνίου 2017. Στην έρευνά τους, χρησιμοποίησαν μοντέλα τυχαίων επιδράσεων για να εκτιμήσουν τα συγκεντρωτικά ποσοστά επιβίωσης σε διάστημα ενός, τριών, πέντε και δέκα ετών. Προκειμένου να διερευνηθούν οι πιθανές πηγές ετερογένειας των αποτελεσμάτων, έγινε ανάλυση των δεδομένων σε υποομάδες και χρησιμοποιήθηκαν κατάλληλα μοντέλα παλινδρόμησης. Οι ετερογένειες μεταξύ διαφορετικών μελετών στα ποσοστά επιβίωσης 1, 3, 5 και 10 ετών ήταν σημαντικά υψηλές. Τα συνολικά ποσοστά επιβίωσης 1, 3, 5 και 10 ετών σε γυναίκες με καρκίνο του

μαστού ήταν 92% (με διάστημα εμπιστοσύνης (Confidence Interval-CI) 95%, 90 - 94), 75% (95% CI, 71 - 79), 73% (95% CI, 71 - 75), και 61% (95% CI, 54 - 67), αντίστοιχα. Η ανάλυση των υποομάδων αποκάλυψε ότι τα ποσοστά επιβίωσης ποικίλλουν σε διάφορες περιοχές σύμφωνα με τον Παγκόσμιο Οργανισμό Υγείας (ΠΟΥ), ανάλογα με την ηλικία, το στάδιο διάγνωσης, το έτος της μελέτης και τον βαθμό ανάπτυξης των χωρών που συμμετέχουν στη μελέτη. Τα μοντέλα παλινδρόμησης που χρησιμοποιήθηκαν έδειξαν ότι το έτος διεξαγωγής της μελέτης και ο βαθμός ανάπτυξης της εκάστοτε χώρας αποτέλεσαν πιθανές πηγές ετερογένειας. Ένα γενικό συμπέρασμα αποτελεί το γεγονός ότι το ποσοστό επιβίωσης από τον καρκίνο του μαστού βελτιώθηκε τις τελευταίες δεκαετίες. Ωστόσο, είναι χαμηλότερο στις αναπτυσσόμενες χώρες από τις ανεπτυγμένες [8].

2.1.1 Στοιχεία σε παγκόσμιο επίπεδο

Μία από τις μεγαλύτερες προκλήσεις της Ιατρικής και Βιοϊατρικής επιστήμης είναι η βέλτιστη αντιμετώπιση του καρκίνου και φυσικά η εξάλειψή του. Συγκεκριμένα, ο καρκίνος του μαστού στις γυναίκες, όπως αποδεικνύεται από τα σύγχρονα δεδομένα, παίρνει συνεχώς αυξανόμενες διαστάσεις δημιουργώντας ένα σημαντικό πρόβλημα δημόσιας υγείας [9], [10]. Η ζωή εκατομμυρίων γυναικών παγκοσμίως επηρεάζεται δραστικά από αυτήν τη νόσο καθώς αποτελεί ύψιστης σημασίας κίνδυνο για μια μεγάλη μερίδα γυναικών, όλων των ηλικιών, δίνοντας το κυριότερο κίνητρο για την όσο το δυνατόν πληρέστερη μελέτη και κατανόηση της ασθένειας.

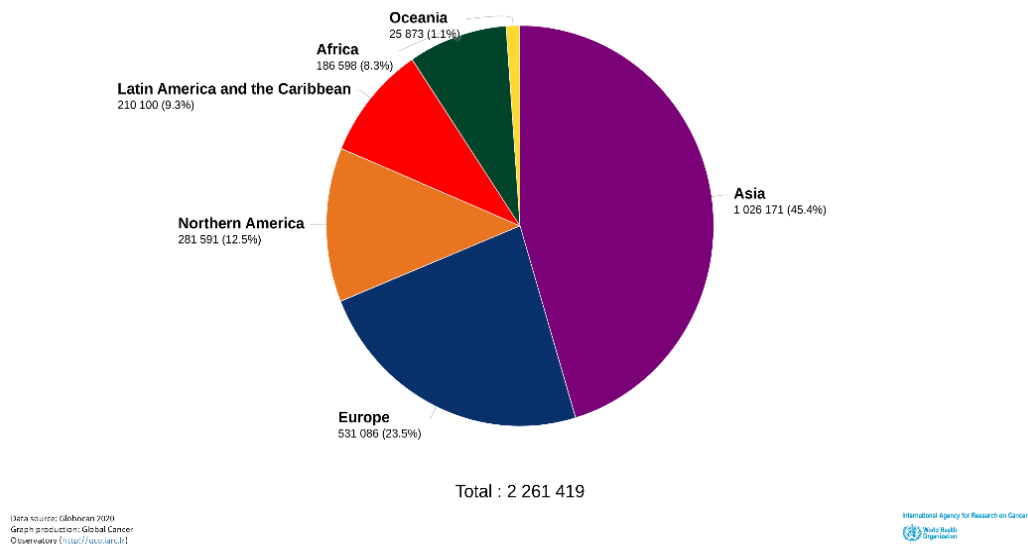
Ο καρκίνος του μαστού είναι εξαιρετικά ετερογενής και ποικίλλει ως προς την αιτιολογία και τα παθολογικά χαρακτηριστικά του. Ανεξάρτητα από το επίπεδο ανάπτυξης των χωρών, ο καρκίνος του μαστού είναι ο συχνότερος καρκίνος που επηρεάζει τον γυναικείο πληθυσμό, όπως διαπιστώνεται από τα ετήσια νέα περιστατικά. Το GLOBOCAN ενημερώνει ότι ο καρκίνος του μαστού ήταν ο συχνότερος τύπος καρκίνου στις γυναίκες σε 157 χώρες το 2020, όπως παρουσιάζεται στην Εικόνα 1 [2].



Εικόνα 1: Ο καρκίνος του μαστού, ο πιο συχνός τύπος καρκίνου στις γυναίκες σε 157 χώρες το 2020 [2]

Σύμφωνα με τον ΠΟΥ¹, το 2020, σε όλο τον κόσμο, σημειώθηκαν 2,3 εκατομμύρια γυναίκες που διαγνώστηκαν με καρκίνο του μαστού και 685.000 θάνατοι. Ο καρκίνος του μαστού θεωρείται ως ο επικρατέστερος καρκίνος στις γυναίκες, καθώς μέχρι το τέλος του 2020, 7,8 εκατομμύρια γυναίκες διαγνώστηκαν με καρκίνο του μαστού τα τελευταία 5 χρόνια. Σύμφωνα με τα παγκόσμια στατιστικά στοιχεία, οι γυναίκες που διαγνώστηκαν με καρκίνο του μαστού είχαν περισσότερα DALYs (Disability - Adjusted Life Years) από αυτές που έπασχαν από άλλους τύπους καρκίνου. Τα DALYs για μια ασθένεια είναι το άθροισμα των ετών ζωής που χάθηκαν λόγω του πρόωρου θανάτου και των ετών που έζησαν οι ασθενείς με αναπηρία λόγω της νόσου [11]. Ο καρκίνος του μαστού μπορεί να εμφανιστεί σε γυναίκες κάθε ηλικίας μετά την εφηβεία, αλλά η εμφάνισή του επηρεάζεται από πολλαπλούς παράγοντες κινδύνου.

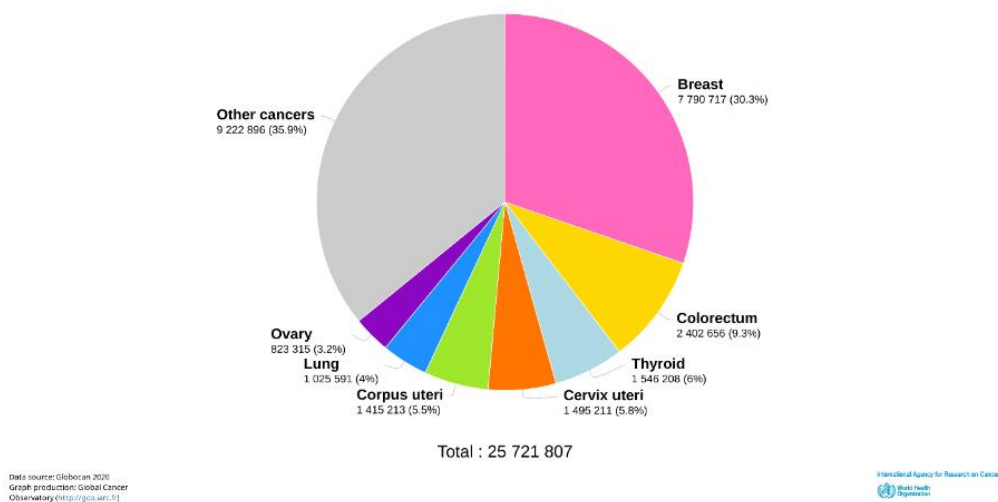
Σύμφωνα με τις στατιστικές μετρήσεις του GLOBOCAN, σχεδόν 1,7 εκατομμύρια γυναίκες διαγνώστηκαν με καρκίνο του μαστού παγκοσμίως το 2012. Αυτές οι περιπτώσεις οδήγησαν σε 522.000 θανάτους. Σε σύγκριση με τις αντίστοιχες τιμές του 2008, η αύξηση της συχνότητας εμφάνισης καρκίνου του μαστού και θνησιμότητας έφτασε το 18%. Σύμφωνα με τα τελευταία στατιστικά στοιχεία που παρουσιάζονται στην Εικόνα 2, το 2020, η συχνότητα εμφάνισης καρκίνου του μαστού στην Ευρώπη εκτιμάται ότι είχε φτάσει τα 531.086 νέα περιστατικά [2].



Εικόνα 2: Εκτιμώμενος αριθμός νέων περιπτώσεων καρκίνου του μαστού παγκοσμίως το 2020 στον γενικό πληθυσμό [2]

Το GLOBOCAN προβλέπει επίσης ότι τα νέα περιστατικά εμφάνισης γυναικείου καρκίνου του μαστού παγκοσμίως θα φτάνουν περίπου τα 3,2 εκατομμύρια ετησίως έως το 2050 [2]. Αυτοί οι αριθμοί αντικατοπτρίζουν τον τεράστιο αντίκτυπο του καρκίνου του μαστού στην παγκόσμια κοινότητα καθώς και την ανάγκη για επείγουσα διεπιστημονική συνεργασία για την πρόληψη, την έγκαιρη διάγνωση και την αποτελεσματική θεραπεία της νόσου. Ο εκτιμώμενος αριθμός περιστατικών καρκίνου του μαστού σε γυναίκες όλων των ηλικιών παγκοσμίως αναμένεται να πλησιάσει τα 8 εκατομμύρια στα επόμενα 5 χρόνια, από τις τελευταίες εκτιμήσεις του GLOBOCAN το 2020³, όπως παρουσιάζεται στην Εικόνα 3 [2].

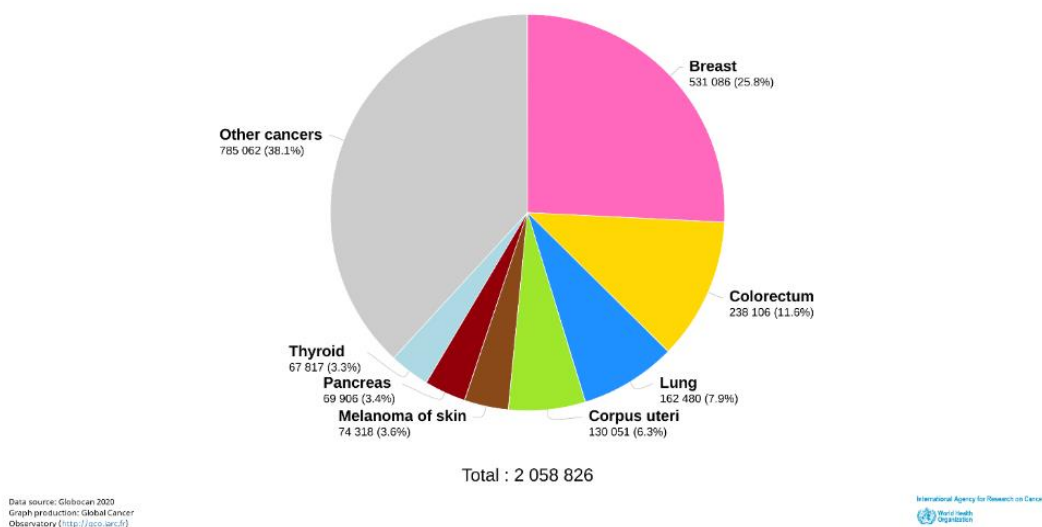
Estimated number of prevalent cases (5-year) in 2020, worldwide, females, all ages



Εικόνα 3: Εκτιμώμενος αριθμός περιστατικών καρκίνου του μαστού σε γυναίκες όλων των ηλικιών παγκοσμίως στα επόμενα 5 χρόνια από το 2020 [2]

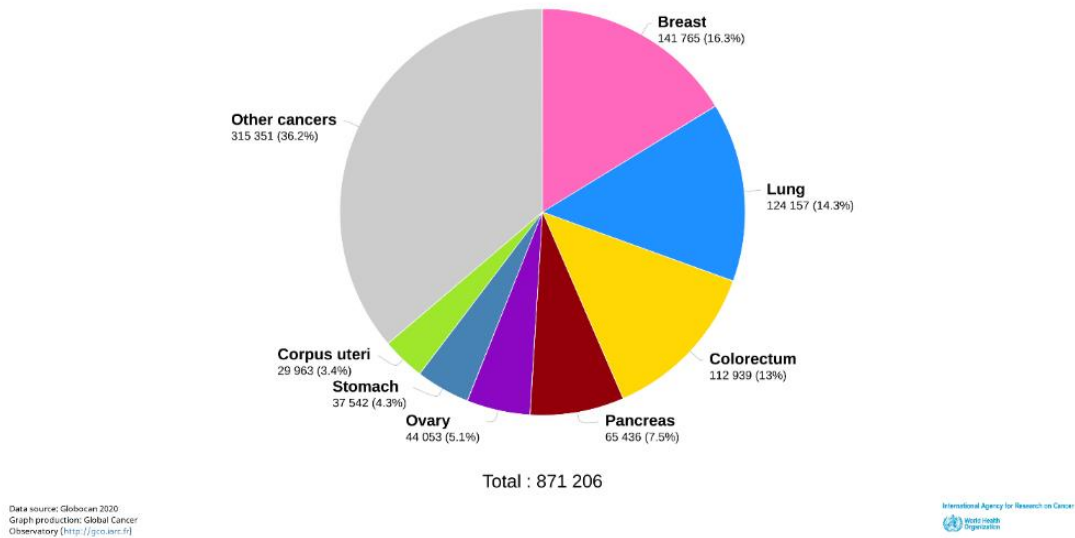
2.1.2 Ευρώπη

Σύμφωνα με τα τελευταία στοιχεία του GLOBOCAN, η Ευρώπη καταλαμβάνει την δεύτερη θέση σε αριθμό νέων περιστατικών καρκίνου του μαστού ανάμεσα στις 6 ηπείρους (η Ανταρκτική δεν συμπεριλαμβάνεται στη σύγκριση), για τη χρονιά 2020, με 531.086 γυναίκες ασθενείς και ποσοστό 23,5% επί των συνολικών περιπτώσεων παγκοσμίως, μετά την Ασία που ηγείται με 45,4% όπως φαίνεται παραπάνω στην Εικόνα 2. Όπως βλέπουμε στην Εικόνα 4, ο επικρατέστερος τύπος καρκίνου για τις γυναίκες στην Ευρώπη είναι ο καρκίνος του μαστού με ποσοστό 25,8% έναντι των άλλων τύπων καρκίνου [2].



Εικόνα 4: Εκτιμώμενος αριθμός νέων περιπτώσεων καρκίνου του μαστού το 2020 στις γυναίκες στην Ευρώπη [2]

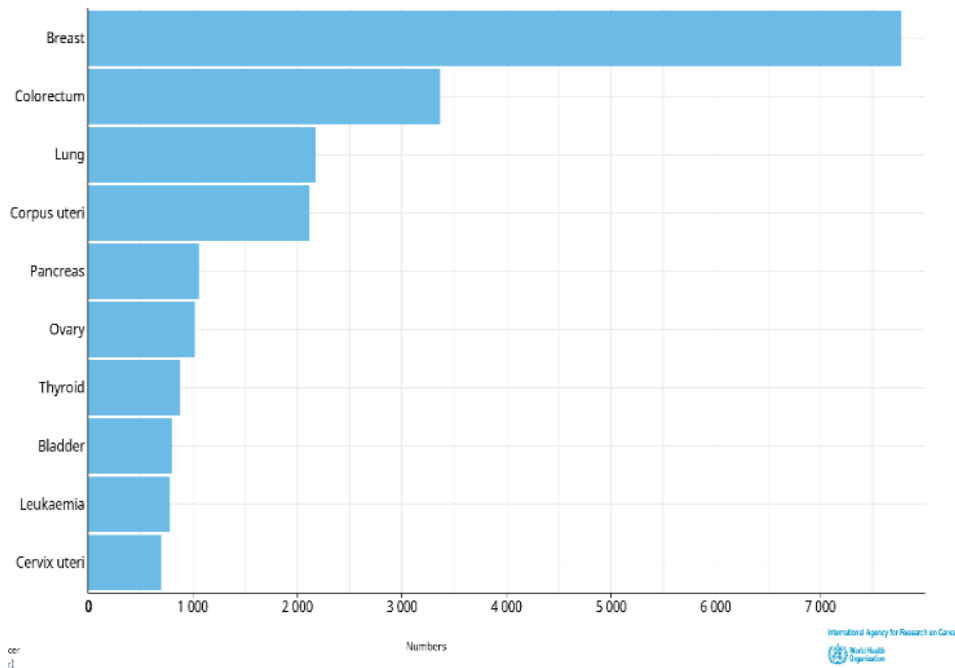
Όσον αφορά στους θανάτους των γυναικών, στην Ευρώπη, λόγω καρκίνου του μαστού, σύμφωνα με τα δεδομένα του 2020, 141.765 γυναίκες έχασαν τη μάχη με τη νόσο. Το ποσοστό θανάτων λόγω καρκίνου του μαστού είναι το υψηλότερο για τις γυναίκες της Ευρώπης όπως φαίνεται στην Εικόνα 5, φτάνοντας το 16,3% με δεύτερο το ποσοστό θανάτων λόγω του καρκίνου στον πνεύμονα που αγγίζει το 14,3% [2].



Εικόνα 5: Εκτιμώμενος αριθμός θανάτων από καρκίνο για τις γυναίκες όλων των ηλικιών στην Ευρώπη [2]

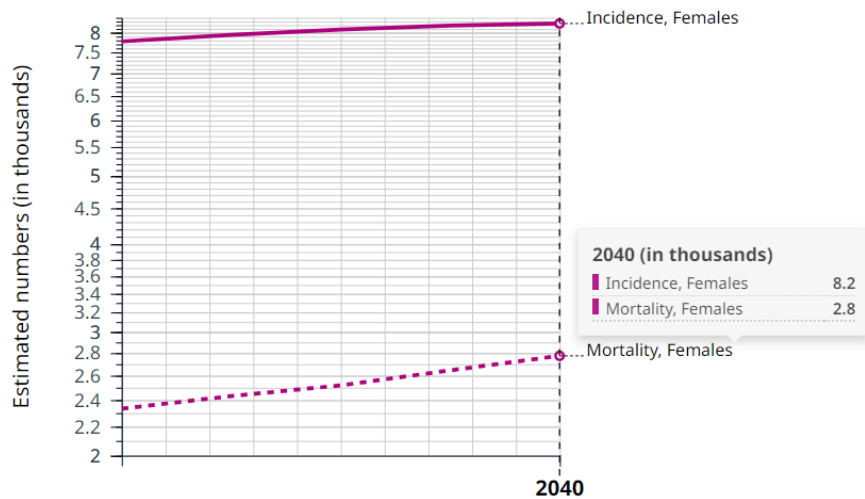
2.1.3 Ελλάδα

Ο καρκίνος του μαστού είναι ο συχνότερος τύπος καρκίνου στις γυναίκες στην Ελλάδα, όπως παρουσιάζεται στην Εικόνα 6, και υπεύθυνος για τους περισσότερους θανάτους λόγω καρκίνου. Στην πραγματικότητα, είναι η τρίτη κύρια αιτία θανάτου από καρκίνο στον γενικό πληθυσμό. Αυτό μεταφράζεται σε 7.772 γυναίκες που διαγνώστηκαν για πρώτη φορά με καρκίνο του μαστού το 2020 και 2.333 που πέθαναν από τη νόσο, όπως παρουσιάζεται από το GLOBOCAN. Από αυτές, οι 1.492 ήταν κάτω των 49 ετών, οι 3.165 ανήκαν στην ηλικιακή ομάδα 50 - 69 και οι υπόλοιπες ήταν γυναίκες άνω των 70 ετών. Με βάση τις αναφερόμενες περιπτώσεις, συμπεραίνουμε ότι περίπου το 40% όλων των ασθενών ανήκαν στη μεσαία ηλικιακή ομάδα. Ένα αρκετά ενθαρρυντικό στοιχείο αποτελεί το γεγονός ότι η επίπτωση του καρκίνου του μαστού σε γυναίκες κάτω των 34 ετών είναι εξαιρετικά μικρή [2].



Εικόνα 6: Εκτιμώμενος αριθμός περιστατικών καρκίνου του μαστού το 2020 στην Ελλάδα για γυναίκες όλων των ηλικιών [2]

Στην Εικόνα 7 αποτυπώνεται ο εκτιμώμενος αριθμός νέων περιστατικών καρκίνου του μαστού και θανάτων σε γυναίκες όλων των ηλικιών στην Ελλάδα, έως το 2040. Όπως βλέπουμε, τα νέα περιστατικά καρκίνου του μαστού αναμένεται να ξεπεράσουν τα 8.200 έως το 2040 και οι θάνατοι εξαιτίας της νόσου θα αγγίξουν τους 2.800 [2].



Εικόνα 7: Εκτιμώμενος αριθμός νέων περιστατικών καρκίνου του μαστού και θανάτων σε γυναίκες όλων των ηλικιών στην Ελλάδα έως το 2040 [2]

Το 2018, οι E.Patsea et al. [12] δημοσίευσαν την πρώτη έκθεση μητρώου καρκίνων για την πενταετή περίοδο 2009 - 2013 στην Ελλάδα. Τα μητρώα καρκίνου είναι απαραίτητα στην υγειονομική περίθαλψη, καθώς επιτρέπουν τον ακριβέστερο σχεδιασμό των απαραίτητων υπηρεσιών υγείας και την αξιολόγηση των προγραμμάτων για την πρόληψη και τον έλεγχο του καρκίνου. Η Ελληνική Παθολογική Εταιρεία (ΕΠΕ) έχοντας αναγνωρίσει την έλλειψη αυτών των σημαντικών πληροφοριών ανέλαβε το έργο της δημιουργίας ενός 5ετούς μητρώου καρκίνου με βάση την παθολογία στην Ελλάδα (2009 - 2013). Στη μελέτη αυτή [12], περισσότερα από το 95% όλων των παθολογικών εργαστηρίων των νοσοκομείων του εθνικού συστήματος υγείας και 100% των παθολογικών εργαστηρίων των ιδιωτικών νοσοκομείων, καθώς και περισσότερα από 80% των ιδιωτικών παθολογικών εργαστηρίων έχουν συνεισφέρει τα δεδομένα τους. Οι συχνότεροι τύποι καρκίνου συνολικά ήταν οι εξής: καρκίνος του μαστού (18,26%), καρκίνος του παχέος εντέρου (15,49%), καρκίνος του προστάτη (13,49%) και καρκίνος του πνεύμονα (10,24% όλων των καταγεγραμμένων καρκίνων). Στους άνδρες, οι συχνότερες νεοπλασίες ήταν οι εξής: καρκίνος του προστάτη (13,19%), καρκίνος του παχέος εντέρου (8,77%), καρκίνος του πνεύμονα (8,31%) και καρκίνος του μυελού των οστών (4,39%). Στις γυναίκες, οι συχνότερες νεοπλασίες ήταν: καρκίνος του μαστού (18,07%), καρκίνος του παχέος εντέρου (6,72%), καρκίνος του θυρεοειδούς (4,86%) και καρκίνος του τραχήλου της μήτρας (4,29%). Τα στοιχεία για την επιβάρυνση από τον καρκίνο στην Ελλάδα, που παρουσιάζονται στην μελέτη των E.Patsea et al. [12], καλύπτουν το κενό πληροφοριών για τον καρκίνο στην Ελλάδα, που επηρεάζει την υγειονομική περίθαλψη όχι μόνο σε εθνικό αλλά και σε ευρωπαϊκό επίπεδο.

2.2 Ανατομία και Φυσιολογία του Μαστού

Ο μαστός είναι ένα όργανο του οποίου η μορφή αποτελεί κύριο στοιχείο για τον διαχωρισμό του φύλου ενός ατόμου. Ο γυναικείος μαστός θεωρείται επικουρικό γεννητικό όργανο εξαιτίας του ρόλου που διαδραματίζει τόσο κατά τη διάρκεια της εγκυμοσύνης, όσο και κατά τον θηλασμό. Επιπλέον, αποτελεί βασικό μέρος του ενδοκρινικού συστήματος. Οι γυναίκες διαθέτουν δύο μαστούς στην πρόσθια επιφάνεια της θωρακικής περιοχής, με το τμήμα μεταξύ αυτών να ονομάζεται μεσομάστιος κόλπος. Συγκεκριμένα, ο μαστός είναι μια ημισφαιρική, λεία και λιπώδης προπέτεια, δηλαδή ένα εξόγκωμα που εκτείνεται στην ενήλικη γυναίκα μεταξύ της 2^{ης} και 6^{ης} πλευράς και από το έξω στερνικό χείλος μέχρι την πρόσθια μασχαλιαία γραμμή. Η ανάπτυξη του ξεκινάει στην εφηβεία όπου αυξάνονται τα επίπεδα οιστρογόνων στα θήλεα άτομα. Πιο συγκεκριμένα, ενεργοποιείται η ανάπτυξη του λίπους και του ιστού, με την παράλληλη πάχυνση του συστήματος των γαλακτοφόρων πόρων που θα περιγράψουμε στη συνέχεια. Προς τη μασχάλη, ο μαστός σχηματίζει μια πυραμοειδή προεκβολή η οποία ονομάζεται ουρά του Spence και λανθασμένα συχνά εκλαμβάνεται σαν όγκος καθώς εισέρχεται στην μασχαλιαία κοιλότητα [13].

Η μπροστινή επιφάνεια του μαστού καλύπτεται από λείο και λεπτό δέρμα. Στο κεντρικό τμήμα ενός μαστού βρίσκεται η θηλή, η οποία περιβάλλεται από μία συνήθως σκουρόχρωμη περιοχή, την θηλαία άλω. Το χρώμα αυτό οφείλεται στην άφθονη παρουσία μελανίνης. Η θηλή είναι ένα κωνικό έπαρμα του δέρματος στο οποίο εκβάλλουν οι γαλακτοφόροι πόροι μέσω 15 - 20 στομίων, με αποτέλεσμα να παρατηρούνται στην κορυφή ρωγμές. Το ύψος της θηλής είναι 1 - 1,5 εκατοστά και βρίσκεται λίγο πιο κάτω και έξω από το μέσο του μαστού. Η θέση της θηλής σε σχέση με τον θώρακα εξαρτάται τόσο από το μέγεθος του μαστού, όσο και από το στάδιο χαλάρωσής του [13].

Η θηλαία άλως είναι η υποστρόγγυλη και ελαφρώς επηρμένη περιοχή γύρω από την θηλή και έχει διάμετρο συνήθως 1.5 – 6 εκατοστά. Στην επιφάνειά της παρατηρούνται 10 - 15 αλωαία οζίδια, τα οποία είναι μικρά επάρματα διατεταγμένα σε κύκλο γύρω από την θηλή. Τα οζίδια αυτά έχουν ως υπόθεμα τους αλωαίους αδένες του Montgomery. Οι αλωαίοι αδένες είναι κυρίως οσμηγόνιοι, αλλά και σμηγματογόνοιοι και υποτυπώδεις γαλακτικοί αδένες και υποβοηθούν στη λίπανση και την προστασία της περιοχής της θηλής. Η οπίσθια επιφάνεια του μαστού έχει σχήμα ελαφρώς υπόκοιλο και επικάθεται στην περιτονία του θωρακικού μυός. Τέλος, παρατηρούμε την περιφέρεια του μαστού η οποία μεταβαίνει βαθμιαία στο γύρω δέρμα (βλ. Εικόνα 8) [13].

2.2.1 Η Ανατομία του Μαστού

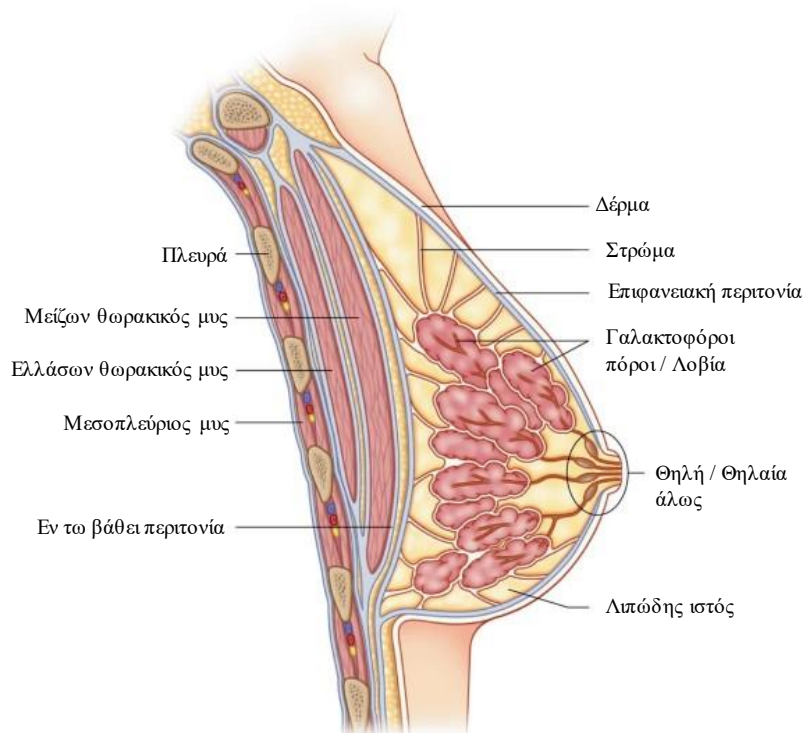
Η περιγραφική ανατομική του μαστού περιλαμβάνει τα παρακάτω:

- ✚ Δέρμα: Το δέρμα στον μαστό είναι λεπτό και διαφανές. Διαθέτει πολλούς σμηγματογόνους και ιδρωτοποιούς αδένες, ενώ συνδέεται στενά με το περιμαστικό λίπος. Στη θηλή και τη θηλαία άλω υπάρχουν, όπως αναφέρθηκε, χρωστικοφόρα κύτταρα, που δίνουν λόγω της μελανίνης την πιο σκούρα χροιά στην περιοχή αυτή. Επιπλέον, στο χόριο του δέρματος σε αυτά τα σημεία παρατηρούνται πολλές λείες μυϊκές ίνες, που προκαλούν σκλήρυνση των μορφωμάτων αυτών κατά την προστριβή ή το θηλασμό.
- ✚ Περιμαστικό Λίπος: Το περιμαστικό λίπος είναι αφθονότερο στο πρόσθιο τμήμα του μαστού και βρίσκεται κάτω από το δέρμα και γύρω από τον μαστικό αδένα. Αποτελεί συνέχεια του υποδόριου λίπους και είναι υποκίτρινο και μαλακό. Στην ανώτερη στοιβάδα του περιμαστικού λίπους, συγκεντρώνονται σε ορισμένα σημεία συνδετικές ίνες, που διέρχονται από τον μαστικό αδένα, και σχηματίζουν τους καλούμενους κρεμαστήρες συνδέσμους του Cooper, οι οποίοι συμβάλλουν στη στήριξη των μαστών συνδέοντας τον εξωτερικό ιστό με το εσωτερικό λίπος. Επιπλέον, καθορίζουν τη σκληρότητα, την έκταση και το σχήμα του μαστού. Οι σύνδεσμοι του Cooper, σε περίπτωση που διηθηθούν από καρκίνο, χάνουν την ελαστικότητά τους με συνέπεια τη βράχυνσή τους. Έτσι

έχουμε εισολκή του δέρματος που είναι σοβαρό κλινικό σημείο, στη διάγνωση του καρκίνου του μαστού.

- ✚ **Μαστικός Αδένας:** Ο μαστικός ή μαζικός αδένας που περιέχεται σε έναν μαστό περιβάλλεται από το περιμαστικό λίπος. Ο αδένας αυτός έχει σχήμα δισκοειδές και μέγεθος που εξαρτάται κυρίως από το αν η γυναίκα είναι άτοκη ή όχι. Έχει επιφάνεια υπόκυρτη και ανώμαλη και είναι σχετικά ελαστικός. Στο πίσω μέρος, ο μαστικός αδένας χωρίζεται από την περιτονία του μείζονα θωρακικού μυός με λίγο ινολιπώδη ιστό. Η περιφέρεια του μαζικού αδένου παρουσιάζει προς τα έξω τη στερνική απόφυση και προς τα έξω την άνω και την κάτω μασχαλιαία απόφυση. Από τις τρεις η τελευταία είναι η μεγαλύτερη. Ο μαστικός αδένας αποτελείται από 15 - 20 λοβούς οι οποίοι έχουν σχήμα πυραμοειδές και χωρίζονται μεταξύ τους από συνδετικές ινώδεις δεσμίδες, που καταλήγουν στους προαναφερθέντες συνδέσμους Cooper. Οι περιοχές μεταξύ των λοβών γεμίζουν με περιμαστικό λίπος, με εξαίρεση το σημείο της θηλής και της θηλαίας άλω. Ο λιπώδης αυτός ιστός δίνει στο μαστό τόσο το περίγραμμα όσο και το μέγεθός του, καθώς σε αυτόν οφείλονται οι διαφορές στο μέγεθος μεταξύ των γυναικών. Κάθε ένας από τους λοβούς του μαστικού αδένου καταλήγει σε έναν κύριο γαλακτοφόρο πόρο, ο οποίος εκβάλλει στην κορυφή της θηλής. Πριν από την εκβολή του, κάθε γαλακτοφόρος πόρος παρουσιάζει ένα ανεύρυσμα, τον γαλακτοφόρο κόλπο, όπου συγκεντρώνεται το γάλα. Επιπλέον, κάθε λοβός υποδιαιρείται, στη συνέχεια, σε 20 - 40 λόβια, διαμορφωμένα σε συμπλέγματα γύρω από μικρούς πόρους και αδενικά λοβίδια, τα οποία είναι υπεύθυνα για την παραγωγή του γάλακτος. Το σύνολο αυτών αποτελεί την τελική πορολοβιακή μονάδα του μαζικού αδένου. Τέλος, το εκκριτικό τμήμα του αδένου αποτελείται από τις αδενοκυψέλες. Οι αδενοκυψέλες που εκβάλλουν σε έναν μεγάλο κλάδο του γαλακτοφόρου πόρου, αποτελούν ένα λόβιο μαζικού αδένου. Τα παραπάνω αδενικά στοιχεία βρίσκονται μέσα σε ένα υπόστρωμα από ινώδη συνδετικό ιστό, που είναι χαλαρότερος γύρω από τις αδενοκυψέλες [14].

Εμφανισιακά μιλώντας, παρατηρούνται διαφορές μεταξύ των ηλικιακών ομάδων, καθώς με την είσοδο μιας γυναίκας στην προ εμμηνόπαυσιακή περίοδο, η έκκριση προγεστερόνης και οιστρογόνων μειώνεται οδηγώντας σε εκφυλισμό των πόρων του μαστού και ατροφία του αδενικού ιστού, με αποτέλεσμα την αντικατάστασή του από ινώδη ιστό που αυξάνει σε πυκνότητα και λίπος. Το γεγονός αυτό κάνει και την διαδικασία της μαστογραφίας πιο εύκολη, λόγω της μαλθακότητας του ιστού. Σε πολλές γυναίκες το στήθος συρρικνώνεται με την πάροδο του χρόνου. Τέλος, η χαλάρωση των συνδέσμων Cooper είναι υπαίτια για την αλλαγή του στήθους στις εμμηνόπαυσιακές γυναίκες [15]. Οι λεπτομέρειες της ανατομίας και φυσιολογίας του μαστού φαίνονται στην Εικόνα 8 που ακολουθεί.



Εικόνα 8: Η ανατομία και φυσιολογία του μαστού [16]

2.2.2 Η Φυσιολογία και η Ανάπτυξη του Μαστού

Ο μαστός είναι ένα ορμονοεξαρτώμενο όργανο. Η λειτουργία και η ανάπτυξή του επηρεάζεται κατά κύριο λόγο από τους εξής παράγοντες [17]:

- ✚ Την υπόφυση, η οποία παράγει την προλακτίνη και την ωκυτοκίνη, ορμόνες που συμμετέχουν στην ανάπτυξη των μαστών. Πιο συγκεκριμένα, η προλακτίνη παίζει ρόλο κατά τη διαφοροποίηση των κυττάρων των τελικών γαλακτοφόρων πόρων, ενώ ελέγχει και τη σύνθεση των πρωτεϊνών του γάλακτος και την αύξηση αλλά και εκκριτική δραστηριότητα των κυττάρων των αδενοκυψελών κατά την εγκυμοσύνη. Η δράση της ωκυτοκίνης είναι η πρόκληση σύσπασης των κυττάρων των γαλακτοφόρων πόρων, συμβάλλοντας στην έξοδο του εκκρινόμενου γάλακτος.
- ✚ Τον υποθάλαμο, ο οποίος εκτός του γεγονότος ότι σχετίζεται με την υπόφυση, ρυθμίζει δύο βασικές ορμόνες του μαστού, την αυξητική και την θυροειδοτρόπο.
- ✚ Τις ωοθήκες, των οποίων η παραγωγή προγεστερόνης και οιστρογόνων συνδέεται στενότερα με τον μαστό. Η μεν προγεστερόνη, σε μικρές δόσεις, προκαλεί την ανάπτυξη των αδενοκυψελών προετοιμάζοντας το στήθος για θηλασμό και σε μεγάλες, έκκριση που μοιάζει με πρωτόγαλα. Τα δε οιστρογόνα βοηθούν στην ανάπτυξη όλων των στοιχείων, συμπεριλαμβανομένων και των μαστών, που κάνουν τη γυναίκα να διαφοροποιείται από τον άνδρα. Η παραγωγή τους μειώνεται με την πάροδο της ηλικίας.

- ✚ Τον θυρεοειδή αδένα, ο οποίος παράγει τη θυροξίνη που μπορεί να αυξήσει την παραγωγή γάλακτος σε περίπτωση υπογαλακτίας, γιατί αυξάνει την έκκριση της προλακτίνης.
- ✚ Τα επινεφρίδια.
- ✚ Τον πλακούντα. Ο πλακούντας μπορεί να θεωρηθεί μεταξύ άλλων και ως ένας ενδοκρινής αδένας και αποτελεί σημαντική πηγή οιστρογόνων ορμονών και χοριακής γοναδοτρόπου ορμόνης, ενώ σε μικρότερες ποσότητες εκκρίνει προγεστερόνη και πλακούντιο γαλακτογόνο ορμόνη. Η τελευταία μοιάζει με την αυξητική ορμόνη της υπόφυσης, αλλά δρα και όπως η προλακτίνη, αφού διεγείρει την παραγωγή γάλακτος και αναπτύσσει τον μαστικό αδένα.

Η ανάπτυξη του μαστού μπορεί να διακριθεί σε πέντε στάδια: το εμβρυικό, το νεογνικό, το εφηβικό, το ενήλικο και το μετεμμηνοπαυσιακό [17]. Πιο αναλυτικά, στο εμβρυικό στάδιο έχουμε αρχικά την ανάπτυξη ενός πρωταρχικού βλαστού περίπου την 4η εβδομάδα, οποίος αναπτύσσεται στη συνέχεια σε έναν στοιχειώδη αδένα, που εκτείνεται από τη μέση μασχαλιαία χώρα έως τη βουβωνική περιοχή. Κατά τη διάρκεια της ενδομήτριας ζωής, ο εμβρυικός μαστός βρίσκεται υπό τον ορμονικό έλεγχο των οιστρογόνων και της προγεστερόνης που παράγονται από τον πλακούντα. Τα οιστρογόνα προάγουν την ανάπτυξη του ιστού και τη δημιουργία του δικτύου των πόρων, ενώ η προγεστερόνη προάγει τη διαφοροποίηση των λοβών και των κυψελίδων. Επιπλέον, και οι δύο αυτές ορμόνες αναστέλλουν τη σύνθεση της προλακτίνης. Ακριβώς μετά τη γέννηση, κατά το νεογνικό δηλαδή στάδιο, λόγω των υψηλών επιπέδων της προλακτίνης στο νεογνό και της έκθεσης του αδένα στις υψηλές συγκεντρώσεις πλακουντιακών οιστρογόνων κατά την κύηση, ο μαστός του νεογνού μπορεί να εκκρίνει μερικές σταγόνες γάλακτος. Στη συνέχεια, ο μαστός γίνεται πιο περίπλοκος, αποκτώντας όλα τα γνωστά ανατομικά στοιχεία εκτός από κυψέλες και παραμένει αδρανής, χωρίς περαιτέρω ωρίμανση, έως την εφηβεία. Σε προεφηβικό επίπεδο, η ανάπτυξη του μαστού ξεκινάει πάλι, προτού τα επίπεδα των ορμονών φύλου να μπορούν να διακριθούν μεταξύ των δύο φύλων [17].

2.3 Η Παθολογία του Μαστού

2.3.1 Καλοήθεις Παθήσεις Μαστού

Η μεγάλη πλειονότητα των αλλοιώσεων που εμφανίζονται στο μαστό είναι καλοήθεις. Οι κακοήθεις αλλοιώσεις εγείρουν μεγάλη ανησυχία επειδή ο καρκίνος του μαστού είναι η πιο συχνή κακοήθεια στις γυναίκες στις δυτικές χώρες, ωστόσο, οι καλοήθεις αλλοιώσεις του μαστού είναι πολύ πιο συχνές από τις κακοήθεις. Είναι γεγονός ότι η πλειονότητα των καλοηθών αλλοιώσεων δεν σχετίζεται με αυξημένο κίνδυνο για επακόλουθο καρκίνο του μαστού, επομένως θα πρέπει να αποφεύγονται οι περιττές χειρουργικές επεμβάσεις. Είναι σημαντικό για τους χειρουργούς, τους ακτινολόγους και τους ογκολόγους να αναγνωρίζουν τις καλοήθεις αλλοιώσεις, τόσο για να τις διακρίνουν από τον *in situ* και τον διηθητικό

καρκίνο του μαστού όσο και για να αξιολογούν τον κίνδυνο μιας ασθενούς να αναπτύξει καρκίνο του μαστού, έτσι ώστε να μπορεί να καθοριστεί ο καταλληλότερος τρόπος θεραπείας για κάθε περίπτωση [18].

Ο όρος «καλοήθεις παθήσεις του μαστού» περιλαμβάνει μια ετερογενή ομάδα αλλοιώσεων που μπορεί να παρουσιάζουν ευρύ φάσμα συμπτωμάτων ή να ανιχνεύονται ως τυχαία μικροσκοπικά ευρήματα. Η εμφάνιση των καλοηθών βλαβών του μαστού αρχίζει να αυξάνεται κατά τη δεύτερη δεκαετία της ζωής και κορυφώνεται κατά την τέταρτη και πέμπτη δεκαετία, σε αντίθεση με τις κακοήθεις παθήσεις, για τις οποίες η εμφάνιση συνεχίζει να αυξάνεται μετά την εμμηνόπαυση, αν και με λιγότερο γρήγορο ρυθμό [18]. Οι κυριότερες καλοήθεις παθήσεις που μπορεί να προκύψουν στο μαστό αναλύονται παρακάτω:

Καλοήθεις ασβεστώσεις

Είναι συνήθως άλατα ασβεστίου που πολύ συχνά ανευρίσκονται στον μαστογραφικό έλεγχο ή μπορεί να αντιστοιχούν σε μια επίσης καλοήθη κατάσταση, που ονομάζεται σκληρυντική αδένωση. Θα πρέπει πάντα να διερευνώνται όταν υπάρχει τάση συρροής (ομαδοποίησης) ή εμφάνιση σε περιοχή του μαστού που δεν προϋπήρχαν, διότι μπορεί να αντιστοιχούν και στην εμφάνιση μιας προκαρκινικής βλάβης [18].

Ινοαδένωμα

Το ινοαδένωμα είναι ο συνηθέστερος καλοήθης όγκος του μαστού. Εμφανίζεται κυρίως κατά την ανάπτυξη των μαστών στην εφηβεία καθώς και σε νεαρές γυναίκες μεταξύ 20 - 30 ετών. Είναι ένας ινοεπιθηλιακός όγκος και θεωρείται μια παρέκκλιση της φυσιολογικής ανάπτυξης των λοβίων. Είναι καλά περιγεγραμμένος όγκος με ομαλή επιφάνεια, δεν έχει σχέση με τον μαζικό αδένα ή τη σύμφυση με το δέρμα. Είναι ανώδυνος, ευκίνητος και συχνά μπορεί να υπάρχει πολλαπλή ανάπτυξη ή και αμφοτερόπλευρη. Το μέγεθός του μπορεί να αυξηθεί κατά την κύηση, τη γαλουχία ή πριν την εμμηνόπαυση. Η διάμετρος του συνήθως είναι 1 - 3 εκ. ενώ σπάνια μπορεί να φτάσει σε σημαντικά μεγάλο μέγεθος 5 - 15 εκ. και τότε ονομάζεται γιγάντιο ινοαδένωμα. Ιστολογικά αποτελείται από επιθηλιακό και ινώδη, συνδετικό ιστό με περισωληναριακή και ενδοσωληναριακή ανάπτυξη. Το ινοαδένωμα δεν εξελίσσεται σε καρκίνο, σπάνια όμως (3 - 5%) μπορεί να βρεθεί καρκίνος στην περιφέρεια ή εντός του ινοαδενώματος, συνήθως πρόκειται για μη διηθητικό λοβιακό καρκίνο (Lobular Carcinoma In Situ - LCIS) και σπανιότερα για μη διηθητικό πορογενές καρκίνωμα (Ductal Carcinoma In Situ - DCIS) [18].

Φυλλοειδής όγκος

Ο φυλλοειδής όγκος ή φυλλοειδές κυστεοσάρκωμα, όταν έχει μικρό μέγεθος, μοιάζει κλινικά και απεικονιστικά με το ινοαδένωμα. Συνήθως όμως αναπτύσσεται γρήγορα και φτάνει σε πολύ μεγαλύτερο μέγεθος. Εμφανίζεται σε ασθενείς ηλικίας 40 - 50 ετών. Από παθολογοανατομικής πλευράς περιέχει στοιχεία τόσο επιθηλιακής όσο και μεσεγχυματικής προέλευσης. Μπορεί να συνοδεύεται από δορυφόρους

όζους με αποτέλεσμα την πολύ συχνή υποτροπή. Στο μεγαλύτερο ποσοστό έχει καλοήγη συμπεριφορά και η απλή εκτομή του επί υγιών ορίων, είναι αρκετή ως θεραπευτική αντιμετώπιση. Σπανιότερα μπορεί να έχει «τοπικά κακοήγη» συμπεριφορά, υψηλό ποσοστό υποτροπής και χρήζει ευρύτερης τοπικής εκτομής ή τεταρτεκτομής. Ακόμα πιο σπάνια μπορεί να είναι κακοήγης, με σαρκωματώδη εξαλλαγή και τότε απαιτείται απλή μαστεκτομή για τον τοπικό έλεγχο της νόσου. Όταν ο φυλλοειδής όγκος διαπιστωθεί παθολογοανατομικά, τότε η ασθενής τίθεται σε μια πιο συχνή (συνήθως ετήσια) κλινικοεργαστηριακή παρακολούθηση από εξειδικευμένο γιατρό [18].

Ενδοπορικό θήλωμα

Τα θηλώματα αναπτύσσονται εντός των εκφορητικών πόρων της θηλής και μπορεί να είναι ψηλαφητά όταν εντοπίζονται σε μεγάλους, κυστικά διατεταμένους πόρους. Το κύριο κλινικό τους σύμπτωμα είναι το ορώδες ή και αιμορραγικό έκκριμα από τη θηλή. Αυτό πρέπει να διαχωρίζεται από την κακοήθεια πίσω από τους πόρους της θηλής, η οποία εκδηλώνεται επίσης με αιμορραγικό έκκριμα, όπως το ενδοπορικό θηλώδες καρκίνωμα [18].

Λιπώματα

Τα λιπώματα του μαστού είναι αρκετά σπάνιοι όγκοι, δεν δημιουργούν πρόβλημα στη διαφορική διάγνωση και αφαιρούνται συνήθως όταν έχουν πολύ μεγάλο μέγεθος κυρίως διότι προκαλούν αισθητικό και όχι διαγνωστικό πρόβλημα [18].

Αμαρτώματα

Είναι σπάνιοι καλοήγεις όγκοι του μαστού, που αποτελούνται από μεικτά στοιχεία αδενικού, λιπώδους και συνδετικού ιστού. Ονομάζονται και αδενολιπώματα και δεν υπάρχουν ενδείξεις κακοήγους εξαλλαγής τους και κατά συνέπεια δεν χρειάζεται η αφαίρεσή τους, πλην ελαχίστων εξαιρέσεων [18].

Κύστες

Οι κύστες στον μαστό είναι ένα πολύ συνηθισμένο εύρημα, μπορεί να είναι πολλαπλές και αρκετά συχνά ψηλαφητές ανάλογα με το μέγεθός τους και το μέγεθος του μαστού. Οι απλές καλοήγεις κύστεις δεν οδηγούν σε κακοήθεια και τις περισσότερες φορές δεν χρειάζονται καμία παρέμβαση. Η πιο συχνή τους επιπλοκή είναι η φλεγμονή και τότε πρέπει να αντιμετωπίζονται με εκκενωτική παρακέντηση και αντιβιοτική αγωγή. Συχνά εμφανίζονται αμφοτερόπλευρα, έχουν ομαλό περίγραμμα και είναι απόλυτα ευκίνητες. Η εξέταση εκλογής για την παρακολούθησή τους είναι το υπερηχογράφημα [18].

Σκληρυντική αδένωση

Πρόκειται για μία καλοήγη κατάσταση η οποία απεικονιστικά εμφανίζεται με τη μορφή διάσπαρτων μικροαποτιτανώσεων και στους δύο μαστούς, μπορεί να αντιστοιχούν σε ψηλαφητή σκληρία και να εκδηλώνονται με μαστοδυνία. Ιδίως σε νέες γυναίκες, προτείνεται τακτική παρακολούθηση και πιο εξειδικευμένες τεχνικές απεικόνισης όπως είναι η τομοσύνθεση σε περιοχές ψηλαφητής σκληρίας ή ακόμα και βιοψία όταν εμφανιστούν περιοχές αποτιτανώσεων με τάση συρροής οι οποίες δεν προϋπήρχαν [18].

Ακτινωτή ουλή

Η ακτινωτή ουλή αναφέρεται και ως σύνθετη σκληρυντική αλλοίωση. Στη μαστογραφία μιμείται την εικόνα του καρκίνου, αλλά δεν πρόκειται για κακοήθεια. Χρήζει χειρουργικής εξαίρεσης χωρίς την ανάγκη περαιτέρω θεραπείας μετά την αφαίρεση. Όπου υπάρχει ακτινωτή ουλή συχνά συνυπάρχουν και άλλες αλλοιώσεις του μαστού. Δεν γνωρίζουμε σαφώς εάν αυξάνει τον κίνδυνο εμφάνισης καρκίνου μαστού [18].

Ινοκυστική μαστοπάθεια

Η ινοκυστική μαστοπάθεια (ή αλλιώς οι ινοκυστικές αλλοιώσεις του μαστού) αποτελεί μια πολύ συνηθισμένη διαταραχή του μαστού. Το 60% των γυναικών εμφανίζουν αυτή τη διαταραχή κλινικά και 90% ιστολογικά. Οφείλεται σε διαταραχή, διέγερση και στη συνέχεια αδυναμία υποστροφής, των κυκλικών αλλοιώσεων που υφίσταται ο μαστός κατά τη διάρκεια της έμμηνου ρύσης [18].

Υπερπλασία

Περιγράφει τον αυξημένο πολλαπλασιασμό των κυττάρων (αυξημένες μιτώσεις), εντός των λοβίων ή των πόρων. Υπάρχουν δύο τύποι υπερπλασίας: χωρίς ατυπία και με ατυπία. Όταν δεν συνυπάρχει ατυπία, η υπερπλασία δεν χρήζει θεραπείας. Επί παρουσίας ατυπίας όμως χρήζει χειρουργικής αφαίρεσης. Η υπερπλασία με παρουσία ατυπίας αυξάνει τον κίνδυνο εμφάνισης καρκίνου του μαστού [18].

Νέκρωση λίπους

Περιγράφει ψηλαφητό ογκίδιο το οποίο συνήθως προκαλείται μετά από τραυματισμό στον μαστό, χειρουργείο ή ακτινοθεραπεία. Σπάνια μπορεί να μην αναφερθεί από τη γυναίκα τραύμα ή άλλο αίτιο. Αποτελείται από νεκρό λιπώδη ιστό και με την πάροδο του χρόνου μπορεί να μετατραπεί σε ουλώδη ιστό. Το δέρμα που περιβάλλει την περιοχή μπορεί να είναι ερυθρό, μελανό ή οιδηματώδες. Είναι συχνότερο σε γυναίκες με μεγάλο μέγεθος μαστού. Μπορεί να προσομοιάζει με καρκίνο στην κλινική εξέταση ή/και στη μαστογραφία και η διαφοροδιάγνωση γίνεται με βιοψία. Δεν χρειάζεται θεραπεία. Μπορεί να απορροφηθεί με την πάροδο του χρόνου. Δεν αυξάνει τον κίνδυνο εμφάνισης καρκίνου του μαστού [18].

Μαστίτιδα

Πρόκειται για επώδυνη φλεγμονή του μαστού και το στήθος είναι ερυθρό και οιδηματώδες. Μπορεί να προκαλέσει και γενικευμένα συμπτώματα λοίμωξης του οργανισμού. Συνήθως εμφανίζεται σε θηλάζουσες γυναίκες λόγω παγίδευσης γάλακτος εντός κάποιου πόρου. Χρήζει θεραπείας με αντιβιοτικά ενώ εάν συνυπάρχει απόστημα τότε μπορεί να χρειαστεί αναρρόφηση με λεπτή βελόνα υπό καθοδήγηση υπερήχου ή ακόμα και χειρουργική παροχέτευση. Δεν αυξάνει τον κίνδυνο εμφάνισης καρκίνου του μαστού [18]

2.3.2 Καλοήθειες και Κακοήθειες Αποτιτανώσεις

Οι αποτιτανώσεις είναι ευρήματα τα οποία εντοπίζονται στην εξέταση της μαστογραφίας και τις περισσότερες φορές δεν συνοδεύονται από ψηλαφητό εύρημα. Το μέγεθός τους δεν ξεπερνά το 1,5 - 2 χιλ. και στη μαστογραφία φαίνονται ως μικρά άσπρα στίγματα ή σαν κόκκοι ζάχαρης. Διακρίνονται σε καλοήθειες και κακοήθειες.

Οι καλοήθειες αποτιτανώσεις οφείλονται κυρίως σε σχηματισμό αλάτων ασβεστίου στη περιοχή λόγω σκληρυντικής αδένωσης, προηγηθείσας χειρουργικής επέμβασης, νέκρωσης λίπους καθώς και άλλων καταστάσεων, που όμως δεν χρειάζονται καμία παρέμβαση παρά μόνο παρακολούθηση [18].

Οι κακοήθειες αποτιτανώσεις είναι μία βλάβη του μαστού, η οποία συνήθως ταυτοποιείται ιστολογικά και ανευρίσκεται στον μαστογραφικό έλεγχο. Η βλάβη στις περιπτώσεις αυτές δεν έχει διηθήσει το επιθήλιο των πόρων και υπάρχει μεγάλη διακύμανση όσον αφορά στο χρονικό διάστημα που μπορεί να συμβεί αυτό. Όταν υπάρξει διήθηση του πόρου τότε έχουμε πλέον διηθητικό καρκίνωμα. Έχει τεράστια σημασία λοιπόν η έγκαιρη ανίχνευση του DCIS διότι είναι μία πάθηση, ίσως και η μοναδική στο μαστό, που σχετίζεται με κακοήθεια, η οποία αν αντιμετωπισθεί έγκαιρα τότε μπορούμε να μιλάμε για απόλυτη ίαση. Σε αυτό έχουν ωφελήσει πολύ οι καινούριες μέθοδοι απεικόνισης και ελέγχου των μαστών που μπορούν να αναδείξουν πρώιμα τέτοιες ύποπτες βλάβες και να απαλλάξουν τις γυναίκες από πολύ σοβαρότερα προβλήματα στο μέλλον. Η εξέταση διαλογής για την ανάδειξη ύποπτων, συρρεουσών μικροαποτιτανώσεων που πιθανόν να αντιστοιχούν σε DCIS, είναι η ψηφιακή μαστογραφία και σπανιότερα ο υπέρηχος ή η μαγνητική τομογραφία (ιδίως όταν το DCIS είναι εκτεταμένο ή συνοδεύεται από ογκόμορφη αλλοίωση). Σπανιότερα μπορεί να είναι τυχαίο εύρημα σε περιπτώσεις όπου γίνεται βιοψία μαστού για κάποια άλλη αλλοίωση. Η θεραπευτική αντιμετώπιση του DCIS διαφέρει ανάλογα με τον τύπο (π.χ. κεντρική νέκρωση Comedotype), το βαθμό κακοήθειας (Grade I, II ή III), την εντόπιση σε ένα ή περισσότερα σημεία του μαστού (πολυκεντρικό ή πολυεστιακό), την έκταση σε σχέση με το μέγεθος του μαστού κ.α. [19].

2.3.3 Ο Καρκίνος του Μαστού

Ο όρος «καρκίνος του μαστού» αναφέρεται στην ορμονοεξαρτώμενη ασθένεια κατά την οποία αναπτύσσεται κακοήθης όγκος στην ευρύτερη περιοχή του μαστού. Οι περισσότερες κακοήθειες του μαστού προέρχονται από επιθηλιακά στοιχεία και κατηγοριοποιούνται ως καρκινώματα. Προσβάλλει τους ιστούς του μαστού, συνήθως τους γαλακτοφόρους πόρους ή τους λοβούς. Ο κακοήθης όγκος σχηματίζεται από άναρχο και ασταμάτητο πολλαπλασιασμό παθολογικών κυττάρων στον μαστό. Πρόκειται για μία κυτταρική νόσο, της οποίας μία από τις χειρότερες επιπλοκές είναι η διήθηση, δηλαδή η διείσδυση και προσβολή γειτονικών ιστών, γεγονός που μπορεί να προκαλέσει σοβαρές βλάβες σε όλον τον οργανισμό. Πιο αναλυτικά, ο καρκίνος ξεκινά από ένα κύτταρο το οποίο έχει μεταλλαχθεί. Ένα κύτταρο θεωρείται καρκινικό όταν χάσει τη δυνατότητά του να σταματήσει να διαιρείται, όταν δεν προσκολλάται σε άλλα κύτταρα ώστε να παραμείνει εκεί για όπου προορίζεται, και όταν δεν πεθαίνει την κατάλληλη βιολογικά ώρα. Το γεγονός αυτό οφείλεται σε μετάλλαξη του γονιδίου που ορίζει τον αναφερόμενο ως «προγραμματισμένο βιολογικό θάνατο» του κυττάρου [20].

2.4 Κατηγοριοποίηση και Τύποι του Καρκίνου του Μαστού

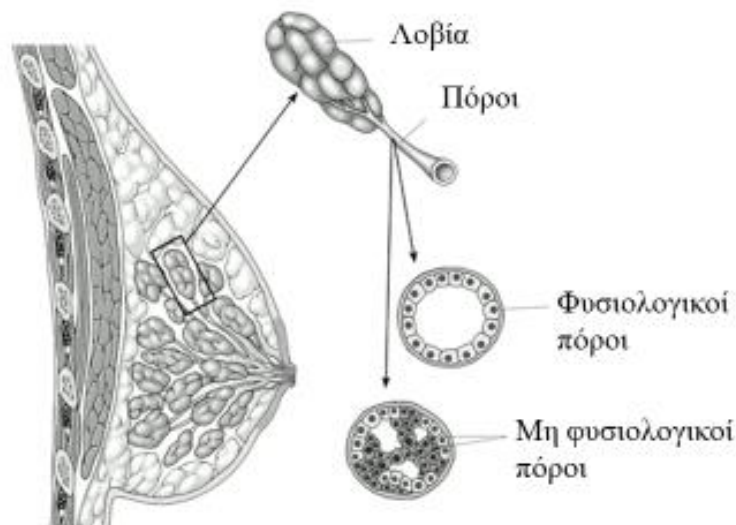
Η ταξινόμηση ενός κακοήθους όγκου του μαστού σε διαφορετικούς τύπους δεν έχει μόνο περιγραφική αξία, αλλά προσδιορίζει την πρόγνωση και συμβάλλει στην κατανόηση του καρκίνου του μαστού. Τα είδη του καρκίνου του μαστού είναι πολλά και εξαρτώνται από τον τρόπο με τον οποίο γίνεται η κατάταξη μιας περίπτωσης. Οι καρκίνοι του μαστού κατηγοριοποιούνται βάσει του τύπου του ιστού από τον οποίο διασπείρονται τα καρκινικά κύτταρα, αλλά και την εξάπλωσή τους. Περαιτέρω ταξινόμηση μπορεί να γίνει ανάλογα με τον βαθμό κακοήθειας ή με τον βαθμό διαφοροποίησης των κυττάρων. Επιπλέον, σημαντικό ρόλο παίζει το μέγεθος του όγκου, η ύπαρξη λεμφαδενικών μεταστάσεων, καθώς και το επίπεδο έκφρασης ορισμένων γονιδιακών παραγόντων [20].

2.4.1 Ιστολογική Ταξινόμηση του Καρκίνου του Μαστού

Η μορφολογική ταξινόμηση γίνεται, αρχικά, σε μη διηθητικούς (in situ - δηλαδή στην αρχική τους θέση) και σε διηθητικούς καρκίνους, ανάλογα με το αν εντοπίζεται διήθηση των ιστών γύρω από τον όγκο. Όσον αφορά στην προέλευση των καρκινικών κυττάρων, η διάκριση γίνεται μεταξύ πορογενών ή λοβιακών καρκινωμάτων. Οι περισσότεροι καρκίνοι του μαστού είναι αδενοκαρκινώματα και αναπτύσσονται στο τελικό τμήμα των πόρων του μαστού. Ο καρκίνος του μαστού μπορεί να παραμένει μη διηθητικός ή να γίνει διηθητικός χωρίς μεταστάσεις για μεγάλες χρονικές περιόδους [20].

Οι πιο συνηθισμένοι τύπο καρκίνου του μαστού είναι οι εξής:

- ✚ **Μη διηθητικό πορογενές καρκίνωμα του μαστού (Ductal Carcinoma In Situ - DCIS).** Πρόκειται για ένα πρώιμο στάδιο καρκίνου του μαστού, μη διηθητικό, και αποτελεί περίπου το ένα πέμπτο (1/5) των νέων περιστατικών που διαγιγνώσκονται. Στην μορφή αυτή τα καρκινικά κύτταρα εντοπίζονται μόνο στον γαλακτοφόρο πόρο και δεν υπάρχει εξάπλωση. Η διάγνωση σε αυτό το στάδιο σχεδόν εγγυάται την αποτελεσματική αντιμετώπιση και τις περισσότερες φορές την ίαση της ασθένειας [21].



Εικόνα 9: Μορφή κυττάρων σε τύπο καρκίνου DCIS [21]

- ✚ **Διηθητικό πορογενές καρκίνωμα του μαστού (Invasive Ductal Carcinoma - IDC).** Ο τύπος IDC είναι ο πιο συνηθισμένος, με οχτώ στις δέκα (8/10) περιπτώσεις διηθητικού καρκίνου να είναι αυτής της μορφής. Στην περίπτωση αυτή, η καρκινογένεση ξεκινάει στην εσωτερική μεμβράνη του γαλακτοφόρου πόρου, ο οποίος διοχετεύει το γάλα στη θηλή, αλλά τα καρκινικά κύτταρα εξαπλώνονται εκτός του πόρου και περνούν στον λιπώδη ιστό του μαστού. Στα καρκινώματα αυτά τη μεγαλύτερη προγνωστική αξία έχει ο βαθμός πυρηνικής διαφοροποίησης. Στον IDC τύπο, ο κίνδυνος για τα κύτταρα αυτά να εξαπλωθούν σε άλλα όργανα του σώματος, μέσω του λεμφικού συστήματος και της κυκλοφορίας του αίματος, είναι αρκετά αυξημένος [22].
- ✚ **Διηθητικό λοβιακό καρκίνωμα του μαστού (Invasive Lobular Carcinoma - ILC).** Είναι ο δεύτερος πιο συνηθισμένος τύπος μετά τον IDC, με ποσοστό περίπου 10%. Εδώ τα καρκινικά κύτταρα δημιουργούνται στα λόβια, τους αδένες που παράγουν το γάλα, και έπειτα εξαπλώνονται στους γειτονικούς ιστούς του μαστού. Παρότι μπορεί να προσβάλλει γυναίκες όλων των ηλικιών,

συναντάται πιο συχνά σε μεγαλύτερες γυναίκες. Μερικές έρευνες έχουν συμπεράνει ότι η ορμονική θεραπεία κατά την διάρκεια ή μετά την εμμηνόπαυση αυξάνει τον κίνδυνο εμφάνισης ILC [23].

Από τους πιο σπάνιους τύπους καρκίνου του μαστού είναι τα εξής:

- ✚ **Μη διηθητικό λοβιακό καρκίνωμα του μαστού (*Lobular Carcinoma In Situ - LCIS*)**. Πρόκειται για μία ασυνήθιστη πάθηση που χαρακτηρίζεται από την ανάπτυξη και παρουσία μη φυσιολογικών κυττάρων στους λοβούς. Παρότι χρησιμοποιείται ο όρος καρκίνωμα, δεν είναι πραγματικός καρκίνος. Υποδηλώνει, ωστόσο, τον αυξημένο κίνδυνο ένα άτομο να αναπτύξει μελλοντικά κάποια μορφή διηθητικού καρκίνου. Επειδή δεν έχει συμπτώματα και δεν είναι εμφανές συνήθως στην μαστογραφία, δεν είναι βέβαιο το ποσοστό των ατόμων που εμφανίζουν τη πάθηση αυτή [24].
- ✚ Δύο άλλες άτυπες μορφές πορογενούς καρκίνου του μαστού είναι ο **φλεγμονώδης καρκίνος του μαστού (*Inflammatory Breast Cancer - IBC*)** [25] και η **Νόσος Paget της θηλής** [26]. Ο φλεγμονώδης καρκίνος του μαστού είναι μια συστηματική νόσος και είναι ο πιο επιθετικός από όλους τους τύπους καρκίνου. Παρουσιάζεται σαν οίδημα του δέρματος του μαστού και δίνει όψη φλοιού πορτοκαλιού, ενώ προκαλείται ερυθρότητα και αίσθηση καύσου. Επειδή συνήθως δεν υπάρχει μόνο ένα ογκίδιο ή όγκος, ορισμένες φορές ερμηνεύεται εσφαλμένα ως μαστίτιδα ή ως κάποια φλεγμονή στο στήθος. Η νόσος του Paget της θηλής του μαστού είναι ένας σπάνιος τύπος καρκίνου του μαστού, ο οποίος ξεκινάει από τους πόρους του μαστού και έπειτα διηθεί το επιθήλιο της θηλής. Χαρακτηρίζεται ιστολογικά από ανάπτυξη των νεοπλασματικών κυττάρων μέσα στην επιδερμίδα του δέρματος της θηλής. Συχνά παρατηρείται ξεφλούδισμα, φολίδωση και ερυθρότητα της επιδερμίδας γύρω από τη θηλή, με σημεία εξέλκωσης ή εκροής αιμορραγικού υγρού.

2.4.2 Βαθμός Κακοήθειας ή Διαφοροποίησης – Tumor Grade

Ο προσδιορισμός του βαθμού κακοήθειας είναι σημαντικό στοιχείο προγνώσεως. Βασίζεται στην ετερογένεια των καρκινικών κυττάρων, στην αρχιτεκτονική δομή του ιστού που σχηματίζουν και τη συχνότητα των μιτώσεων (κυτταρική διαίρεση) των καρκινικών κυττάρων. Ουσιαστικά, γίνεται μια σύγκριση μεταξύ της εμφάνισης των καρκινικών κυττάρων με αυτή των υγίων. Ένας καλά διαφοροποιημένος καρκινικός όγκος χαρακτηρίζεται ως βαθμού 1 (grade 1) και έχει χαμηλή ετερογένεια των κυττάρων, διατηρημένη αρχιτεκτονική δομή και λίγες μιτώσεις. Ένας αδιαφοροποίητος όγκος βαθμού 3 (grade 3) έχει υψηλή ετερογένεια, απώλεια της αρχιτεκτονικής και πολλές μιτώσεις. Ένας μέτρια διαφοροποιημένος όγκος βαθμού 2 (grade 2) είναι μεταξύ των grade 1 και grade 3. Όσο μικρότερος είναι ο βαθμός διαφοροποίησης, τόσο καλύτερη είναι η πρόγνωση. Πιο συγκεκριμένα, ο βαθμός κακοήθειας είναι σημαντικός δείκτης επιβίωσης των ασθενών με καρκίνο μαστού. Ασθενείς με βαθμό κακοήθειας 1

έχουν πάνω από 85% πιθανότητα δεκαετούς επιβίωσης σε αντίθεση με τις ασθενείς με βαθμό κακοήθειας 3, όπου το ποσοστό είναι μικρότερο του 45% [27].

2.4.3 Υπότυποι και Μοριακή Ταξινόμηση του Καρκίνου του Μαστού

Η ανάλυση του γονιδιακού προφίλ του καρκίνου βοηθά ώστε να προσδιοριστούν με μεγαλύτερη σαφήνεια τα ιδιαίτερα χαρακτηριστικά της νόσου. Έτσι μπορεί και η θεραπευτική αντιμετώπιση να είναι στοχευμένη και αποτελεσματική. Υπάρχουν πέντε κύριοι εγγενείς ή μοριακοί υπότυποι καρκίνου του μαστού που βασίζονται στα γονίδια που εκφράζει ο καρκίνος [28]:

- i. **Luminal A.** Αυτός ο τύπος αφορά στο 40% των καρκίνων του μαστού, έχει θετικούς ορμονικούς υποδοχείς οιστρογόνων και προγεστερόνης, είναι HER2 (Human Epidermal growth factor Receptor) αρνητικός και έχει χαμηλά επίπεδα πρωτεΐνης Ki-67, που βοηθά στον έλεγχο του ρυθμού ανάπτυξης των καρκινικών κυττάρων. Πρόκειται για καρκίνους μικρού βαθμού κακοήθειας, τείνουν να αναπτύσσονται αργά και έχουν την καλύτερη πρόγνωση.
- ii. **Luminal B.** Πρόκειται για καρκίνο με θετικούς ορμονικούς υποδοχείς οιστρογόνων και προγεστερόνης και είναι είτε HER2 θετικός είτε HER2 αρνητικός, αλλά με υψηλά επίπεδα Ki-67. Η μορφή αυτή συναντάται σε ποσοστό 20% των περιπτώσεων, είναι ελαφρώς πιο σοβαρή από την Luminal A, με ταχύτερη ανάπτυξη και ελάχιστα πιο δυσμενή πρόγνωση.
- iii. Ο τύπος καρκίνου του μαστού **Τριπλά Αρνητικός/Βασικού Τύπου (Triple Negative/Basal-like)** έχει αρνητικούς ορμονικούς υποδοχείς, είναι HER2 αρνητικός και είναι κακής πρόγνωσης. Χαρακτηρίζεται από πολύ υψηλό βαθμό έκφρασης γονιδίων πολλαπλασιασμού. Αυτός ο τύπος καρκίνου είναι πιο συχνός σε γυναίκες με μεταλλάξεις γονιδίων BRCA1. Χωρίς να είναι γνωστή η αιτία, αυτός ο τύπος καρκίνου είναι πιο κοινός μεταξύ των νεαρών γυναικών και των αφροαμερικανών γυναικών.
- iv. Οι **πλούσιοι σε HER2 (HER2-enriched)** υπότυποι αφορούν στο 10% - 15% των περιπτώσεων. Είναι καρκίνος του μαστού με αρνητικούς υποδοχείς ορμονών, αλλά HER2 θετικός. Οι εμπλουτισμένοι με HER2 καρκίνοι τείνουν να αναπτύσσονται ταχύτερα από τους καρκίνους τύπου Luminal και έχουν αρκετά χειρότερη πρόγνωση. Ωστόσο, συχνά αντιμετωπίζονται επιτυχώς με θεραπείες που στοχεύουν στην πρωτεΐνη HER2.
- v. **Normal-like.** Πρόκειται για έναν υπότυπο παρόμοιο με τον Luminal A. Έχει, επομένως, θετικούς ορμονικούς υποδοχείς, είναι HER2 αρνητικός και έχει χαμηλά επίπεδα πρωτεΐνης Ki-67. Η πρόγνωση στην μορφή αυτή είναι καλή, αλλά παραμένει χειρότερη από εκείνη μιας Luminal A περίπτωσης.

2.4.4 Σταδιοποίηση του Καρκίνου του Μαστού

Προκειμένου να γίνει ολόπλευρη εκτίμηση του τύπου του καρκίνου του μαστού και να προσδιοριστούν οι κίνδυνοι, καθώς και η πρόγνωση για την/τον ασθενή, χρησιμοποιείται το σύστημα **TNM**. Το T αναφέρεται στον όγκο από την αγγλική λέξη Tumor και το μέγεθος που έχει τη στιγμή της διάγνωσης. Το N εκφράζει το αν καρκινικά κύτταρα έχουν διηθηθεί σε λεμφαδένες (Lymph Nodes), ενώ η ύπαρξη ή μη μετάστασης (Metastasis), δηλαδή η διασπορά της νόσου σε άλλα όργανα, δηλώνεται με το M.

Το στάδιο του καρκίνου είναι θεμελιώδες για τις αποφάσεις που αφορούν στη θεραπεία και την αντιμετώπιση της νόσου. Όσο λιγότερο προχωρημένο είναι το στάδιο, τόσο καλύτερη είναι η πρόγνωση. Η σταδιοποίηση συνήθως γίνεται δύο φορές: μετά την κλινική και ακτινολογική εξέταση, όπως επίσης και μετά το χειρουργείο. Εάν πραγματοποιηθεί χειρουργείο, η σταδιοποίηση μπορεί, επίσης, να επηρεαστεί από την εργαστηριακή εξέταση του αφαιρούμενου όγκου και των λεμφαδένων.

Υπάρχουν 4 στάδια, από 0 έως 4 (0 - IV) που καθορίζονται από το συνδυασμό των παραπάνω παραγόντων αλλά και από το αν ο καρκίνος είναι διηθητικός ή όχι. Τα στάδια, μαζί με τις αντίστοιχες υποκατηγορίες τους, παρουσιάζονται στον παρακάτω πίνακα [29]:

Πίνακας 1: Τα στάδια του καρκίνου του μαστού

ΣΤΑΔΙΟ	ΟΡΙΣΜΟΣ
Στάδιο 0	0 Μη διηθητικό πορογενές καρκίνωμα.
Στάδιο I	Ia Η διάμετρος του όγκου είναι μέχρι 2 εκ. και έχει αρνητικό λεμφαδένα φρουρό.
	Ib Η διάμετρος του όγκου είναι μέχρι 2 εκ. και υπάρχει μικρομετάσταση (0,2mm - 2mm) στον λεμφαδένα φρουρό.
Στάδιο II	IIa Ο όγκος έχει διάμετρο μέχρι 2 εκ. και θετικό για μετάσταση φρουρό λεμφαδένα ή ο όγκος έχει διάμετρο 2 - 5 εκ. με αρνητικό φρουρό λεμφαδένα.

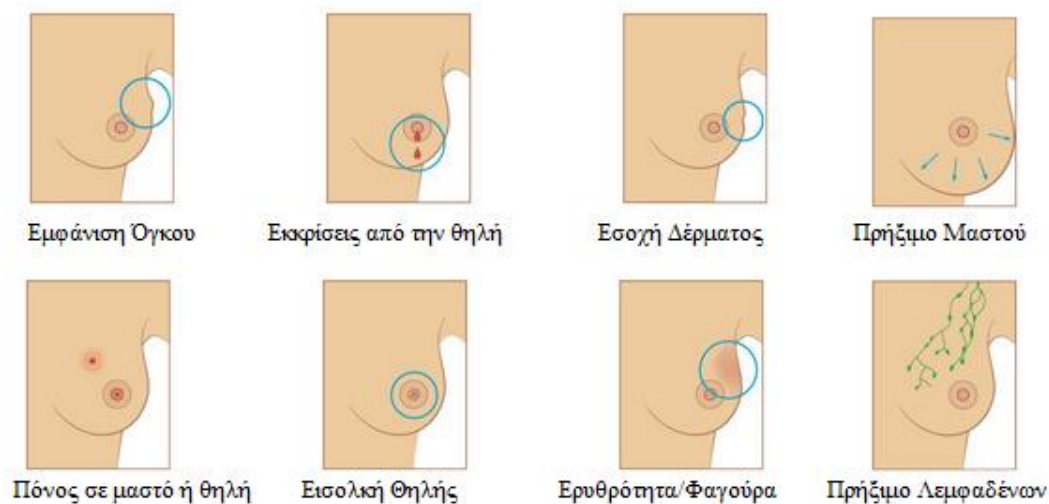
	<p>IIIb</p> <p>Όγκος 2 - 5 εκ. με μικρομετάσταση στον λεμφαδένα φρουρό ή όγκος 2 - 5 εκ. με 1 - 3 θετικούς λεμφαδένες ή όγκος μεγαλύτερος από 5 εκ. με αρνητικό λεμφαδένα φρουρό.</p>
Στάδιο III	<p>IIIa</p> <p>Όγκος ανεξαρτήτως μεγέθους με 4-9 θετικούς μασχάλιους λεμφαδένες ή όγκος μεγαλύτερος των 5 εκ. με μικρομετάσταση στον λεμφαδένα φρουρό ή όγκος μεγαλύτερος από 5 εκ. με 1 - 3 θετικούς μασχάλιους λεμφαδένες.</p>
	<p>IIIb</p> <p>Όγκος ανεξαρτήτως μεγέθους με διήθηση του θωρακικού τοιχώματος ή του δέρματος του μαστού (συμπεριλαμβάνεται και ο φλεγμονώδης καρκίνος μαστού) ή διήθηση σε περισσότερους από 9 λεμφαδένες.</p>
	<p>IIIc</p> <p>Διήθηση δέρματος ή θωρακικού τοιχώματος και μετάσταση σε 10 ή περισσότερους μασχάλιους λεμφαδένες ή μετάσταση σε υπερκλειδίους λεμφαδένες.</p>
Στάδιο IV	<p>IV</p> <p>Διηθητική κακοήθεια μαστού και μετάσταση σε άλλο όργανο του σώματος.</p>

Οι πιθανότητες για επιβίωση είναι διαφορετικές για κάθε στάδιο του καρκίνου του μαστού. Οι μη διηθητικοί καρκίνοι του μαστού, αλλά και ο διηθητικός πρώιμος καρκίνος μαστού (στάδια I και II), έχουν καλύτερη πρόγνωση από όγκους σε πιο προχωρημένο στάδιο (στάδιο III και IV). Η έγκαιρη διάγνωση είναι μείζονος σημασίας στην εξέλιξη της νόσου [29].

2.5 Συμπτώματα του Καρκίνου του Μαστού

Το βασικότερο βήμα για τη διατήρηση της υγιούς κατάστασης των μαστών μιας γυναίκας είναι να γνωρίζει η ίδια τη φυσιολογική όψη και υφή τους. Τα συμπτώματα και πρώτα δείγματα εμφάνισης καρκίνου στους μαστούς μπορεί να διαφέρουν πολύ σε κάθε περίπτωση. Αυτό οφείλεται στην ετερογένεια της ασθένειας που συνεπάγεται διαφορετική κλινική εκδήλωση. Το μεγαλύτερο ποσοστό των καρκίνων του μαστού διαγιγνώσκονται μέσω απεικονιστικών μεθόδων και κατά τη διάρκεια εξετάσεων. Ωστόσο τα πιο συνηθισμένα σημάδια που πιθανώς υποδηλώνουν την ύπαρξη καρκίνου του μαστού είναι τα εξής [30]:

- ✚ Το βασικότερο σημάδι είναι ένας νέος όγκος ή μάζα που έχει διαφορετική αίσθηση από τον περιβάλλοντα ιστό. Μια μη επώδυνη και σκληρή ακανόνιστη μάζα είναι πιο πιθανό να είναι καρκίνωμα, σε σχέση με ένα στρογγυλό και μαλακό όζο, όμως υπάρχουν και εξαιρέσεις.
- ✚ Ακόμα και στην περίπτωση που δεν υπάρχει εμφανής μάζα, μπορεί να διαπιστωθεί πρήξιμο σε μέρος ή και σε ολόκληρο τον μαστό. Το πρήξιμο αυτό οδηγεί σε αλλαγή του μεγέθους και της συνήθους μορφής του μαστού και σε αίσθηση τραβήγματος στο δέρμα.
- ✚ Η εσοχή ή βαθούλωμα σε οποιοδήποτε σημείο του μαστού.
- ✚ Η αλλαγή στην όψη και την υφή του δέρματος, όπως για παράδειγμα πάχυνση σε σημεία, ή δερματικές αλλοιώσεις και ερεθισμός παρόμοιος αφυδάτωσης, ιδιαίτερα κοντά στη θηλή και τη θηλαία άλω, πιθανώς να είναι δείγμα ορισμένων μορφών καρκίνου.
- ✚ Στις δερματικές αλλοιώσεις συμπεριλαμβάνονται και η ερυθρότητα, η φαγούρα του μαστού και της θηλής, καθώς και το ξεφλούδισμα και η όψη φλοιού πορτοκαλιού.
- ✚ Ο καρκίνος του μαστού προκαλεί αλλαγές στα κύτταρα του δέρματος δημιουργώντας συχνά πόνο και ευαισθησία στον μαστό και τη θηλή.
- ✚ Η εισολκή της θηλής είναι ένα σύνθηρες δείγμα της νόσου.
- ✚ Η έκκριση υγρού από τη θηλή σε κατάσταση μη γαλουχίας δεν είναι φυσιολογική και παρότι μπορεί να οφείλεται σε πολλούς μη καρκινικούς παράγοντες, δεν πρέπει να αγνοείται.
- ✚ Το ασυνήθιστο πρήξιμο ή ένα νέο ογκίδιο στην μασχαλιαία περιοχή και τους λεμφαδένες μπορεί να είναι δείγμα διήθησης καρκινικών κυττάρων. Επίσης παρατηρούνται αλλαγές και στην περιοχή της κλείδας.



Εικόνα 10: Συμπτώματα του καρκίνου του μαστού [31]

Πολλά από τα παραπάνω μπορεί να έχουν αθώα ή συνηθισμένα αίτια, όπως κάποια μόλυνση, ωστόσο δεν πρέπει να αμελείται σε καμία περίπτωση η εξέταση και ο έλεγχος τους από τον ειδικό κλινικό ιατρό. Το πιο συνηθισμένο σημάδι καρκίνου του μαστού που μπορεί να παρατηρηθεί από την ασθενή είναι ένα ανώδυνο εξόγκωμα στο στήθος, που ενδεχομένως υποδηλώνει όγκο. Είναι γεγονός ότι κατά τη στιγμή της διάγνωσης, οι περισσότερες γυναίκες δεν έχουν εμφανή σημάδια καρκίνου του μαστού. Αντιθέτως, τα πρώιμα σημάδια του ανακαλύπτονται σε μια απεικονιστική εξέταση. Όταν ο όγκος είναι μικρός στα αρχικά στάδια του καρκίνου του μαστού, σπάνια γίνεται αντιληπτός με την αφή ή με γυμνό μάτι. Αυτός είναι ο λόγος για τον οποίο οι μαστογραφίες παίζουν τόσο σημαντικό ρόλο στη διάγνωση του καρκίνου του μαστού [32].

Οι μαστογραφίες μπορούν να βοηθήσουν στην ανίχνευση ενός όγκου στο μαστό πολύ πριν αυτός γίνει αρκετά μεγάλος ώστε να αρχίσει να προκαλεί συμπτώματα και σε αυτή την περίπτωση ο καρκίνος είναι πιθανό να είναι πιο εύκολα αντιμετωπίσιμος. Οι όγκοι μπορεί να είναι τόσο μικροί όσο η άκρη ενός μολυβιού (1mm) ή τόσο μεγάλοι όσο ένα λεμόνι (50mm). Οι μαστογραφίες έχουν τη δυνατότητα να εντοπίζουν όγκους στο μικρότερο άκρο του φάσματος, όταν δεν μπορούν ακόμη να φανούν ή να ψηλαφηθούν. Οι τακτικές μαστογραφίες είναι ο πιο αξιόπιστος τρόπος για την έγκαιρη διάγνωση του καρκίνου του μαστού, σύμφωνα με την Αμερικανική Αντικαρκινική Εταιρεία, παρόλα αυτά οι μαστογραφίες δεν μπορούν να εντοπίσουν κάθε περίπτωση καρκίνου του μαστού, γι' αυτό είναι σημαντικό οι γυναίκες να δίνουν σημασία στις αλλαγές στο στήθος τους.

2.6 Αιτιολογία του Καρκίνου του Μαστού και Παράγοντες Κινδύνου

Ακόμα και σήμερα δεν είναι ξεκάθαροι οι λόγοι που αναπτύσσεται ο καρκίνος του μαστού. Σύμφωνα με την Εταιρεία Ουκολόγων Παθολόγων Ελλάδος, πρόκειται για μια ασθένεια που δεν έχει σαφή αιτιολογία, ούτε μία μόνο συγκεκριμένη μορφή [33]. Είναι γνωστό ότι ορισμένες γυναίκες έχουν περισσότερες πιθανότητες να νοσήσουν από άλλες. Μέσω εντατικής και διεπιστημονικής έρευνας έχουν βρεθεί και προσδιοριστεί κάποιοι παράγοντες κινδύνου που επηρεάζουν την εμφάνιση της νόσου.

Ένας παράγοντας κινδύνου αυξάνει τον κίνδυνο εμφάνισης του καρκίνου, αλλά δεν είναι ούτε αναγκαίος ούτε και αρκετός ώστε να προκαλέσει καρκίνο από μόνος του. Παρόλα αυτά, η γνώση και μελέτη των παραγόντων αυτών μπορεί να βοηθήσει στην αξιολόγηση των γυναικών υψηλού κινδύνου, ώστε αυτές να παρακολουθούνται σχολαστικά και να αντιμετωπίζονται έγκαιρα. Επιπλέον, η διερεύνηση των παραγόντων κινδύνου θα συνδράμει στην ευρύτερη κατανόηση της πολυπαραγοντικής αυτής νόσου. Οι πιο σημαντικοί παράγοντες κινδύνου που προέκυψαν μετά από εκτεταμένη μελέτη της διεθνούς βιβλιογραφίας αναλύονται στη συνέχεια του υποκεφαλαίου.

Ηλικία και αναπαραγωγικοί παράγοντες

Ηλικία

Πολλές μελέτες έδειξαν ότι από τη στιγμή που γίνεται η διάγνωση του καρκίνου του μαστού, η νεότερη ηλικία αποτελεί αρνητικό παράγοντα κινδύνου καθώς φαίνεται πως οι νεαρές γυναίκες αποτελούν περιπτώσεις με λιγότερες πιθανότητες επιβίωσης καθώς παρουσιάζουν μεγαλύτερους όγκους σε πιο προχωρημένο στάδιο και συχνά έχουν θετικούς λεμφαδένες [34]. Σε έρευνα που διεξήχθη από τους A. Toriola et al. [35], στις Ηνωμένες Πολιτείες Αμερικής, το κατά προσέγγιση ηλικιακό όριο με τη χειρότερη πρόγνωση μετά την εμφάνιση καρκίνου του μαστού είναι τα 45 χρόνια. Επίσης, εκτίμησαν ότι η μέση πιθανότητα μιας 30χρονης γυναίκας να αναπτύξει καρκίνο του μαστού σε μια περίοδο 10 ετών είναι 1 στα 232 (0,43%), περίπου 10 φορές μικρότερη από αυτήν που έχει μια 70χρονη γυναίκα, με 1 στις 27 πιθανότητες (3,74%) [35].

Η ερευνητική ομάδα των C. DeSantis et al. [36] διαπίστωσε ότι μετά από μια αργή αύξηση που παρατηρήθηκε από το 1975 έως το 1990 (0,4% ετησίως), τα ποσοστά θανάτου από καρκίνο του μαστού ετησίως μειώθηκαν κατά 2,2% κατά την περίοδο 1990 - 2007 για γυναίκες όλων των ηλικιών. Αυτή η μείωση ήταν υψηλότερη στις γυναίκες ηλικίας κάτω των 50 ετών (3,2% ετησίως) σε σύγκριση με τις γυναίκες μεγαλύτερης ηλικίας [36].

Ο Β. Yankaskas και οι συνεργάτες του [37], συγκέντρωσαν δεδομένα από έξι διαφορετικά Κέντρα Μαστού στις Ηνωμένες Πολιτείες, μέλη της Κοινοπραξίας για την επιτήρηση του καρκίνου του μαστού. Στην έρευνα συμπεριελήφθησαν 117.738 γυναίκες ηλικίας 18 - 39 ετών που έκαναν την πρώτη τους διαγνωστική μαστογραφία κατά την περίοδο 1995 - 2005 και παρακολούθηθηκαν για 1 έτος μετά τη διεξαγωγή της προκειμένου να προσδιοριστεί η ακρίβεια της αξιολόγησης της μαστογραφίας. Στην έρευνα μετρήθηκε το ποσοστό των γυναικών που κλήθηκαν να κάνουν συμπληρωματικές απεικονιστικές εξετάσεις, όπως επίσης και η ευαισθησία, η ειδικότητα, η θετική προγνωστική αξία και το ποσοστό ανίχνευσης καρκίνου για όλες τις μαστογραφίες που διεξήχθησαν.

Κατά τον προσυμπτωματικό έλεγχο, δεν ανιχνεύθηκε κανένας καρκίνος σε 637 μαστογραφίες για γυναίκες ηλικίας 18 - 24 ετών. Για τις γυναίκες ηλικίας 35 - 39 ετών που είχαν τον μεγαλύτερο αριθμό μαστογραφιών προσυμπτωματικού ελέγχου (n = 73.335) σε αυτή τη μελέτη, το ποσοστό ανάκλησης ήταν 12,7% (95% CI = 12,4% έως 12,9%), η ευαισθησία ήταν 76,1% (95% CI = 69,2% έως 82, 6%), η ειδικότητα ήταν 87,5% (95% CI = 87,2% έως 87,7%), η θετική προγνωστική αξία ήταν 1,3% (95% CI = 1,1% έως 1,5%) και το ποσοστό ανίχνευσης καρκίνου ήταν 1,6 καρκίνοι ανά 1000 μαστογραφίες (95% CI = 1,3 έως 1,9 καρκίνοι ανά 1000 μαστογραφίες). Οι περισσότερες (67.468 [77,7%]) από τις 86.871 γυναίκες που εξετάστηκαν ανέφεραν ότι δεν είχαν οικογενειακό ιστορικό καρκίνου του μαστού. Για τις διαγνωστικές μαστογραφίες, τα προσαρμοσμένα στην ηλικία ποσοστά σε όλες τις ηλικιακές ομάδες ήταν: ευαισθησία 85,7% (95% CI = 82,7% έως 88,7%), ειδικότητα 88,8% (95% CI = 88,4% έως 89,1%), θετική προγνωστική αξία 14,6% (95% CI = 13,3% έως 15,8%) και ποσοστό ανίχνευσης καρκίνου 14,3 καρκίνοι ανά 1.000 μαστογραφίες (95% CI = 13,0 έως 15,7 καρκίνοι ανά 1.000 μαστογραφίες). Η απόδοση της μαστογραφίας, εκτός από την ειδικότητα, βελτιώθηκε με την παρουσία όγκου στο μαστό.

Οι Β. Yankaskas et al. [37] κατέληξαν στο ότι οι νεότερες γυναίκες έχουν πολύ χαμηλά ποσοστά εμφάνισης καρκίνου του μαστού και μετά τη μαστογραφία εμφανίζουν υψηλά ποσοστά ανάκλησης και υψηλά ποσοστά πρόσθετων απεικονιστικών εξετάσεων. Στην εν λόγω έρευνα, δεν διαπιστώθηκε καρκίνος σε γυναίκες ηλικίας κάτω των 25 ετών ενώ διαπιστώθηκαν χαμηλά ποσοστά εμφάνισης της νόσου για τη μεγάλη ομάδα γυναικών ηλικίας 35 - 39 ετών. Διαπιστώθηκε λοιπόν, ότι σε αναγωγή σε έναν θεωρητικό πληθυσμό 10.000 γυναικών ηλικίας 35 - 39 ετών, 1.266 γυναίκες που υποβάλλονται σε προληπτικό έλεγχο θα υποβληθούν σε περαιτέρω εξετάσεις, με 16 καρκίνους να ανιχνεύονται και 1.250 γυναίκες να λαμβάνουν ψευδώς θετικό αποτέλεσμα [37].

Εμμηνος ρύση, αριθμός τοκετών, ηλικία τεκνοποίησης, θηλασμός

Η γνώση της επιστημονικής κοινότητας σχετικά με τους παράγοντες κινδύνου για την εμφάνιση του καρκίνου του μαστού, ειδικά όσον αφορά σε αναπαραγωγικούς παράγοντες, έχει εξελιχθεί πολύ τα τελευταία χρόνια.

Σύμφωνα με τους Nelson et al. [38] οι όλο και περισσότεροι τοκετοί μπορούν δυνητικά να μειώσουν τον κίνδυνο εμφάνισης του καρκίνου του μαστού, συγκεκριμένα κάθε τοκετός μπορεί να μειώσει τον κίνδυνο καρκίνου του μαστού, ιδιαίτερα όταν πρόκειται για όγκους θετικούς στους υποδοχείς οιστρογόνων (Estrogen receptor - ER +), κατά 7%. Αυτή η «προστασία» από τον κίνδυνο ανάπτυξης καρκίνου του μαστού κατά την εγκυμοσύνη οφείλεται πιθανώς στη διαδικασία διαφοροποίησης του ιστού του μαστού μετά την εγκυμοσύνη [39]. Ωστόσο παρατηρείται ότι, αντίθετα, με όσα αναφέρθηκαν παραπάνω, ο αυξημένος αριθμός τοκετών μπορεί να αυξήσει σημαντικά τον κίνδυνο TNBC (triple-negative breast cancer) μετά την εμμηνόπαυση [38].

Σύμφωνα με τους Toriola et al.[35], η καθυστερημένη τεκνοποίηση και η γέννηση θα μπορούσαν να αυξήσουν τον κίνδυνο έτσι ώστε, κάθε 1 έτος αύξησης της ηλικίας πρώτου τοκετού όπως επίσης και κάθε 1 έτος «καθυστέρησης» της πρώτης εμμηνόρροιας μπορεί να αυξήσει τον κίνδυνο καρκίνου του μαστού κατά 1,7 και 5%, αντίστοιχα [38]. Οι αναπαραγωγικοί παράγοντες περιλαμβάνουν την ηλικία έναρξης της εμμηνόρροιας, τον αριθμό τοκετών, την ηλικία του πρώτου τοκετού, το θηλασμό, την από του στόματος χρήση αντισυλληπτικών χαπιών, την ηλικία έναρξης της εμμηνόπαυσης και την ενδεχόμενη θεραπεία εμμηνόπαυσιακής ορμόνης. Αντίθετα, σύμφωνα με τον H.D. Nelson και τους συνεργάτες του, στη μετα-ανάλυση 13 ερευνών που πραγματοποίησαν, συμπέραναν ότι η εμμηνόρροια σε ηλικία μεγαλύτερη των 15 ετών συσχετίστηκε με μειωμένο κίνδυνο για καρκίνο του μαστού σε σύγκριση με την ηλικία αναφοράς έναρξης της εμμηνόρροιας που είναι τα 13 έτη [38]. Επιπλέον, οι Nelson et al. στη μετα-ανάλυση 14 μελετών που πραγματοποίησαν, διαπίστωσαν ότι ο θηλασμός σχετίστηκε με μειωμένο κίνδυνο εμφάνισης καρκίνου του μαστού, ιδιαίτερα όταν ο θηλασμός συνεχίστηκε για περισσότερο από 12 μήνες [38].

Ηλικία τοκετού

Μετά από τη μετα-ανάλυση 17 άλλων μελετών, η ερευνητική ομάδα των Nelson et al. [38] αναφέρει πως οι γυναίκες που δεν είχαν τεκνοποιήσει είχαν σημαντικά αυξημένο κίνδυνο για καρκίνο του μαστού σε σύγκριση με τις γυναίκες που τεκνοποίησαν έστω μία φορά στη ζωή τους. Οι ίδιοι, σε άλλη μετα-ανάλυση, 13 μελετών, κατέληξαν πως ο κίνδυνος ανάπτυξης του καρκίνου του μαστού μειώθηκε σημαντικά για γυναίκες με περισσότερες από τρεις γεννήσεις σε σύγκριση με τις άτεκνες γυναίκες. Σε μια ακόμη μετα-ανάλυση 5 μελετών, από την ίδια ερευνητική ομάδα, διαφαίνεται ότι οι γυναίκες που πραγματοποίησαν

τον πρώτο τους τοκετό σε ηλικία μεγαλύτερη των 30 ετών παρουσίασαν υψηλότερο κίνδυνο ανάπτυξης καρκίνου του μαστού σε σύγκριση με την ομάδα αναφοράς που χρησιμοποιήθηκε (γυναίκες ηλικίας 25 - 29 ετών) αλλά ελαφρώς χαμηλότερο κίνδυνο σε σχέση με τις άτεκνες γυναίκες [38].

Ο Wei-Chu Chie και οι συνεργάτες του [40] ανέλυσαν δεδομένα από δύο πολυκρατικές πληθυσμιακές μελέτες ελέγχου περιπτώσεων προκειμένου να διερευνήσουν τη σχέση μεταξύ της ηλικίας, οποιασδήποτε σε σειρά, ολοκληρωμένης εγκυμοσύνης και του κινδύνου εμφάνισης καρκίνου του μαστού. Η μελέτη περιλάμβανε περιπτώσεις καρκίνου του μαστού σε γυναίκες ηλικίας 20 - 79 ετών που ανήκαν σε τέσσερα πολιτειακά μητρώα καρκίνου όπως επίσης και τυχαία επιλεγμένους ελέγχους που διεξήχθησαν μέσω συνεντεύξεων από το 1988 έως το 1996. Πλήρεις πληροφορίες σχετικά με ένα ολοκληρωμένο σύνολο παραγόντων κινδύνου για τον καρκίνο του μαστού ήταν διαθέσιμες για 9.891 περιπτώσεις και 12.271 ελέγχους. Ο μεγάλος αριθμός περιστατικών επέτρεψε την αποτελεσματική εφαρμογή διαφόρων προσεγγίσεων μοντελοποίησης. Όπως διαπιστώθηκε, κάθε 5ετής αύξηση στην ηλικία της πρώτης ολοκληρωμένης εγκυμοσύνης συσχετίστηκε με OR (ODD RATIO - λόγος απόδοσης) 1,07 (95% CI: 1,01, 1,13) για καρκίνο του μαστού. Οι αντίστοιχες εκτιμήσεις OR ήταν 1,02 (95% CI: 1,00, 1,05) για την ηλικία από τη δεύτερη έως και την ένατη ολοκληρωμένη εγκυμοσύνη. Όσον αφορά στην ηλικία κατά την τελευταία ολοκληρωμένη εγκυμοσύνη, η εκτίμηση της επίδρασης αυτού του παράγοντα (OR = 1,01, 95% CI: 0,97, 1,06) δεν διέφερε από εκείνη για τις άλλες ολοκληρωμένες εγκυμοσύνες μετά την πρώτη. Σε αυτή την ανάλυση, παρατηρήθηκε επίσης μια μέτρια και παροδική αύξηση του κινδύνου εμφάνισης καρκίνου του μαστού μετά τον τοκετό. Η σχετικά μεγαλύτερη επίδραση της ηλικίας κατά τον πρώτο τοκετό συνάδει με την ύπαρξη μακροπρόθεσμης επίδρασης του πρώτου τοκετού σε νεαρή ηλικία στη διαφοροποίηση των μαστικών κυττάρων, με αποτέλεσμα αυτά να γίνονται λιγότερο ευάλωτα στην καρκινογένεση [40].

Μια ακόμα σημαντική μελέτη ελέγχου περιπτώσεων πραγματοποιήθηκε στη Σουηδία [41], σε γυναίκες που γεννήθηκαν μεταξύ 1925 και 1960. Στόχος της εν λόγω έρευνας ήταν να διερευνηθούν οι συσχετίσεις μεταξύ της τεκνοποίησης και του κινδύνου εμφάνισης καρκίνου του μαστού, καθώς και η σχετική σημασία της ηλικίας των γυναικών κατά τον πρώτο και κατά τον τελευταίο τοκετό. Στη μελέτη συμπεριλήφθηκαν συνολικά 12.782 γυναίκες με καρκίνο του μαστού και πενταπλάσιος αριθμός γυναικών που αποτέλεσαν τον πληθυσμό ελέγχου και αντιστοιχούσαν στα ηλικιακά επίπεδα των γυναικών με καρκίνο μαστού. Όλες οι συμμετέχουσες στην έρευνα ήταν κάτω των 60 ετών και υπήρχαν διαθέσιμες οι πληροφορίες για τη γονιμότητά τους [41].

Όπως διαπιστώθηκε από τους M. Lambe et al. [41], η αύξηση της τεκνοποίησης συσχετίστηκε με έντονη μείωση του κινδύνου εμφάνισης καρκίνου του μαστού, καθώς κάθε πρόσθετος τοκετός επέφερε μείωση του κινδύνου κατά 10% (OR = 0,90 [95% CI 0,88-0,91]). Σε μια ανάλυση που περιορίστηκε σε γυναίκες

με δύο ή περισσότερους τοκετούς και αφού λήφθηκε υπόψιν η επίδραση της ηλικίας των γυναικών κατά τις ενδιάμεσες γεννήσεις, διαπιστώθηκε ότι ο κίνδυνος καρκίνου του μαστού αυξήθηκε κατά περίπου 13% για κάθε αύξηση της ηλικίας πρώτου τοκετού κατά πέντε έτη (OR = 1,13 [1,08-1,19]). Για κάθε πενταετή αύξηση της ηλικίας κατά τον τελευταίο τοκετό παρατηρήθηκε μικρή αύξηση του κινδύνου εμφάνισης καρκίνου του μαστού με οριακή στατιστική σημασία (OR = 1,05 [1,01-1,09]) [41]. Τα ευρήματα των M. Lambe et al. [41] έρχονται σε αντίθεση με τους ισχυρισμούς ότι η ηλικία κατά τον τελευταίο τοκετό έχει ισχυρότερη επίδραση από την ηλικία κατά τον πρώτο τοκετό στον κίνδυνο καρκίνου του μαστού. Η σημασία της ηλικίας κατά τον πρώτο τοκετό ως διαμορφωτή του κινδύνου είναι πιθανό να αντανάκλα την προστασία που παρέχει η τελική διαφοροποίηση των κυττάρων του μαστού που προκαλείται κατά τη διάρκεια της πρώτης εγκυμοσύνης [41].

Μια ακόμα έρευνα με αντιφατικά αποτελέσματα δημοσιεύτηκε το 2019 [42], όπου αναφέρεται ότι η τεκνοποίηση έχει διπλή επίδραση στον κίνδυνο καρκίνου του μαστού. Με πολλούς τρόπους, η καρκινογένεση του μαστού μιμείται διάφορους μηχανισμούς που συνήθως ενεργοποιούνται κατά τη διάρκεια της εγκυμοσύνης, συμπεριλαμβανομένου του αυξημένου κυτταρικού πολλαπλασιασμού, των μεταβολών στην αποβολή των κυττάρων, της μειωμένης κυτταρικής απόπτωσης, της αλλαγής της γονιδιακής έκφρασης και των τροποποιήσεων της εξωκυττάριας μήτρας. Οι ερευνητές της εν λόγω μελέτης υπογραμμίζουν την ύπαρξη επιδημιολογικών μελετών όπου παρουσιάζονται τα αντικαρκινικά οφέλη της εγκυμοσύνης, σύμφωνα με τα οποία η μικρότερη ηλικία εγκυμοσύνης μειώνει τον κίνδυνο ανάπτυξης καρκίνου του μαστού, ωστόσο οι ίδιοι διαπίστωσαν διπλή επίδραση [42]. Αν και η κατανόηση των μοριακών μηχανισμών που διέπουν αυτά τα φαινόμενα βρίσκεται ακόμη σε νηπιακό στάδιο, η διαλεύκανσή τους θα ανοίξει νέους δρόμους για τη κατανόηση του καρκίνου του μαστού.

Αμβλώσεις και Αποβολές

Τα αποτελέσματα από προηγούμενες μελέτες σχετικά με τις αποβολές ή τις προκαλούμενες αμβλώσεις και τον κίνδυνο εμφάνισης μεταγενέστερου καρκίνου του μαστού είναι αντιφατικά, πιθανώς λόγω των διαφορετικών μεθόδων μελέτης

Η ερευνητική ομάδα του Erlandsson και των συνεργατών του [43], διερεύνησε τη σχέση μεταξύ των αποβολών και του κινδύνου εμφάνισης καρκίνου του μαστού μεταγενέστερα, σε μια μελέτη ελέγχου περιπτώσεων χρησιμοποιώντας πληροφορίες που καταγράφηκαν σε δεύτερο χρόνο. Ο πληθυσμός της μελέτης περιελάμβανε γυναίκες που καταγράφηκαν στο πληθυσμιακό ιατρικό μητρώο γεννήσεων της Σουηδίας μεταξύ του 1973 και του 1991. Οι περιπτώσεις καθορίστηκαν με σύνδεση του μητρώου γεννήσεων με το Σουηδικό Μητρώο Καρκίνου και οι έλεγχοι επιλέχθηκαν τυχαία από το μητρώο γεννήσεων. Από τα αρχεία προγεννητικής φροντίδας των συμμετεχουσών αντλήθηκαν πληροφορίες

σχετικά με τις αποβολές ή τις προκαλούμενες αμβλώσεις καθώς και μια σειρά πιθανών κρίσιμων παραγόντων. Ο σχετικός κίνδυνος εμφάνισης καρκίνου του μαστού εκτιμήθηκε με λόγους απόδοσης (Odds ratio - OR) και με 95% διαστήματα εμπιστοσύνης (95% CI). Παρατηρήθηκε μειωμένος κίνδυνος καρκίνου του μαστού για τις γυναίκες με ιστορικό τουλάχιστον μίας αμβλώσεως σε σύγκριση με αυτές που δεν είχαν υποβληθεί σε καμία αμβλώση (προσαρμοσμένο OR = 0,84, 95% CI = 0,72 - 0,99). Ο προσαρμοσμένος λόγος πιθανοτήτων OR μειώνεται σταδιακά όσο αυξάνεται ο αριθμός των αμβλώσεων, συγκεκριμένα φτάνει το 0,59 (95% CI = 0,34 - 1,03) για 3 ή περισσότερες αμβλώσεις σε σύγκριση με καμία αμβλώση. Τα αποτελέσματα είναι παρόμοια για τις αποβολές. Συμπερασματικά, ούτε το ιστορικό των αποβολών ούτε αυτό των προκαλούμενων αμβλώσεων σχετίζεται με αυξημένο κίνδυνο καρκίνου του μαστού. Τα δεδομένα της παραπάνω έρευνας υποδηλώνουν ξεκάθαρα την προστατευτική επίδραση των κυήσεων ανεξάρτητα από την έκβασή τους.

Μια ακόμα ερευνητική ομάδα, των Reeves et al. [44], εξέτασε το ρόλο των αποβολών και των προκαλούμενων αμβλώσεων στον κίνδυνο εμφάνισης καρκίνου του μαστού. Στη μελέτη συμμετείχαν 267.361 γυναίκες και τα στοιχεία τους ελήφθησαν από την Ευρωπαϊκή Προοπτική Έρευνα για τον Καρκίνο και τη Διατροφή μεταξύ του 1992 και του 2000. Τα δεδομένα συλλέχθηκαν από 20 κέντρα, σε 9 χώρες, και περιλάμβαναν πληροφορίες για συνολικά 4.805 γυναίκες με καρκίνο του μαστού, εκ των οποίων οι 1.657 ανέφεραν ότι δεν τους είχε συμβεί ποτέ κάποια αποβολή. Συνολικά, ο σχετικός κίνδυνος εμφάνισης καρκίνου του μαστού σε γυναίκες που ανέφεραν ότι είχαν μία αποβολή στο παρελθόν δεν ήταν σημαντικά αυξημένος σε σύγκριση με τις γυναίκες που ανέφεραν ότι δεν είχαν αποβάλει ποτέ (RR = 1,07, 95% CI = 0,99-1,14). Ωστόσο, υπήρχαν κάποιες ενδείξεις για ελαφρά αύξηση του κινδύνου εμφάνισης καρκίνου του μαστού στις γυναίκες που ανέφεραν ότι είχαν βιώσει 2 ή περισσότερες αποβολές (1,20, 1,07-1,35). Ο σχετικός κίνδυνος καρκίνου του μαστού μεταξύ των γυναικών που ανέφεραν ότι είχαν υποβληθεί σε προκαλούμενη αμβλώση σε σύγκριση με τις γυναίκες που ανέφεραν ότι δεν είχαν υποβληθεί σε αμβλώση ήταν 0,95 (0,87-1,03). Συνολικά, τα ευρήματα συνηγορούν επίσης για την έλλειψη δυσμενούς επίδρασης της προκαλούμενης αμβλώσεως στον κίνδυνο καρκίνου του μαστού [44].

Τα αποτελέσματα της παραπάνω έρευνας συμπληρώνουν οι παρατηρήσεις της μετα-ανάλυσης που πραγματοποίησε η ερευνητική ομάδα των Yongchun Deng, et al. [45]. Τα αποτελέσματα της μετα-ανάλυσης που βασίστηκε σε 25 μελέτες, συμπεριλαμβανομένων 5 μελετών με Κινέζες ασθενείς, έδειξαν ότι δεν υπήρχε συσχέτιση της προκαλούμενης αμβλώσεως με τον καρκίνο του μαστού (OR = 1,08, 95% CI 0,98-1,19, P = 0,1).

Λήψη αντισυλληπτικών χαπιών

Οι Nelson et al. [38], στη μετα-ανάλυση που πραγματοποίησαν, αξιολογήσαν δώδεκα μελέτες για την, από του στόματος, λήψη αντισυλληπτικής αγωγής, και παρείχαν εκτιμήσεις για τη συσχέτιση καρκίνου του μαστού με τη διαρκή χρήση αντισυλληπτικών χαπιών σε σύγκριση με την καθόλου χρήση. Άλλες 8 μελέτες παρείχαν εκτιμήσεις για τον ρόλο της συχνότητας της χρήσης στην εμφάνιση καρκίνου του μαστού. Ωστόσο, σύμφωνα με τους Nelson et al. [38], καμία από αυτές τις συσχετίσεις δεν ήταν στατιστικά σημαντική. Από την άλλη, τα δεδομένα του BCSC (Breast Cancer Surveillance Consortium - Κοινοπραξία Επιτήρησης για τον Καρκίνο του Μαστού) έδειξαν σημαντικά αυξημένο κίνδυνο εμφάνισης καρκίνου του μαστού σε γυναίκες με τρέχουσα, από του στόματος, αντισυλληπτική αγωγή σε σύγκριση με γυναίκες που έκαναν προηγούμενη ή καθόλου χρήση [38].

Ηλικία Εμμηνόπαυσης και Ορμονική θεραπεία

Τα στοιχεία του BCSC έδειξαν μειωμένο κίνδυνο καρκίνου του μαστού για γυναίκες που βρίσκονται σε εμμηνοπαυσιακή φάση (είτε χειρουργική είτε μη χειρουργική εμμηνόπαυση) σε σύγκριση με τις γυναίκες που βρίσκονται σε προεμμηνοπαυσιακή φάση. Σύμφωνα με τα εν λόγω στοιχεία, οι γυναίκες που είχαν υποστεί υστερεκτομή και λάμβαναν θεραπεία με εμμηνοπαυσιακή ορμόνη είχαν μειωμένο κίνδυνο για ανάπτυξη καρκίνου του μαστού σε σύγκριση με εκείνες που δε λάμβαναν την εν λόγω θεραπεία. Οι γυναίκες που βρίσκονταν σε εμμηνόπαυση, χωρίς να έχουν υποστεί υστερεκτομή, δεν φαίνεται να παρουσίασαν διαφορές με λήψη ή όχι τέτοιας θεραπείας, πιθανώς, γιατί οι γυναίκες που έχουν υποστεί υστερεκτομή λαμβάνουν μόνο θεραπεία οιστρογόνων ενώ εκείνες που δεν έχουν υποστεί λαμβάνουν οιστρογόνα σε συνδυασμό με προγεστερόνη [38].

Οι Colditz et al. [46], σε σχετική ερευνά τους διαπίστωσαν ότι η αυξημένη χρήση των μετεμμηνοπαυσιακών ορμονών από τις γυναίκες συνέβαλε σημαντικά στη σταθερή αύξηση της συχνότητας εμφάνισης καρκίνου του μαστού θετικού σε υποδοχείς οιστρογόνων/ υποδοχείς προγεστερόνης (Estrogen receptor - ER / Progesterone receptor - PR) μέχρι και τη δεκαετία του 1980, καθώς αυτές οι ορμόνες φαίνεται να διευκόλυναν την ανάπτυξη αυτών των, κατά τα άλλα αδρανών, ορμονικά ανταποκρινόμενων όγκων. Μεταγενέστερα παρατηρήθηκε απότομη μείωση της συχνότητας εμφάνισης καρκίνου του μαστού που θα μπορούσε να οφείλεται στην αυξημένη ευαισθητοποίηση του κοινού και των γιατρών σχετικά με τις επιδράσεις των μετεμμηνοπαυσιακών ορμονών και στην απόφαση για μειωμένη χρήση τους [46].

Σύμφωνα με την έρευνα των Beral et al. [47], από τα 15.759 περιστατικά γυναικών με καρκίνο του μαστού, τα 7.107 αφορούσαν γυναίκες που έκαναν τρέχουσα χρήση ορμονικής θεραπείας. Η συχνότητα

εμφάνισης καρκίνου του μαστού αυξήθηκε τόσο στις γυναίκες που έκαναν τρέχουσα χρήση ορμονικής θεραπείας, αλλά και σε αυτές που διέκοψαν τη χρήση, καθώς εμφανίστηκε καρκίνος του μαστού μερικά χρόνια αργότερα. Σύμφωνα με την εν λόγω ερευνητική ομάδα, οι σχετικοί κίνδυνοι για ανάπτυξη καρκίνου του μαστού σε γυναίκες με τρέχουσα ορμονική θεραπεία ήταν μεγαλύτεροι εάν η ορμονική θεραπεία ξεκίνησε πριν ή λίγο μετά την εμμηνόπαυση σε σχέση με τις γυναίκες που το διάστημα λήψης της ορμονικής θεραπείας απείχε σημαντικά χρονικά από την εμμηνόπαυση.

Συγκεκριμένα όσον αφορά στις γυναίκες με τρέχουσα λήψη ορμονικής θεραπείας, παρατηρήθηκε ότι, για αυτές που λάμβαναν θεραπεία αποκλειστικά με οιστρογόνα, στις περιπτώσεις που η χρήση ξεκίνησε 5 ή περισσότερα χρόνια μετά την εμμηνόπαυση, σημειώθηκε μικρή ή καθόλου αύξηση του κινδύνου αλλά ο κίνδυνος αυξήθηκε στατιστικά εάν η χρήση ξεκίνησε σε λιγότερο από 5 χρόνια μετά την εμμηνόπαυση. Ένα παρόμοιο μοτίβο παρατηρήθηκε μεταξύ των γυναικών που λάμβαναν τρέχουσα θεραπεία σκευασμάτων οιστρογόνου - προγεστερόνης. Σε ηλικίες 50 - 59 ετών, τα ποσοστά συχνότητας εμφάνισης καρκίνου του μαστού ήταν 0,30% για γυναίκες που δεν είχαν κάνει ποτέ λήψη ορμονικής θεραπείας, 0,43% για γυναίκες που λάμβαναν θεραπεία οιστρογόνων σε λιγότερο από 5 χρόνια από την έναρξη της εμμηνόπαυσης και 0,61% αντίστοιχα, για τις γυναίκες που λάμβαναν θεραπεία οιστρογόνων - προγεστερόνης [47].

Φυλετικός Παράγοντας

Η επίπτωση του καρκίνου του μαστού όλων των σταδίων συνεχίζει να αυξάνεται στις λευκές γυναίκες ηλικίας 50 ετών και άνω, ενώ η τάση της νόσου είναι σταθερή για την ίδια ηλικιακή ομάδα των αφροαμερικανίδων γυναικών. Από την άλλη πλευρά, η τάση της νόσου είναι σταθερή στις λευκές γυναίκες ηλικίας κάτω των 50 ετών, ενώ μειώνεται για τις αφροαμερικανίδες της ίδιας ηλικίας. Μια ενδιαφέρουσα στατιστική ανωμαλία είναι ότι επειδή οι γυναίκες ηλικίας 50 ετών και άνω αντιπροσωπεύουν περίπου 73 % του γυναικείου πληθυσμού, τα συνολικά ποσοστά επίπτωσης φαίνεται να είναι σημαντικά υψηλότερα σε αυτή την ηλικιακή ομάδα (375,0 ανά 100.000 γυναίκες) σε σύγκριση με τις γυναίκες ηλικίας κάτω των 50 ετών (42,5 ανά 100.000 γυναίκες), παρόλο που περίπου 23 % των καρκίνων του μαστού διαγιγνώσκονται σε γυναίκες ηλικίας κάτω των 50 ετών. Οι φυλετικές διαφορές είναι σαφώς εμφανείς σε πολλές μελέτες, οι οποίες έδειξαν ότι παρά τα χαμηλότερα ποσοστά επίπτωσης του καρκίνου του μαστού στις Αφροαμερικανίδες, τα ποσοστά θνησιμότητας είναι 37 % υψηλότερα από ό,τι στις λευκές γυναίκες [48].

Παρομοίως, έχει επισημανθεί ότι στις ΗΠΑ, τα ποσοστά θανάτου από καρκίνο του μαστού μειώνονται από τις αρχές της δεκαετίας του 1990 για όλες τις γυναίκες εκτός από τις Ινδιάνες/Ιθαγενείς της Αλάσκα, μεταξύ των οποίων τα ποσοστά παρέμειναν σταθερά. Έχει εκτιμηθεί ότι το μέσο ετήσιο ποσοστό θανάτου από

καρκίνο του μαστού στις γυναίκες μεταξύ 2003 και 2007 ήταν υψηλότερο στις Αφροαμερικανίδες (32,4 θάνατοι ανά 100.000 γυναίκες) και χαμηλότερο στις Ασιατοαμερικανίδες (12,2 θάνατοι ανά 100.000 γυναίκες). Επίσης, οι Αφροαμερικανίδες έχουν το χαμηλότερο ποσοστό 5ετούς επιβίωσης από τον καρκίνο του μαστού (77,5 %) από κάθε άλλη φυλετική ή εθνοτική ομάδα (Σχήμα 2), ενώ οι γυναίκες από την Ασία και τις Νήσους του Ειρηνικού έχουν το υψηλότερο ποσοστό 5ετούς επιβίωσης από τον καρκίνο του μαστού (90,3 %) [49].

Είναι διαπιστωμένο ότι ενώ η συχνότητα εμφάνισης καρκίνου του μαστού είναι υψηλότερη στις ανεπτυγμένες χώρες και επίσης σε γυναίκες με υψηλότερη κοινωνικοοικονομική κατάσταση, η θνησιμότητα λόγω καρκίνου του μαστού είναι υψηλότερη σε γυναίκες από φτωχότερες χώρες και επίσης χαμηλότερης κοινωνικοοικονομικής κατάστασης. Αυτό μπορεί να οφείλεται εν μέρει στα χαμηλότερα ποσοστά προσυμπτωματικού ελέγχου στις φτωχές γυναίκες σε σύγκριση με τις σχετικά εύπορες γυναίκες, παρά την μεγάλη πρόοδο στην αύξηση της χρήσης μαστογραφίας [50].

Δείκτης Μάζας Σώματος και Παχυσαρκία

Τα ευρήματα της έρευνας των S.R.Shawon et al.[45], που αφορούν στη σχέση μεταξύ του μεγέθους του σώματος στην πρόωμη ζωή και του κινδύνου του καρκίνου του μαστού συμφωνούν με τις περισσότερες από τις επιδημιολογικές μελέτες, τόσο σε κατεύθυνση όσο και σε μέγεθος. Τα αποτελέσματα από παγκόσμιες, μεγάλες μελέτες που διεξάχθηκαν με πραγματικά κλινικά δεδομένα [51], [52] υποστηρίζουν ότι διάφορα μέτρα αξιολόγησης του σωματικού λίπους στην παιδική ηλικία όπως είναι ο Δείκτης Μάζας Σώματος (Body Mass Index - BMI) σχετίζονται αντιστρόφως ανάλογα με τον κίνδυνο καρκίνου του μαστού στις ενήλικες.

Κάποιες αναδρομικές μελέτες ελέγχου περιπτώσεων [53], [54] ανέφεραν επίσης τον «προστατευτικό» ρόλο του σωματικού λίπους σε μικρές ηλικίες. Σύμφωνα με μελέτη που συμπεριελάμβανε 117.415 γυναίκες στη Δανία, ο χαμηλός BMI φαίνεται να αποτελεί ανεξάρτητο παράγοντα κινδύνου για τον καρκίνο του μαστού [55]. Στη Σουηδία, σε μια ανεξάρτητη μελέτη ελέγχου περιπτώσεων που πραγματοποιήθηκε, παρατηρήθηκε ότι το μεγάλο μέγεθος σώματος στην ηλικία των 7 ετών συσχετίστηκε με 27% μείωση του κινδύνου μετεμμηνοπαυσιακού καρκίνου του μαστού σε σύγκριση με το μικρό μέγεθος σώματος [56]. Μετά την ενσωμάτωση του BMI ενηλίκων σε αυτή τη μελέτη, οι συσχετίσεις παρέμειναν σημαντικές και δεν διέφεραν αισθητά, υποδηλώνοντας ότι το μεγαλύτερο μέγεθος σώματος στην παιδική ηλικία μπορεί να παρέχει ανεξάρτητη μακροχρόνια «προστασία» κατά του καρκίνου του μαστού.

Το μεγάλο μέγεθος σώματος την περίοδο πριν από την εφηβεία μπορεί να επηρεάσει τη σεξουαλική ωρίμανση και την εφηβική ανάπτυξη της γυναίκας, γεγονός που εξηγεί την αντίστροφη σχέση με τον

κίνδυνο ανάπτυξης καρκίνου του μαστού στην ενηλικίωση. Μια μεγάλη έρευνα που διεξήχθη στη Γαλλία αναφέρει ότι είναι μοναδική η επίδραση του σωματικού λίπους κατά την έναρξη της εμμηνόρροιας ως προς την καρκινογένεση στο μαστό, και δε σχετίζεται με τη παχυσαρκία σε άλλες ηλικίες [57].

Ομοίως, τα δεδομένα των S.R.Shawon et al.[45] μας δείχνουν ότι το μεγαλύτερο μέγεθος σώματος στην ηλικία των 7 ετών συσχετίζεται σημαντικά με την μικρότερη ηλικία έναρξης της εμμηνόρροιας. Δεδομένου ότι η μικρότερη ηλικία έναρξης της εμμηνόρροιας (πρώιμη εφηβεία) θεωρείται ότι αυξάνει τη συσσωρευτική έκθεση σε οιστρογόνα, η παχυσαρκία στην παιδική ηλικία θα πρέπει να συνδέεται διαισθητικά με τον αυξημένο κίνδυνο καρκίνου του μαστού. Ωστόσο, η πρώιμη εφηβεία φαίνεται επίσης να οδηγεί σε μικρότερο τελικό ύψος για την γυναίκα, το οποίο είναι επίσης παράγοντας κινδύνου για καρκίνο του μαστού, σε κάποιες μελέτες [58].

Πυκνότητα μαστών

Στην απεικόνιση της μαστογραφίας, σε αντίθεση με τον ακτινοδιαπερατό λιπώδη ιστό, το επιθήλιο και το στρώμα του μαστού εμφανίζονται ακτινοσκοιρά και αποτυπώνονται ως μαστογραφική πυκνότητα. Περίπου το 40% των γυναικών ηλικίας 40 έως 74 ετών έχουν ετερογενείς ή εξαιρετικά πυκνούς μαστούς και αυτό έχει βρεθεί ότι σχετίζεται αντιστρόφως ανάλογα με την ηλικία, τον BMI, την εγκυμοσύνη και τον θηλασμό. Επίσης έχει παρατηρηθεί ότι κατά την περίοδο λήψης θεραπείας ορμονικής υποκατάστασης, ιδίως της οξικής νορεθινδρόνης, οι γυναίκες εμφανίζουν αυξημένη μαστογραφική πυκνότητα, γεγονός που υποδηλώνει ότι η μαστογραφική πυκνότητα επηρεάζεται πιθανότατα τόσο από ενδογενείς όσο και από εξωγενείς ορμόνες [59]. Η μαστογραφική πυκνότητα σχετίζεται με αυξημένο κίνδυνο εμφάνισης καρκίνου του μαστού, τόσο κατά το διαγνωστικό έλεγχο όσο και σε ορισμένα χρονικά διαστήματα μετά από μια αρνητική μαστογραφία προσυμπτωματικού ελέγχου. Σύμφωνα με μελέτη που διεξήχθη το 2007 από τον Boyd και τους συνεργάτες του [60], στην οποία εξετάστηκαν 1.112 αντιστοιχισμένα ζεύγη περιπτώσεων-ελέγχου, διαπιστώθηκε ότι οι γυναίκες που εμφάνιζαν πυκνούς μαστούς στη μαστογραφία είχαν σημαντικά αυξημένο κίνδυνο καρκίνου του μαστού (OR= 4,7), ανεξάρτητα αν ο καρκίνος ανιχνεύθηκε με μαστογραφικό διαγνωστικό έλεγχο ή με άλλα μέσα. Έχει διαπιστωθεί ότι οι καρκίνοι του μαστού που προέρχονται από περιοχές πυκνού ιστού συνδέονται συχνότερα με υψηλού βαθμού καρκίνο ιστολογικά, λεμφαγγειακή διείσδυση και προχωρημένο στάδιο σε σύγκριση με εκείνους που προκύπτουν σε πιο ακτινοδιαπερατούς ιστούς [60]. Επιπλέον, ο μαστογραφικά πυκνός ιστός του μαστού, όπως είναι αναμενόμενο, καθιστά την ανίχνευση του καρκίνου πιο δύσκολη. Η αυξημένη μαστογραφική πυκνότητα μειώνει την ευαισθησία της παραδοσιακής μαστογραφίας. Αυτό το ζήτημα έχει τουλάχιστον εν μέρει βελτιωθεί με την ψηφιακή μαστογραφία, η οποία μπορεί να τροποποιήσει την αντίθεση της εικόνας για να βελτιώσει την ανίχνευση των μαζών που ενδεχομένως υπάρχουν στον πυκνό ιστό του μαστού.

Συμπερασματικά, η πυκνότητα των μαστών σχετίζεται με αυξημένο κίνδυνο ανάπτυξης καρκίνου του μαστού, ακόμα και στις περιπτώσεις που η μαστογραφία διαλογής έτεινε υπέρ αρνητικού αποτελέσματος [61]. Η πρόσφατη υποχρεωτική αναφορά της πυκνότητας στη μαστογραφία σε αρκετά κράτη παγκοσμίως δεν έχει ακόμη διαπιστωθεί ότι βελτιώνει το κλινικό αποτέλεσμα. Παρά το επιχείρημα ότι η αναφορά της πυκνότητας θα βοηθήσει στον εντοπισμό των γυναικών που χρήζουν απεικόνισης εκτός μαστογραφίας, εξακολουθεί να τονίζεται η χρησιμότητα της παραδοσιακής μαστογραφίας κατά τον προσυμπτωματικό έλεγχο [61].

Μέγεθος Στήθους

Ο Jansen και οι συνεργάτες του [62], πραγματοποίησαν συστηματική έρευνα για τη σχέση του μεγέθους του στήθους με τον καρκίνο του μαστού, το 1950 έως τον Νοέμβριο του 2010. Η έρευνα ενημερώθηκε εκ νέου τον Φεβρουάριο του 2014. Η εν λόγω ερευνητική ομάδα αναζήτησε στη βιβλιογραφία την αξιολόγηση της σχέσης μεταξύ των ακόλουθων μεταβλητών και του καρκίνου του μαστού: 1) μέγεθος στήθους 2) μείωση στήθους 3) αύξηση στήθους και 4) προφυλακτική υποδόρια μαστεκτομή. Στη συστηματική ανασκόπηση που πραγματοποίησαν συμπεριέλαβαν 50 έρευνες. Σύμφωνα με τα συγκεντρωτικά αποτελέσματα που παρουσίασαν, η αύξηση του μεγέθους του στήθους φαίνεται να αποτελεί παράγοντα κινδύνου για τον καρκίνο του μαστού, αλλά οι μελέτες περιορίζονται από την αναδρομική τους φύση, τις ατελείς τεχνικές μέτρησης του μεγέθους και τις συγκεχυμένες μεταβλητές που συμπεριέλαβαν. Τα στοιχεία είναι ισχυρότερα όσον αφορά στη μείωση του κινδύνου ανάπτυξης καρκίνου του μαστού με μείωση του μεγέθους του μαστού, συμπεριλαμβανομένης της ριζικής προφυλακτικής υποδόριας μαστεκτομής. Υπάρχουν άμεσα και έμμεσα στοιχεία που καταδεικνύουν ότι το μέγεθος του στήθους αποτελεί σημαντικό παράγοντα στον κίνδυνο ανάπτυξης καρκίνου του μαστού [62].

Μια ακόμα ερευνητική ομάδα που διεξήγαγε έρευνα μεταξύ γυναικών με καρκίνο του μαστού στη Σουηδία, έλαβε μετρήσεις μεγέθους στήθους αλλά και άλλων σωματικών μεγεθών προεγχειρητικά από 772 ασθενείς [63]. Οι μετρήσεις διεξήχθησαν από ειδικά εκπαιδευμένο προσωπικό με τη βοήθεια πλαστικών θηκών που χρησιμοποιούν οι πλαστικοί χειρουργοί για να κάνουν μειώσεις στήθους. Τα κλινικά δεδομένα ελήφθησαν από τα διαγράμματα ασθενών και τις παθολογικές αναφορές. Σύμφωνα με τα αποτελέσματα της εν λόγω μελέτης, για τις ασθενείς με όγκο δεξιού και αριστερού μαστού μεγαλύτερο από 850 mL ήταν πιο πιθανό να έχουν μεγαλύτερα μεγέθη καρκινικού όγκου, σε πιο προχωρημένο ιστολογικό βαθμό και με περισσότερη μασχαλαία συμμετοχή. Σημειώθηκαν 62 συμβάντα καρκίνου του μαστού κατά τη διάρκεια της επταετούς παρακολούθησης. Ο όγκος του μαστού μεγαλύτερος των 850 mL συσχετίστηκε με μικρότερη επιβίωση χωρίς τη νόσο (disease-free survival) και μεγαλύτερη επιβίωση χωρίς μετάσταση (metastasis-free survival) σε ασθενείς με όγκους θετικούς στους υποδοχείς οιστρογόνων (ER) ανεξάρτητα

από άλλες ανθρωπομετρικές μετρήσεις και ηλικία. Σε ασθενείς με όγκους ER-θετικούς, το μέγεθος του μαστού ήταν ένας ανεξάρτητος προγνωστικός παράγοντας με μικρότερη επιβίωση χωρίς τη νόσο (disease-free survival) και μεγαλύτερη επιβίωση χωρίς μετάσταση (metastasis-free survival), προσαρμοσμένος στα χαρακτηριστικά του όγκου, τον BMI, την ηλικία και το είδος θεραπείας. Το κύριο εύρημα αυτής της μελέτης ήταν ότι το στήθος με όγκο μεγαλύτερο των 850 mL σε γυναίκες με ER-θετικούς όγκους σχετίζεται με σημαντικά μικρότερο χρόνο επιβίωσης εκτός ασθένειας σε σύγκριση με τις γυναίκες με μικρότερο στήθος, ανεξάρτητα από ηλικία, τον BMI, το μέγεθος της λεκάνης και την περιφέρεια της μέσης [63]. Τα αποτελέσματα της μελέτης ήταν συνεπή με τα ευρήματα προηγούμενων μελετών [64], [65] όπου γυναίκες με μεγαλύτερο στήθος παρουσίαζαν πιο επιθετικά χαρακτηριστικά όγκου από τις γυναίκες με μικρότερο στήθος.

Οικογενειακό ιστορικό

BMI & Οικογενειακό ιστορικό

Ο BMI είναι ένας ενδιαφέρων παράγοντας κινδύνου για την εμφάνιση καρκίνου του μαστού καθώς σύμφωνα με την έρευνα των J. L. Hopper et al. [68], η σχέση του με την ασθένεια εξαρτάται από το αν η γυναίκα βρίσκεται στην προεμμηνοπαυσιακή ή μετεμμηνοπαυσιακή φάση. Έχει διαπιστωθεί ότι ο υψηλότερος BMI σχετίζεται με αυξημένο κίνδυνο ανάπτυξης καρκίνου του μαστού μόνο για τις γυναίκες που βρίσκονται σε εμμηνόπαυση. Αντίθετα, για τις προεμμηνοπαυσιακές γυναίκες, ακόμη και για τις έφηβες, έχει διαπιστωθεί ότι ο υψηλότερος BMI σχετίζεται με μειωμένο κίνδυνο ανάπτυξης της νόσου. Αυτά τα ευρήματα ήταν συνεπή σε διαφορετικές φυλετικές/εθνοτικές υποομάδες και έρευνες βασισμένες σε μελέτη περιπτώσεων αλλά και διεθνούς βεληγκεούς αναλύσεις δεδομένων. Ο BMI αποτελεί έναν σημαντικό παράγοντα κινδύνου επειδή είναι δυνητικά τροποποιήσιμος. Το γεγονός ότι ο υψηλός BMI φαίνεται να λειτουργεί «προστατευτικά» σε μικρές ηλικίες, αλλά δρα αντίθετα στη μετέπειτα ζωή, ενδεχομένως αποτελεί πρόβλημα στη εξαγωγή συμπερασμάτων για τη συσχέτισή του με τον καρκίνο του μαστού. Επομένως, οι συνέπειες του παράγοντα αυτού χρειάζεται να ποσοτικοποιηθούν [68].

Μια προοπτική μελέτη [66] αλλά και μια μελέτη που βασίστηκε σε έλεγχο περιπτώσεων [67] έδειξαν ότι ο αυξημένος κίνδυνος εμφάνισης καρκίνου του μαστού που σχετίζεται με υψηλότερο BMI αυξάνεται χρόνο με το χρόνο μετά την εμμηνόπαυση, αλλά είναι δεν είναι εμφανής η επιρροή έως 10 χρόνια μετά την εμμηνόπαυση [68]. Μια καλύτερη κατανόηση του τρόπου με τον οποίο σχετίζεται ο BMI με τον καρκίνο του μαστού θα ήταν εξαιρετικά χρήσιμη καθώς η σχέση των δύο ποικίλλει ανάλογα με την ηλικία και την κατάσταση της εμμηνόπαυσης. Ιδιαίτερο ενδιαφέρον παρουσιάζει το γεγονός ότι μελέτες που σχετίζονται με το ανθρώπινο γονιδίωμα, και συγκεκριμένα με τον υπολογισμό του κινδύνου ανάπτυξης καρκίνου του μαστού λόγω γενετικών παραγόντων με βάση τους πολυμορφισμούς των νουκλεοτιδίων (Single-

Nucleotide Polymorphisms - SNPs), προβλέπουν ότι ο υψηλότερος BMI στην παιδική ηλικία ή την εφηβεία σχετίζεται με χαμηλότερο κίνδυνο τόσο του προεμμηνόπαυσιακού όσο και του μετεμμηνόπαυσιακού καρκίνου του μαστού [69].

Το οικογενειακό ιστορικό είναι ένας ακόμα σημαντικός παράγοντας κινδύνου για την εμφάνιση καρκίνου του μαστού στις περιπτώσεις που διαφαίνεται υποκείμενος οικογενειακός κίνδυνος για τη νόσο. Ο υποκείμενος οικογενειακός κίνδυνος μπορεί να προβλεφθεί από τη βοήθεια γνωστών μοντέλων [70] εκτίμησης κινδύνου που χρησιμοποιούν γενεαλογικές πληροφορίες συμπεριλαμβανομένης της ηλικίας της εμφάνισης των πληγέντων, από τη νόσο, συγγενών. Γίνεται όλο και περισσότερο δυνατόν να διαφοροποιηθούν καλύτερα οι γυναίκες σύμφωνα με τον υποκείμενο γενετικό κίνδυνο χρησιμοποιώντας βαθμολογίες με βάση τους πολυμορφισμούς των SNPs [71], [72]. Η πρόβλεψη οικογενειακού κινδύνου είναι πιθανό να βελτιωθεί με μεγαλύτερη έρευνες που σχετίζονται με τη μελέτη ολόκληρου του ανθρώπινου γονιδιώματος και τη χρήση πιο εξελιγμένων στατιστικών μεθόδων για τη δημιουργία καλύτερων μοντέλων υπολογισμού του κινδύνου βασισμένων στα SNPs και το αναλυτικό οικογενειακό ιστορικό [68].

Με τη χρήση του γονιδιακού τεστ για ανίχνευση υψηλού κινδύνου μεταλλάξεων σε γνωστά γονίδια που «ενοχοποιούνται» για τον καρκίνο του μαστού [73], [74], οι κλινικές που ασχολούνται με τον οικογενειακό καρκίνο πρόκειται να διαγνώσουν όλο και μεγαλύτερο αριθμό γυναικών με οικογενειακό ιστορικό καρκίνου του μαστού. Ωστόσο, διαπιστώνεται ότι η συντριπτική πλειονότητα αυτών των γυναικών δεν φαίνεται να φέρει κάποια επιβλαβή γονιδιακή μετάλλαξη. Επομένως, ένα βασικό και σημαντικό κλινικό ζήτημα είναι η στρατηγική διαχείρισης των εν λόγω γυναικών που έχουν «επιβαρυσμένο» οικογενειακό ιστορικό χωρίς όμως να φέρουν κάποια μετάλλαξη υψηλού κινδύνου για την εμφάνιση καρκίνου του μαστού. Προκειμένου να επιλυθεί αυτό το υψηλής σημασίας ζήτημα, είναι σημαντικό να γνωρίζουμε εάν η προαναφερθείσα κατηγορία γυναικών έχει τους ίδιους παράγοντες κινδύνου, αλλά και αν οι συσχετίσεις κινδύνου είναι ίδιας σημασίας με αυτές των γυναικών που ανήκουν στο γενικό πληθυσμό [68].

Το ζήτημα της πολλαπλασιαστικής και πρόσθετης αλληλεπίδρασης του οικογενειακού κινδύνου θα πρέπει να λαμβάνεται υπόψη για κάθε παράγοντα κινδύνου προκειμένου να δοθούν κατάλληλες κατευθυντήριες οδηγίες σε γυναίκες που βρίσκονται σε όλο το «φάσμα» του οικογενειακού κινδύνου. Η γνώση του βαθμού συσχέτισης της εμφάνισης της ασθένειας με τον οικογενειακό κίνδυνο θα επιφέρει εγκυρότερα και καταλληλότερα μέτρα πρόληψης για τις γυναίκες [75].

O J. L. Hopper και οι συνεργάτες του [68], διεξήγαγαν προοπτική μελέτη για γυναίκες διαφόρων ηλικιών, με δεδομένη την ύπαρξη οικογενειακού κινδύνου, προκειμένου να διαπιστώσουν την ενδεχόμενη συσχέτιση του BMI με την ηλικία της γυναίκας, την φάση της εμμηνόπαυσης και τον υποκείμενο οικογενειακό κίνδυνο. Αρχικά, για κάθε συμμετέχουσα, υπολογίστηκε ο κίνδυνος διεισδυτικού καρκίνου

του μαστού για ένα έτος και η ο κίνδυνος εμφάνισης της νόσου κατά τη διάρκεια ζωής της γυναίκας χρησιμοποιώντας τον αλγόριθμο BOADICEA (Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm) και συγκεκριμένα την Τρίτη έκδοση, όπου χρησιμοποιούνται γενεαλογικές πληροφορίες στη βάση της ανάλυσης. Ο αλγόριθμος BOADICEA χρησιμοποιεί πληροφορίες για τον καρκίνο του μαστού για γυναίκες και για άντρες, τον καρκίνο των ωοθηκών και την ηλικία διάγνωσης για συγγενείς πρώτου, δεύτερου και τρίτου βαθμού, την ημερομηνία γέννησης, τη γενικότερη κατάσταση υγείας, την ηλικία καταγραφής της «συνέντευξης» για την ένταξη στη μελέτη καθώς και περιστατικά που σχετίζονται με συγκεκριμένη ηλικία ή συγκεκριμένη χώρα για τον υπολογισμό του κινδύνου που αφορά στον καρκίνο του μαστού [76]. Σε περιπτώσεις που ήταν διαθέσιμη η πληροφορία για την ύπαρξη μεταλλάξεων BRCA1 και BRCA2 που «ενοχοποιούνται» για τον καρκίνο του μαστού, συμπεριλαμβάνονταν επίσης στον υπολογισμό του ρίσκου [77]. Η ευαισθησία της τεχνικής ανίχνευσης της μετάλλαξης θεωρήθηκε περίπου ίση με 70% και 80% για τα γονίδια BRCA1 και BRCA2, αντίστοιχα. Στην εν λόγω έρευνα διεθνούς εμβέλειας έρευνα [68] χρησιμοποιήθηκε ένας πολύ μεγάλος αριθμός γυναικών με οικογενειακό ιστορικό καρκίνου του μαστού (16.035 γυναίκες) και διαπιστώθηκε ότι ο απόλυτος βαθμός κινδύνου από τον καρκίνο του μαστού δεδομένου του BMI, αυξάνεται όσο αυξάνεται η ηλικία μετά την εμμηνόπαυση σε γυναίκες με υποκείμενο οικογενειακό κίνδυνο.

Η ερευνητική ομάδα των J. L. Hopper et al. [68] κατέληξε σε τρία βασικά συμπεράσματα, κλινικής και βιολογικής σημασίας. Πρώτον, διαπίστωσαν ότι ο μεγαλύτερος BMI σε νεαρή ηλικία σχετίζεται με μειωμένο κίνδυνο εμφάνισης καρκίνου του μαστού, όπως έχει καταγραφεί και σε άλλες έρευνες [78]. Ωστόσο, απέδειξαν ότι αυτή η αρνητική σχέση κινδύνου και BMI δεν μεταφράζεται σε ουσιαστική επίδραση στον απόλυτο κίνδυνο καρκίνου του μαστού. Δεύτερον, τα μοντέλα που χρησιμοποιήθηκαν επιβεβαιώνουν ότι ο BMI σχετίζεται με αύξηση του κινδύνου όταν μια γυναίκα βρίσκεται σε μετεμμηνοπαυσιακή φάση. Όσον αφορά στις διαφορές στον απόλυτο κίνδυνο εμφάνισης της νόσου, φαίνεται ότι η εν λόγω επίδραση στον απόλυτο κίνδυνο αυξάνεται με την ηλικία στις περιπτώσεις που η γυναίκα βρίσκεται στο τέλος της δεκαετίας των 50 και μετά. Τρίτον, τα μοντέλα των J. L. Hopper et al., προβλέπουν ότι όσο μεγαλύτερος είναι ο οικογενειακός κίνδυνος για μια γυναίκα, τόσο μεγαλύτερη είναι η επίδραση του BMI στον απόλυτο κίνδυνο εμφάνισης του καρκίνου του μαστού μετά την εμμηνόπαυση. Αυτή η διαπίστωση βασίζεται στο γεγονός ότι, σε όρους πολλαπλασιαστικού κινδύνου, η συσχέτιση του καρκίνου του μαστού με τον BMI δεν διέφερε για τις γυναίκες με διαφορετικό οικογενειακό κίνδυνο. Σε αντίθεση με τις περισσότερες άλλες μελέτες, η εμπλουτισμένη μελέτη των Hopper et al., συμπεριλάμβανε επαρκή στατιστικά στοιχεία ώστε να εξετάσει τις αλληλεπιδράσεις των παραγόντων κινδύνου με τον υποκείμενο οικογενειακό κίνδυνο [68].

Γενετικοί παράγοντες

Σύμφωνα με τις εκτιμήσεις του Εθνικού Ινστιτούτου Καρκίνου (NCI -National Cancer Institute) των ΗΠΑ για το 2018, περίπου το 10% των καρκίνων του μαστού συνδέεται με μια παθογόνο γενετική μετάλλαξη σε ένα από πολλά διαφορετικά γονίδια [79]. Περισσότερο από το 50% των παθογόνων γενετικών μεταλλάξεων αφορούν στα γονίδια BRCA1 και BRCA2. Η χρήση γενετικών εξετάσεων για τον εντοπισμό ασθενών που διατρέχουν αυξημένο κίνδυνο να αναπτύξουν καρκίνο του μαστού δίνει τη δυνατότητα στις ασθενείς να λάβουν μέτρα για τη μείωση αυτού του κινδύνου. Υπάρχουν διάφορες στρατηγικές διαχείρισης του κινδύνου που είναι διαθέσιμες για τα άτομα με αυξημένο κίνδυνο (π.χ. χημειοπροφύλαξη μαζί με ενισχυμένο έλεγχο - χειρουργικές επεμβάσεις μείωσης του κινδύνου). Δυστυχώς, ακόμη και σήμερα ένας σημαντικός αριθμός φορέων παθογόνων μεταλλάξεων παραμένει μη ανιχνεύσιμος και αδιάγνωστος. Πρόκειται σε μεγάλο βαθμό για γυναίκες με μεταλλάξεις «μέτριας διεισδυτικότητας», αλλά ακόμα και για γυναίκες με μεταλλάξεις BRCA1 ή 2 που μπορεί να μην εντοπιστούν [79].

Οι Momenimovahed et al. [7], πραγματοποίησαν ενδελεχή μελέτη όσον αφορά στα επιδημιολογικά χαρακτηριστικά και τους παράγοντες κινδύνου του καρκίνου του μαστού. Σύμφωνα με την εν λόγω ερευνητική ομάδα, παρά το γεγονός ότι και αρκετοί γενετικοί παράγοντες συμβάλλουν στη συχνότητα εμφάνισης του καρκίνου του μαστού, περίπου το 40% των κληρονομικών περιπτώσεων καρκίνου του μαστού οφείλεται σε μεταλλάξεις των γονιδίων BRCA1 και BRCA2 που κληρονομούνται με την επικρατούσα αυτοσωμική μέθοδο [7]. Τα αποτελέσματα της μελέτης των I. Godet et al. [80], δείχνουν ότι το 55% - 65% των φορέων της μετάλλαξης BRCA1 και το 45% των φορέων της μετάλλαξης BRCA2 αναπτύσσουν καρκίνο του μαστού μέχρι την ηλικία των 70 ετών. Με βάση μια άλλη προοπτική μελέτη κοόρτης [81], ο κίνδυνος αθροιστικού καρκίνου του μαστού μέχρι την ηλικία των 80 ετών ήταν 72% στους φορείς της μετάλλαξης BRCA1 (95% CI, 65% - 79%) και 69% στους φορείς της μετάλλαξης BRCA2 (95% CI, 61%-77%). Οι μεταβολές της ανθρώπινης ιντερφερόνης α-2b μπορεί να εμπλέκονται στην εμφάνιση και εξέλιξη του καρκίνου του μαστού εκτός από άλλους παράγοντες κινδύνου [82]. Σε μια άλλη μελέτη ελέγχου περιπτώσεων [83], οι πολυμορφισμοί της μεταλλοπρωτεϊνάσης της μήτρας (MMP-2 c-735-T) συσχετίστηκαν με αυξημένο κίνδυνο εμφάνισης καρκίνου του μαστού. Επιπλέον, το αλληλόμορφο MMP-2c μπορεί επίσης να αυξήσει τον κίνδυνο ανάπτυξης καρκίνου του μαστού σε νεότερη ηλικία κατά 1,64 φορές (OR, 1,64; 95% CI, 1,01-2,7) [83].

Σε μια πρόσφατη μελέτη των A. Lee et al. [84] παρουσιάζεται η επέκταση του αλγόριθμου BOADICEA προκειμένου να ενσωματώσει την επίδραση του πολυγονιδιακού κινδύνου (Polygenic Risk Score - PRS) και άλλων γενετικών παραγόντων κινδύνου. Ο αλγόριθμος BOADICEA ενσωματώνει τις επιδράσεις της μετάλλαξης των γονιδίων BRCA1, BRCA2, PALB2, CHEK2 και ATM, και τον PRS (Polygenic Risk Score) βασισμένο σε 313 SNPs που εξηγεί το 20% της πολυγονιδιακής διακύμανσης του καρκίνου του

μαστού. Επιπλέον περιλαμβάνει μια υπολειμματική πολυγονιδιακή συνιστώσα που αντιπροσωπεύει άλλες γενετικές/οικογενειακές επιδράσεις, γνωστούς ορμονικούς/αναπαραγωγικούς παράγοντες κινδύνου και παράγοντες που σχετίζονται με τον τρόπο ζωής και τη μαστογραφική πυκνότητα. Σύμφωνα με τα αποτελέσματα των A. Lee et al. [84], ακριβέστερη εκτίμηση του κινδύνου του καρκίνου του μαστού επιτυγχάνεται όταν όλοι οι γενετικοί παράγοντες και ο τρόπος ζωής, οι ορμονικοί/αναπαραγωγικοί παράγοντες λαμβάνονται υπόψη από κοινού. Συγκεκριμένα, λαμβάνοντας υπόψιν τους ανωτέρω παράγοντες, ο προβλεπόμενος κίνδυνος κατά τη διάρκεια της ζωής των γυναικών στον πληθυσμό του Ηνωμένου Βασιλείου, κυμαινόταν από 2,8% έως 30,6% με 14,7% των γυναικών να προβλέπεται να έχουν κίνδυνο ζωής $\geq 17\%$ και $< 30\%$ (μέτριος κίνδυνος σύμφωνα με τις κατευθυντήριες γραμμές του Εθνικού Ινστιτούτου Υγείας και Φροντίδας (National Institute for Health and Care Excellence - NICE) και το 1,1% είχαν κίνδυνο ζωής $\geq 30\%$ (υψηλός κίνδυνος) [84].

Κάπνισμα

Το κάπνισμα έχει ενοχοποιηθεί για την αιτιολογία μεγάλου αριθμού καρκίνων και υπάρχει πιθανότητα να εμπλέκεται βιολογικά και στην καρκινογένεση του μαστού. Παρόλα αυτά, τα επιδημιολογικά στοιχεία μέχρι σήμερα είναι ασυνεπή. Έρευνα που πραγματοποιήθηκε το 2014 [85], στόχευσε στη διερεύνηση του ρόλου του ενεργού καπνίσματος στον κίνδυνο εμφάνισης διηθητικού καρκίνου του μαστού χρησιμοποιώντας μια προοπτική μελέτη σε γυναίκες από το Καναδικό Μητρώο παρακολούθησης της διατροφής, του τρόπου ζωής και της υγείας. Η έρευνα πραγματοποιήθηκε με τη μέθοδο ελέγχου περιπτώσεων και δημιουργήθηκε μια ηλικιακά κατανομημένη υποομάδα 3314 γυναικών από τις 39.532 γυναίκες συμμετέχουσες που συμπλήρωσαν αναλυτικά ερωτηματολόγια με πληροφορίες για τον τρόπο ζωής και διατροφής τους. Συνολικά εντοπίστηκαν 1096 περιπτώσεις καρκίνου του μαστού στο σύνολο της μελέτης (συμπεριλαμβανομένων 141 περιπτώσεων από την υποομάδα που προαναφέρθηκε) μέσω σύνδεσης με το Καναδικό Μητρώο Καρκίνου. Χρησιμοποιήθηκαν διάφορα μοντέλα παλινδρόμησης για την εκτίμηση του βαθμού επικινδυνότητας του παράγοντα του καπνίσματος και τη συσχέτιση μεταξύ των διαφορετικών περιόδων έκθεσης στο κάπνισμα και του κινδύνου εμφάνισης καρκίνου του μαστού. Αφού εξετάστηκαν προσεκτικά όλες οι περιπτώσεις έκθεσης των γυναικών στο κάπνισμα στην πρόιμη ζωή τους, δεν βρέθηκε καμία συσχέτιση μεταξύ οποιασδήποτε έκθεσης στο κάπνισμα και του κινδύνου εμφάνισης καρκίνου του μαστού σε αυτή τη μελέτη (λόγος κινδύνου = 1,0 - 95%, διάστημα εμπιστοσύνης (Confidence Interval - CI) = 0,87-1,17 για τις καπνίστριες έναντι εκείνων που δεν κάπνισαν ποτέ). Αν και τα αποτελέσματα αυτά δεν μπορούν να αποκλείσουν τη συσχέτιση μεταξύ καπνίσματος και καρκίνου του μαστού, συμφωνούν με την τρέχουσα βιβλιογραφία που υποδηλώνει ότι, αν υπάρχει συσχέτιση, αυτή είναι σχετικά ασθενής [85].

Υπέρταση

Σύμφωνα με τους H. Han et al. [86], οι μελέτες παρατήρησης που εξετάζουν τη σχέση μεταξύ της υπέρτασης και του κινδύνου καρκίνου του μαστού έχουν αναφέρει αντικρουόμενα ευρήματα. Οι H. Han et al. [86] πραγματοποίησαν συστηματική ανασκόπηση και μετα-ανάλυση για να συνοψίσουν τα στοιχεία σχετικά με τη συσχέτιση μεταξύ της υπέρτασης και του κινδύνου εμφάνισης καρκίνου του μαστού. Οι επιλέξιμες μελέτες εντοπίστηκαν μέσω μιας ολοκληρωμένης βιβλιογραφικής αναζήτησης στις βιβλιοθήκες PubMed, EMBASE και Cochrane μέχρι τον Αύγουστο του 2016. Συμπεριέλαβαν μελέτες παρατήρησης που ανέφεραν τους σχετικούς κινδύνους (Relative Risks) με τα αντίστοιχα διαστήματα εμπιστοσύνης (CI - 95%). Τα αποτελέσματα από τις επιμέρους μελέτες εξήχθησαν με τη χρήση μοντέλου τυχαίων επιδράσεων, συγκεκριμένα, 29 άρθρα από 30 μελέτες, με συνολικά 11643 περιπτώσεις καρκίνου του μαστού, επιλέχθηκαν για να συμπεριληφθούν στη μετα-ανάλυση. Παρατηρήθηκε στατιστικά σημαντική συσχέτιση μεταξύ της υπέρτασης και του αυξημένου κινδύνου καρκίνου του μαστού (RR: 1,15, 95% CI: 1,08, 1,22). Σε ανάλυση που πραγματοποιήθηκε σε υποομάδα γυναικών, διαπιστώθηκε θετική συσχέτιση μεταξύ της υπέρτασης και της συχνότητας εμφάνισης καρκίνου του μαστού στις μετεμμηνοπαυσιακές γυναίκες (RR: 1,20; 95% CI: 1,09, 1,31). Αντίθετα, η υπέρταση δεν συσχετίστηκε με τον κίνδυνο εμφάνισης καρκίνου του μαστού στις προεμμηνοπαυσιακές γυναίκες (RR: 0,97; 95% CI: 0,84, 1,12) και σε γυναίκες Ασιατικής καταγωγής (RR: 1,07; 95% CI: 0,94, 1,22). Η εν λόγω μετα-ανάλυση υποδηλώνει συλλογικά μια σημαντική συσχέτιση μεταξύ της υπέρτασης και του κινδύνου εμφάνισης καρκίνου του μαστού, ειδικά για τις μετεμμηνοπαυσιακές γυναίκες [86].

Υψηλή χοληστερίνη

Στόχος της μελέτης ήταν να διεξαχθεί η πρώτη συστηματική ανασκόπηση και μετα-ανάλυση προοπτικών μελετών που διερευνούν τις συσχετίσεις μεταξύ των επιπέδων ολικής χοληστερόλης (Total Cholesterol TC), καλής χοληστερόλης (HDL-C), «βλαβερής» χοληστερόλης (LDL-C) και του κινδύνου εμφάνισης καρκίνου του μαστού. Οι σχετικές μελέτες εντοπίστηκαν στο PubMed (έως τον Ιανουάριο του 2014). Χρησιμοποιήθηκαν μοντέλα τυχαίων επιδράσεων για την εκτίμηση των λόγων κινδύνου (Hazard Ratios - HR) με 95% CI. Συνολικά, το συνοπτικό HR για τη συσχέτιση της ολικής χοληστερόλης με τον κίνδυνο καρκίνου του μαστού ήταν 0,97 και αυτό μεταξύ HDL-C και κινδύνου καρκίνου του μαστού ήταν 0,86, με υψηλή ετερογένεια. Για τις μελέτες που δεν έλαβαν υπόψιν την προκλινική μεροληψία, παρατηρήθηκε αντίστροφη συσχέτιση μεταξύ του κινδύνου καρκίνου του μαστού και της TC χοληστερόλης, και της HDL-C αντίστοιχα. Δεν παρατηρήθηκε συσχέτιση μεταξύ της LDL-C και του κινδύνου εμφάνισης καρκίνου του μαστού (σε τέσσερις μελέτες). Η εν λόγω μετα-ανάλυση επιβεβαιώνει τα στοιχεία για μια μέτρια αλλά στατιστικά σημαντική αντίστροφη συσχέτιση μεταξύ της TC και πιο συγκεκριμένα της HDL-C και του

κινδύνου εμφάνισης καρκίνου του μαστού, η οποία υποστηρίζεται από πολλές πειραματικές μελέτες. Απαιτούνται περαιτέρω μεγάλες προοπτικές μελέτες για να επιβεβαιωθούν τα αποτελέσματα σχετικά με το ρόλο των επιπέδων της χοληστερόλης στην αιτιολογία του καρκίνου του μαστού [87].

Διαταραχές στο θυρεοειδή αδένα

Έχει παρατηρηθεί συσχέτιση μεταξύ του καρκίνου του μαστού και των θυρεοειδικών (αυτοάνοσων) νοσημάτων ή της παρουσίας αντισωμάτων θυρεοειδικής υπεροξειδάσης (TPOAb, δείκτης αυτοάνοσου νοσήματος του θυρεοειδούς) [88]. Ωστόσο, ελάχιστα είναι γνωστά για το κατά πόσον οι γυναίκες με θυρεοειδικές (αυτοάνοσες) νόσους διατρέχουν αυξημένο κίνδυνο να αναπτύξουν καρκίνο του μαστού. Οι J. L. P. Kuijpers et al. [88], πραγματοποίησαν προοπτική μελέτη κοόρτης στην οποία διερεύνησαν κατά πόσον η παρουσία των TPOAb ή η δυσλειτουργία του θυρεοειδούς σχετίζεται με την παρουσία ή την ανάπτυξη καρκίνου του μαστού. Μια μη επιλεγμένη κοόρτη 2.775 γυναικών γύρω στην εμμηνόπαυση ελέγχθηκε για τις θυρεοειδικές παραμέτρους της θυρεοτροπίνης (TSH), της ελεύθερης θυροξίνης (FT(4)) και των TPOAb. Κατά τη διάρκεια της εν λόγω μελέτης, ελήφθησαν λεπτομερείς πληροφορίες σχετικά με προηγούμενες ή τρέχουσες διαταραχές του θυρεοειδούς και καρκίνο του μαστού, καθώς και σχετικά με πιθανούς παράγοντες που σχετίζονται με τον καρκίνο του μαστού και τις διαταραχές του θυρεοειδούς. Η κλινική δυσλειτουργία του θυρεοειδούς ορίστηκε τόσο με παθολογική τιμή της FT4 όσο και της TSH, και η υποκλινική δυσλειτουργία του θυρεοειδούς με παθολογική TSH (και φυσιολογική FT4). Η συγκέντρωση TPOAb ≥ 100 U/ml ορίστηκε ως θετική (TPOAb(+)). Η ομάδα μελέτης συνδέθηκε με το Μητρώο Καρκίνου του Eindhoven για την ανίχνευση όλων των γυναικών με (in situ) καρκίνο του μαστού (κωδικός ICD-O 174) που διαγνώστηκαν μεταξύ 1958 και 1994. Στη συνέχεια, στην προοπτική μελέτη, όλες οι γυναίκες που δεν είχαν καρκίνο του μαστού το 1994 (n = 2.738) παρακολούθηθηκαν μέχρι τον Ιούλιο του 2003 και καταγράφηκαν όλες οι νέες περιπτώσεις (in situ) καρκίνου του μαστού και όλοι, οι σχετιζόμενοι με τον καρκίνο, θάνατοι. Από τις 2.775 γυναίκες, 278 (10,0%) ήταν κατηγορίας TPOAb(+). Κατά τον έλεγχο του 1994, 37 γυναίκες (1,3%) είχαν καρκίνο του μαστού. Τα TPOAbs σχετίζονταν (ανεξάρτητα) με την τρέχουσα διάγνωση καρκίνου του μαστού (OR = 3,3, 95% CI 1,3-8,5). Από τις υπόλοιπες γυναίκες, 61 (2,2%) εμφάνισαν καρκίνο του μαστού. Ο νέος καρκίνος του μαστού σχετιζόταν με: (1) με προγενέστερη διάγνωση υποθυρεοειδισμού (OR = 3,8- 95% CI 1,3-10,9) (2) με τη χρήση θυρεοειδικών φαρμάκων (OR = 3,2- 95% CI 1,0-10,7) και (3) με χαμηλή FT4 (χαμηλότερο δέκατο εκατοστημόριο: OR = 2,3- 95% CI 1,2-4,6). Κατά τα πρώτα 3 έτη παρακολούθησης, η σχέση μεταξύ FT4 και TSH διαταράχθηκε στις γυναίκες με νέα διάγνωση καρκίνου του μαστού. Η παρουσία των TPOAb δεν σχετίστηκε με τον καρκίνο του μαστού κατά τη διάρκεια της παρακολούθησης σύμφωνα με τους J. L. P. Kuijpers et al. [88], οι οποίοι δεν κατέληξαν σε διαπίστωση άμεσης σχέσης μεταξύ αυτοάνοσης θυρεοειδίτιδας και καρκίνου του μαστού

αλλά σημείωσαν ότι ο υποθυρεοειδισμός και η χαμηλή FT4 σχετίζονται με αυξημένο κίνδυνο καρκίνου του μαστού σε μετεμμηνοπαυσιακές γυναίκες.

Η θυροξίνη (T4) έχει συσχετιστεί θετικά με τον πολλαπλασιασμό των καρκινικών κυττάρων, ενώ η επίδραση της τριωδοθυρονίνης (T3) στον πολλαπλασιασμό των κυττάρων δεν έχει τεκμηριωθεί επαρκώς, διότι διαφέρει ανάλογα με τον τύπο της κυτταρικής σειράς που χρησιμοποιείται. Στο Μεξικό, έχει αναφερθεί ότι το 14,5% των ενήλικων γυναικών που έχουν κάποιο είδος δυσλειτουργίας του θυρεοειδούς, έχουν παρατηρήσει την εμφάνιση άλλων ανωμαλιών και ασθενειών, συμπεριλαμβανομένου του καρκίνου του μαστού [89]. Αυτές οι ανωμαλίες ενδέχεται να μεταβάλλονται με το δείκτη μάζας σώματος, καθώς οι θυρεοειδικές ορμόνες εμπλέκονται στη ρύθμιση διαφόρων μεταβολικών μονοπατιών και πιθανώς σχετίζονται με την εμμηνοπαυσιακή κατάσταση, επειδή η παχυσαρκία έχει συσχετιστεί αρνητικά με τον καρκίνο του μαστού στις προεμμηνοπαυσιακές γυναίκες και έχει συσχετιστεί θετικά με αυτόν στις μετεμμηνοπαυσιακές γυναίκες [89].

Η ερευνητική ομάδα των C. O. Olvera et al. [89], είχε ως στόχο να αξιολογήσει τη συσχέτιση μεταξύ της συγκέντρωσης των θυρεοειδικών ορμονών (T4 και T3) στον ορό και του καρκίνου του μαστού αλλά και την επίδραση της παχυσαρκίας στην εν λόγω σχέση σε προεμμηνοπαυσιακές και μετεμμηνοπαυσιακές γυναίκες. Ως εκ τούτου, οι C. O. Olvera et al. [89] μέτρησαν τα επίπεδα των θυρεοειδικών ορμονών και των θυρεοειδικών αντισωμάτων σε 682 ασθενείς με αναφερόμενο καρκίνου του μαστού και 731 γυναίκες στην ομάδα ελέγχου, οι οποίες συμμετείχαν σε μια πληθυσμιακή μελέτη που πραγματοποιήθηκε από το 2004 έως το 2007 σε τρεις πολιτείες του Μεξικού. Κατά τη διάρκεια της μελέτης, έλεγξαν τη συσχέτιση της ολικής T4 (Total T4 - TT4) και της ολικής T3 (Total T3 - TT3), ανάλογα με την κατάσταση εμμηνόπαυσης και τον δείκτη μάζας σώματος, συμπεριλαμβάνοντας και άλλους παράγοντες κινδύνου υγείας και δημογραφικούς παράγοντες χρησιμοποιώντας μοντέλα λογιστικής παλινδρόμησης (Logistic Regression – LR).

Σύμφωνα με τα αποτελέσματα των C. O. Olvera et al. [89], οι υψηλότερες συγκεντρώσεις TT4 στον ορό συσχετίστηκαν με τον καρκίνο του μαστού τόσο στις προεμμηνοπαυσιακές (λόγος πιθανοτήτων (OR) ανά τυπική απόκλιση = 5,98, 95% CI 3,01-11,90) όσο και στις μετεμμηνοπαυσιακές γυναίκες (OR ανά τυπική απόκλιση = 2,81, 95% CI 2,17-3,65). Στις προεμμηνοπαυσιακές γυναίκες, η επίδραση της TT4 μειώθηκε καθώς αυξανόταν ο BMI, ενώ το αντίθετο παρατηρήθηκε στις μετεμμηνοπαυσιακές γυναίκες. Η σημασία της τροποποίησης της επίδρασης ήταν οριακή ($p = 0,059$) στις μετεμμηνοπαυσιακές γυναίκες και δεν ήταν σημαντική στις προεμμηνοπαυσιακές γυναίκες ($p = 0,22$). Οι χαμηλότερες συγκεντρώσεις TT3 συσχετίστηκαν με τον καρκίνο του μαστού τόσο στις προεμμηνοπαυσιακές όσο και στις μετεμμηνοπαυσιακές γυναίκες και δεν παρατηρήθηκε τροποποίηση της επίδρασης. Οι C. O. Olvera et al.

[89], κατέληξαν στο συμπέρασμα ότι υπάρχει ισχυρή συσχέτιση μεταξύ του καρκίνου του μαστού και των συγκεντρώσεων των TT3 και TT4 στον ορό και θα πρέπει να διερευνηθεί περαιτέρω η εν λόγω συσχέτιση.

Μια ακόμη ερευνητική ομάδα, των A. Angelousi et al. [90], μελέτησε τα βασικά επίπεδα θυρεοτροπίνης, ελεύθερης τριωδοθυρονίνης, ελεύθερης θυροξίνης και θυροειδικών αυτοαντισωμάτων σε 97 γυναίκες με πρωτοπαθή καρκίνο του μαστού, 27 γυναίκες με καλοήγη νόσο του μαστού και 4 γυναίκες με άτυπη υπερπλασία των πόρων. Τα βασικά επίπεδα που μετρήθηκαν, συγκρίθηκαν με εκείνα 48 υγιών γυναικών με φυσιολογική μαστογραφία τα τελευταία 2 χρόνια. Οι εν λόγω ερευνητές δε διαπίστωσαν σημαντικές συσχετίσεις μεταξύ του ιστορικού θυροειδικής νόσου και του καρκίνου του μαστού. Τα αρχικά μέσα επίπεδα της τριωδοθυρονίνης και της θυρεοτροπίνης δεν διέφεραν σημαντικά μεταξύ των συγκρινόμενων ομάδων σε αντίθεση με τα επίπεδα ελεύθερης θυροξίνης που βρέθηκαν σημαντικά υψηλότερα στην ομάδα γυναικών με καρκίνο του μαστού, ακόμη και μετά τη λήψη θεραπείας υποκατάστασης της λειτουργίας του θυροειδούς. Η παρουσία θυροειδικών αντισωμάτων δεν διέφερε σημαντικά μεταξύ των συγκρινόμενων ομάδων. Σε μια ανάλυση που πραγματοποίησαν οι A. Angelousi et al. [90] σε υποομάδα γυναικών, διαπιστώθηκε ότι οι περιπτώσεις συνύπαρξης καρκίνου του μαστού με θυροειδική νόσο και ιδιαίτερα υποθυροειδισμό είχαν σημαντικά χαμηλότερη επίπτωση λεμφαδενικών μεταστάσεων σε σύγκριση με τις περιπτώσεις καρκίνου του μαστού χωρίς θυροειδική νόσο. Τα δεδομένα των A. Angelousi et al. [90] επιβεβαίωσαν την πολλαπλασιαστική επίδραση των θυροειδικών ορμονών στα κύτταρα του μαστού, η οποία είχε ήδη αποδειχθεί σε άλλες μέletes *in vitro*. Επιπλέον, η θυροειδική νόσος και ιδιαίτερα η υποθυροειδική λειτουργία φάνηκε να σχετίζεται με χαμηλότερη επίπτωση λεμφαδενικών μεταστάσεων. Εν κατακλείδι, απαιτούνται περαιτέρω μελέτες για τον προσδιορισμό του προγνωστικού ρόλου των θυροειδικών ορμονών στον καρκίνο του μαστού [90].

2.7 Θεραπεία του Καρκίνου του Μαστού

Ο σχεδιασμός του θεραπευτικού πλάνου αντιμετώπισης του καρκίνου του μαστού, γίνεται από μια διεπιστημονική ομάδα ειδικών προκειμένου να αναλυθεί διεξοδικά η εκάστοτε περίπτωση και να ληφθεί η βέλτιστη απόφαση για την πορεία της ασθενούς. Η προσέγγιση για την αντιμετώπιση της κάθε περίπτωσης καρκίνου του μαστού είναι απόλυτα ανεξάρτητη και ξεχωριστή. Ανάλογα την βιολογία, τα παθολογικά χαρακτηριστικά, την συμπεριφορά του καρκίνου και τις προσωπικές επιθυμίες της ασθενούς, επιλέγεται η έκταση και μορφή του θεραπευτικού σχήματος. Μερικοί από τους παράγοντες στους οποίους βασίζεται η απόφαση των ειδικών είναι ο υπότυπος του καρκίνου, συμπεριλαμβανομένης της ορμονικής υποδοχής, το στάδιο, τα διαθέσιμα γενετικά δεδομένα αν υπάρχουν, και φυσικά τα κλινικά στοιχεία της ασθενούς. Αφού, λοιπόν, γίνει η διάγνωση και ταυτοποιηθούν οι παραπάνω παράγοντες, γίνεται η βέλτιστη προσωποποιημένη επιλογή για την ασθενή [91].

2.7.1 Τρόποι Αντιμετώπισης του Καρκίνου του Μαστού

Ένα πλάνο θεραπείας καρκίνου του μαστού περιλαμβάνει επιλογές όπως το χειρουργείο, η ακτινοθεραπεία, η χημειοθεραπεία, η ορμονοθεραπεία, η στοχευμένη θεραπεία και η ανοσοθεραπεία. Το χειρουργείο, όπως και η ακτινοθεραπεία, αποτελούν τοπική θεραπεία (local therapy), κάτι που σημαίνει ότι εκτείνεται στην γύρω περιοχή που εμφανίζεται ο καρκίνος χωρίς να επηρεάζει το υπόλοιπο σώμα. Παράλληλα, υπάρχει η συστηματική θεραπεία (systemic therapy), στην οποία χρησιμοποιούνται ουσίες που, μέσω της κυκλοφορίας του αίματος, επιδρούν στα κύτταρα σχεδόν όλου του σώματος. Ο στόχος της συστηματικής θεραπείας είναι η επίδραση στα καρκινικά κύτταρα, τα οποία μπορεί να έχουν φτάσει σε άλλα σημεία του σώματος. Η χημειοθεραπεία, η ορμονοθεραπεία, η στοχευμένη θεραπεία και η ανοσοθεραπεία ανήκουν στην κατηγορία της συστηματικής θεραπείας. Επειδή, σχεδόν όλες οι περιπτώσεις καρκίνου του μαστού αντιμετωπίζονται αρχικά με εγχείρηση, όλες οι υπόλοιπες μέθοδοι ονομάζονται επικουρικές (adjuvant), είναι δηλαδή συμπληρωματικές του χειρουργείου. Στην περίπτωση που προηγηθεί της επέμβασης κάποια θεραπεία, αυτή χαρακτηρίζεται ως νέο-επικουρική (neoadjuvant). Πιο αναλυτικά, οι διαθέσιμες θεραπείες είναι οι εξής [91]:

✚ Χειρουργική επέμβαση

Η αφαίρεση της ορατής νόσου είναι το πρώτο βήμα για αντιμετώπισή της. Οι περισσότερες γυναίκες με καρκίνο του μαστού, λοιπόν, υπόκεινται σε κάποιο είδος χειρουργικής επέμβασης ως μέρος της θεραπείας τους. Αυτή μπορεί να είναι είτε αφαίρεση όλου του μαστού με ολική μαστεκτομή είτε αφαίρεση μόνο του όγκου. Το δεύτερο ονομάζεται χειρουργείο διατήρησης του μαστού και στόχος είναι η αφαίρεση μόνο του τμήματος του μαστού που περιλαμβάνει τον όγκο, καθώς και ένα μέρος του περιβάλλοντα φυσιολογικού ιστού για ασφάλεια. Εκτός του καρκινικού όγκου, ο χειρουργός καλείται να αναλύσει την κατάσταση των λεμφαδένων. Για αυτό τον λόγο, ελέγχεται με βιοψία ο λεμφαδένας φρουρός, ώστε να καθοριστούν οι διηθημένοι αδένες και επιλέγεται αφαίρεση όσων είναι εγγύτερα στη μασχάλη. Αν η πρώτη ή μετέπειτα βιοψία δείξει καρκινικά κύτταρα σε περισσότερους λεμφαδένες, γίνεται μασχαλιαίος λεμφαδενικός καθαρισμός, όπου αφαιρείται ο μαλακός ιστός κάτω από τη μασχάλη [91].

✚ Ακτινοθεραπεία

Η χρήση υψηλής ενέργειας ακτινοβολίας προκειμένου να θανατωθούν τα καρκινικά κύτταρα είναι τοπική επικουρική θεραπεία. Η ακτινοθεραπεία ξεκινά αφού έχει επουλωθεί το σημείο της επέμβασης και μετά τη χημειοθεραπεία. Η ακτινοβολία προκαλεί τέτοια βλάβη που τα καρκινικά κύτταρα δεν είναι ικανά να επανέλθουν, σε σχέση με τα φυσιολογικά. Η ακτινοθεραπεία συνιστάται σχεδόν σε όλους τους διηθητικούς καρκίνους. Επιπλέον, στις περιπτώσεις χειρουργείου διατήρησης του μαστού, με ακτινοβόληση όλου του μαστού και στις περιπτώσεις μαστεκτομής με

πολύ μεγάλο όγκο και/ή διηθημένους λεμφαδένες, με ακτινοβόληση του θωρακικού τοιχώματος, μειώνεται ο κίνδυνος επανεμφάνισης καρκίνου στο ίδιο στήθος ή στους κοντινούς λεμφαδένες. Το πεδίο ακτινοβόλησης επεκτείνεται πάνω από την κλείδα και πίσω από το στήθος, αν έχουν βρεθεί καρκινικά κύτταρα στους λεμφαδένες αυτών των περιοχών. Η απαραίτητη δόση ακτινοβολίας που λαμβάνεται, χωρίζεται σε ημερήσιες συνεδρίες ακτινοθεραπείας η οποία διαρκεί περίπου 6 - 7 εβδομάδες. Με αυτόν τον τρόπο βελτιώνεται η πιθανότητα μακροπρόθεσμου ελέγχου του όγκου και μειώνεται η βλάβη στους υγιείς ιστούς [91].

✚ Χημειοθεραπεία

Η χημειοθεραπεία είναι μέρος της συστηματικής θεραπείας και μπορεί να είναι επικουρική ή νέο-επικουρική. Η νέο-επικουρική χημειοθεραπεία χορηγείται στις περιπτώσεις που απαιτείται συρρίκνωση του όγκου πριν την αφαίρεση. Στόχος της επικουρικής είναι η καταστροφή κυττάρων που έχουν μεταναστεύσει από τον αρχικό όγκο και δεν ανιχνεύονται με τα απεικονιστικά μέσα. Χορηγείται είτε ενδοφλέβια είτε διά του στόματος. Συνήθως είναι ένας συνδυασμός 2 ή 3 φαρμάκων, τα οποία χορηγούνται βάσει του ιατρικού πρωτοκόλλου. Για τον καρκίνο του μαστού, η θεραπεία δίνεται, συνήθως, σε 4 με 8 κύκλους, ένας κύκλος αφορά στη χρονική περίοδο 2 με 4 εβδομάδων και τα φάρμακα χορηγούνται σε συγκεκριμένη δόση, διάρκεια και σειρά, περιλαμβανομένης και μιας περιόδου ανάπαυσης πριν την έναρξη του επόμενου κύκλου, ώστε να αναρρώνει η ασθενής από τις παρενέργειες [91].

✚ Ορμονοθεραπεία

Η ορμονοθεραπεία, ή αλλιώς ενδοκρινική θεραπεία, γίνεται μέρος του θεραπευτικού πλάνου εάν οι υποδοχείς ορμονών είναι θετικοί. Πρόκειται για επικουρική συστηματική θεραπεία και μπορεί να χρησιμοποιηθεί μόνη ή συνδυαστικά με χημειοθεραπεία και ακτινοθεραπεία. Τα καρκινικά κύτταρα που έχουν θετικούς υποδοχείς μπορεί να χρειάζονται οιστρογόνα και/ή προγεστερόνη για να μεγαλώσουν, οπότε σταματάει η ανάπτυξή τους ή πεθαίνουν όταν αντιμετωπίζονται με ουσίες που αναστέλλουν τη σύνδεση και τη δράση των οιστρογόνων. Όπως και οι υπόλοιπες θεραπείες, βοηθά στη μείωση του κινδύνου επανεμφάνισης του καρκίνου. Υπάρχουν διάφορα είδη ορμονοθεραπείας για τον καρκίνο του μαστού. Τα περισσότερα αυτών είτε μειώνουν το επίπεδο των οιστρογόνων, είτε εμποδίζουν εξολοκλήρου την επίδρασή τους στα καρκινικά κύτταρα. Η επιλογή της ορμονοθεραπείας βασίζεται στην εμμηνοπαυσιακή κατάσταση της ασθενούς [91].

✚ Στοχευμένη θεραπεία

Μία θεραπεία που επιτίθεται σε συγκεκριμένα καρκινικά κύτταρα ονομάζεται στοχευμένη. Χρησιμοποιεί συνήθως φάρμακα ή άλλες ουσίες, όπως τα μονοκλωνικά αντισώματα που εντοπίζουν τις αλλαγές στα κύτταρα που προκαλούν καρκίνο. Οι στοχευμένες θεραπείες μπορεί να έχουν λιγότερες ανεπιθύμητες ενέργειες από άλλους τύπους αντικαρκινικής θεραπείας και να

έχουν αποτελέσματα όταν άλλες μέθοδοι αποτυγχάνουν. Συγκεκριμένα για την HER2, σε μία στις πέντε γυναίκες με καρκίνο του μαστού παρατηρείται υπερέκφραση αυτής στην επιφάνεια των καρκινικών κυττάρων. Η HER2 είναι μία πρωτεΐνη που προσδίδει επιθετικά χαρακτηριστικά στον καρκίνο του μαστού και ενεργοποιεί την ανάπτυξή του. Μια HER2 στοχεύουσα θεραπεία, λοιπόν, αποτελείται από ουσίες που προσκολλώνται στην πρωτεΐνη αυτή και σταματούν την ανάπτυξη των κυττάρων [91].

Ανοσοθεραπεία

Τα φάρμακα ανοσοθεραπείας χρησιμοποιούν τη δύναμη του ανοσοποιητικού συστήματος του σώματος για να επιτεθούν στον καρκίνο. Στην πραγματικότητα ενισχύουν τη δράση του ανοσοποιητικού, ώστε να εντοπίζει και να καταστρέφει τα καρκινικά κύτταρα με μεγαλύτερη αποτελεσματικότητα. Δεν θεωρούνται τα πλέον χρήσιμα στην αντιμετώπιση της νόσου, αλλά έρευνες αποδεικνύουν ότι έχουν τη δυνατότητα να βελτιώσουν την επιβίωση και πρόγνωση των ασθενών [91].

2.7.2 Θεραπευτικό Σχήμα ανάλογα με το στάδιο της νόσου

Η κάθε θεραπεία αποτελείται από μία ή έναν συνδυασμό των παρεμβατικών μεθόδων της προηγούμενης ενότητας. Συνοδεύεται από οφέλη, κινδύνους και αντενδείξεις, για τα οποία η ασθενής πρέπει να είναι πλήρως ενημερωμένη [92].

Πλάνο για Στάδιο 0 - Μη Διηθητικό Πορογενές Καρκίνωμα (DCIS)

Τα συνηθέστερα ενδεχόμενα για μη διηθητικό καρκίνο είναι είτε αφαίρεση του όγκου με χειρουργείο διατήρησης του μαστού, είτε μαστεκτομή. Στην πρώτη περίπτωση ακολουθεί ακτινοθεραπεία του μαστού για να αποφευχθεί ο κίνδυνος υποτροπής, ενώ στη δεύτερη τυπικά δεν κρίνεται απαραίτητη. Στην περίπτωση που ο DCIS έχει θετικούς ορμονικούς υποδοχείς, συστήνεται η λήψη ορμονοθεραπείας με ταμοξιφένη, ένα φάρμακο που εξουδετερώνει τη δράση των οιστρογόνων στον μαστό και μειώνει τον κίνδυνο επανεμφάνισης της νόσου. Η ταμοξιφένη μειώνει, επίσης, τον κίνδυνο ανάπτυξης ετερόπλευρου καρκίνου του μαστού, δηλαδή καρκίνου στον άλλο μαστό. Για μετεμμηνοπαυσιακές γυναίκες χορηγείται αναστολέας αρωματάσης. Όσον αφορά στον LCIS επειδή τεχνικά δεν είναι καρκίνος απαιτείται παρακολούθηση και όχι θεραπεία [91].

Πλάνο για Στάδιο I έως III - Διηθητικό Καρκίνωμα (IDC + ILC)

Η θεραπεία σε αυτές τις περιπτώσεις στοχεύει τον καρκίνο τόσο τοπικά όσο και στα καρκινικά κύτταρα που πιθανώς έχουν επεκταθεί στον υπόλοιπο οργανισμό. Σε όλες τις περιπτώσεις απαιτείται διεξαγωγή χειρουργείου. Για όγκους των οποίων το μέγεθος κριθεί μεγάλο προτιμάται, αρχικά, μία νέο-επικουρική

συστηματική θεραπεία, ώστε να συρρικνωθεί το μέγεθός τους. Αυτό συμβαίνει συνήθως στους καρκίνους Σταδίου III και έτσι διευκολύνεται η τοπική θεραπεία και συχνά η διατήρηση του μαστού.

Κατά την διάρκεια του χειρουργείου γίνεται ογκεκτομή ή μαστεκτομή και ταυτόχρονος έλεγχος του λεμφαδένα φρουρού. Αν είναι δυνατόν αφαιρούνται οι διηθημένοι λεμφαδένες κατά την ίδια επέμβαση, μετά από βιοψία, ωστόσο πολλές φορές είναι αναγκαίο και δεύτερο χειρουργείο για λεμφαδενικό καθαρισμό. Η τοπική θεραπεία ολοκληρώνεται με ακτινοβολία στις περιπτώσεις ογκεκτομής, αλλά και στην ολική μαστεκτομή, αν έχουν εντοπιστεί καρκινικά κύτταρα σε λεμφαδένες. Η ακτινοθεραπεία γίνεται μετά το χειρουργείο, αν δεν απαιτείται επικουρική χημειοθεραπεία, αλλιώς καθυστερεί μέχρι να ολοκληρωθεί εκείνη.

Μετά την επέμβαση, ακολουθεί η επικουρική συστηματική θεραπεία, στόχος της οποίας είναι η επίδραση σε καρκινικά κύτταρα τα οποία μπορεί να έχουν φτάσει σε άλλα σημεία του σώματος. Τα χαρακτηριστικά του όγκου θα καθορίσουν έναν συνδυασμό ορμονικής θεραπείας, χημειοθεραπείας και HER2 στοχεύουσας θεραπείας, προκειμένου να επιτευχθεί η βέλτιστη πρόγνωση για την ασθενή [91].

Πλάνο για Στάδιο IV – Μεταστατικός Καρκίνος

Η αντιμετώπιση για την πλειονότητα των γυναικών με καρκίνο Σταδίου IV γίνεται με συστηματική θεραπεία. Η θεραπεία αυτή αποτελεί μονόδρομο δεδομένης της ευρείας επέκτασης της νόσου σε άλλα όργανα. Αυτό μπορεί να περιλαμβάνει ορμονοθεραπεία, χημειοθεραπεία, στοχευμένη θεραπεία, ανοσοθεραπεία ή έναν συνδυασμό όλων. Το χειρουργείο και η ακτινοθεραπεία βοηθά σε συγκεκριμένες περιπτώσεις. Τα παραπάνω μπορούν να βελτιώσουν τα συμπτώματα και να παρατείνουν τη ζωή της ασθενούς, αλλά αυτοί οι καρκίνοι θεωρούνται ανίατοι. Συχνά, σε ασθενείς με μεταστατικό καρκίνο εφαρμόζονται κλινικές μελέτες με πειραματικά και νέα φάρμακα. Η θεραπεία ακολουθείται μόνο όσο οι παρενέργειες είναι ανεκτές, αφού στόχος είναι να διατηρηθεί ή να βελτιωθεί η ποιότητα ζωής [91].

Κεφάλαιο 3^ο

Διαγνωστικές μέθοδοι για την Ανίχνευση του Καρκίνου του Μαστού – Ανασκόπηση των Τεχνικών και της Απόδοσής τους

3.1 Διαγνωστικές Μέθοδοι για την Ανίχνευση του Καρκίνου του Μαστού

Ο καρκίνος του μαστού είναι μια ετερογενής νόσος που αντανακλάται σε διαφορετικές κλινικές ενδείξεις. Ένα όλο και αυξανόμενο ποσοστό των καρκίνων του μαστού διαγιγνώσκεται σε προκλινική φάση μέσω της διαλογής, δηλαδή πριν από την ανάπτυξη κλινικών ενδείξεων. Το ποσοστό των καρκίνων του μαστού που διαγιγνώσκονται σε προκλινικό στάδιο ποικίλλει σημαντικά μεταξύ των χωρών και εξαρτάται από την παρουσία, την ποιότητα και τη συχνότητα του οργανωμένου μαστογραφικού ελέγχου, την πρόσβαση στα διαφορετικά διαγνωστικά απεικονιστικά μέσα, την ηλικία της ασθενούς και το επίπεδο ενημέρωσης του πληθυσμού για τον καρκίνο του μαστού [93].

Οι διαγνωστικές εξετάσεις ανιχνεύουν, απειλητικές για τη ζωή, ασθένειες σε πρώιμα στάδια. Η έγκαιρη διάγνωση της νόσου μπορεί να οδηγήσει σε βελτιωμένα αποτελέσματα σε όσους πάσχουν και να μειώσει τη συχνότητα εμφάνισης της νόσου σε μεταγενέστερα στάδια. Η πρόληψη έχει εξέχουσα σημασία, ειδικά στον καρκίνο του μαστού, ο οποίος είναι ιάσιμος. Η διατήρηση της υγείας βασίζεται στην έγκαιρη διάγνωση, που εξασφαλίζεται με τη σειρά της από την προληπτική συμπεριφορά και τον συχνό έλεγχο όσον αφορά στα ιατρικά θέματα. Ο καρκίνος του μαστού, με βάση τα επιδημιολογικά δεδομένα, αποτελεί αναμφίβολα ένα μεγάλο ζήτημα υγείας με επιπτώσεις τόσο σε ατομικό όσο και σε κοινωνικό επίπεδο. Το γεγονός αυτό κάνει ακόμα πιο επιτακτική την ανάγκη της πρόληψης και της ανίχνευσης τυχόν κακοήθειας σε πρωταρχικό στάδιο [93], [94].

Στην πρόληψη του καρκίνου του μαστού μπορούν να συνδράμουν ενέργειες που περιορίζουν τους περιβαλλοντικούς παράγοντες κινδύνου. Αυτό συμβαίνει διότι είναι οι μόνοι που μπορούν να ελεγχθούν ως ένα βαθμό από τις ατομικές επιλογές. Για παράδειγμα, η υιοθέτηση ενός υγιεινού τρόπου ζωής που περιλαμβάνει άθληση, αποφυγή βλαβερών ουσιών, όπως το κάπνισμα και το αλκοόλ που ενοχοποιούνται

για εμφάνιση διάφορων καρκίνων, η σωστή διατροφή και η μειωμένη χρήση φαρμακευτικών ουσιών, είναι μια εν δυνάμει προσωπική στρατηγική για την ελαχιστοποίηση της πιθανότητας εμφάνισης της νόσου. Μάλιστα, έρευνες έχουν αποδείξει ότι η μεσαία προς έντονη φυσική άσκηση οδηγεί σε μείωση του κινδύνου καρκίνου του μαστού μέχρι και 14% σε εμμηνόπαυσιακές γυναίκες. Παράλληλα, στην ίδια ηλικιακή ομάδα, η μείωση του βάρους μπορεί να επιφέρει έως και 50% μείωση του κινδύνου. Δεν υπάρχει εγγυημένος τρόπος να αποτρέψει κάποιος την ασθένεια, ωστόσο η αλλαγή στις καθημερινές συνήθειες συνίσταται από τους ειδικούς [94].

Τον καθοριστικότερο ρόλο στην πρόληψη του καρκίνου του μαστού διαδραματίζουν τα διαθέσιμα ιατρικά μέσα προληπτικού ελέγχου και ανίχνευσης καρκινικών όγκων. Οι ανιχνευτικές ή προσυμπτωματικές εξετάσεις (screening tests) επικουρούν στην ταυτοποίηση ατόμων υψηλού κινδύνου για καρκίνο του μαστού, πριν εμφανίσουν συμπτώματα ή συνειδητοποιήσουν ότι διατρέχουν κίνδυνο, ώστε να λάβουν προληπτικά μέτρα. Οι εξετάσεις αυτές αποτελούν σημαντικό μέρος της προληπτικής ιατρικής. Τα screening tests βοηθούν στην ανίχνευση του καρκίνου του μαστού στα πρώιμα και πιο θεραπεύσιμα στάδιά του. Η τελική διάγνωση και η κατάταξη ενός όγκου σε καλοήγη ή κακοήγη προκύπτει μετά την παθολογοανατομική ανάλυση, ωστόσο η διαδικασία που οδηγεί στα τελικά συμπεράσματα περιλαμβάνει μια τριπλή εκτίμηση: το ιστορικό και την κλινική εξέταση, την απεικονιστική μέθοδο και τη βιοψία. Πολλές φορές, όμως, μία πρώιμη ανίχνευση έχει επιτευχθεί χάρη στην πρακτική της αυτοεξέτασης του μαστού από τις ίδιες τις γυναίκες [93].

Σκοπός της πρώιμης διάγνωσης είναι η μείωση της θνησιμότητας εξαιτίας του καρκίνου του μαστού. Σε σύγκριση με τον γενικό πληθυσμό, οι γυναίκες με βεβαρυμμένο οικογενειακό ιστορικό καρκίνου του μαστού έχουν αυξημένο κίνδυνο εμφάνισης της νόσου. Στο πλαίσιο της πρόληψης περιλαμβάνονται τα γενετικά τεστ για εύρεση επικίνδυνων γονιδίων και η συμβουλευτική σε περιπτώσεις που ενοχοποιούνται για οικογενειακό κληρονομικό καρκίνο του μαστού. Οι γνώμες των ειδικών, όμως, είναι διχασμένες σχετικά με τα οφέλη και τα αρνητικά αυτής της πρακτικής. Στις περιπτώσεις πολύ αυξημένου κινδύνου εμφάνισης καρκίνου στο μαστό, η λήψη φαρμάκων ή ακόμα και προληπτικά/προφυλακτικά χειρουργεία υπάρχουν επίσης ως διαθέσιμη επιλογή [95]. Προκειμένου να επιτευχθεί η ανίχνευση σε όσο πιο αρχικό στάδιο γίνεται, στην συγκεκριμένη κατηγορία, οι γενικές οδηγίες περιλαμβάνουν τον συχνό κλινικό έλεγχο του μαστού και ετήσια μαστογραφία ξεκινώντας από την ηλικία των 25 - 30, ή δέκα χρόνια νωρίτερα από την ηλικία εμφάνισης της νόσου σε συγγενικό πρόσωπο. Σε άλλη περίπτωση, δηλαδή μεσαίας επικινδυνότητας, οι ειδικοί συστήνουν ετήσιο έλεγχο με μαστογραφία σε γυναίκες από 40 χρονών και άνω [96].

Υπάρχουν διαφορετικά στάδια πρόληψης και διάγνωσης τυχόν καρκινικών αλλοιώσεων. Ξεκινώντας από το ατομικό επίπεδο και προσωπικό έλεγχο, συνεχίζοντας με εξετάσεις ρουτίνας και καταλήγοντας σε πιο εξειδικευμένες μεθόδους που περιλαμβάνουν απεικονιστικά τεστ και βιοψίες, μια γυναίκα μπορεί να επιτύχει τη βέλτιστη παρακολούθηση της υγείας των μαστών της. Όπως αναφέρθηκε, για να εξαχθούν τα σωστά τελικά διαγνωστικά συμπεράσματα πρέπει να ακολουθείται μία πλήρης και λεπτομερής διαδικασία ελέγχου όπως περιγράφεται στα παρακάτω υποκεφάλαια.

3.1.1 Απεικονιστικά Διαγνωστικά Μέσα

Τα απεικονιστικά διαγνωστικά μέσα αποτελούν μέρος του προληπτικού και συστηματικού ελέγχου του μαστού και είναι σε θέση να εντοπίσουν την ασθένεια πριν ακόμα εμφανιστούν συμπτώματα. Έχοντας ξεκάθαρη εικόνα των ευρημάτων γίνεται μια ορθότερη εκτίμηση της κατάστασης, με στόχο να αποφασιστούν τα επόμενα βήματα. Οι βασικότερες απεικονιστικές εξετάσεις που επιλέγονται για τον καρκίνο του μαστού είναι η μαστογραφία, το υπερηχογράφημα μαστού και η μαγνητική τομογραφία μαστών.

3.1.1.1 Μαστογραφία

Η μαστογραφία είναι η κύρια μέθοδος απεικόνισης του μαστού. Η εν λόγω τεχνική διαφοροποιεί τους ιστούς με βάση την πυκνότητά τους χρησιμοποιώντας ιονίζουσα ακτινοβολία χαμηλής ενέργειας. Η δόση ακτινοβολίας από τη μαστογραφία είναι πολύ χαμηλή (το πολύ 3 mGy). Η επεξεργασία της ψηφιακής μαστογραφίας πλήρους πεδίου βελτιώνει την ποιότητα της εικόνας, επιτρέπει την καλύτερη απεικόνιση του παρεγχύματος του μαστού, του δέρματος και του υποδόριου ιστού, διευκολύνει την προβολή και την αρχειοθέτηση της εικόνας και μειώνει περαιτέρω τη δόση ακτινοβολίας [97].

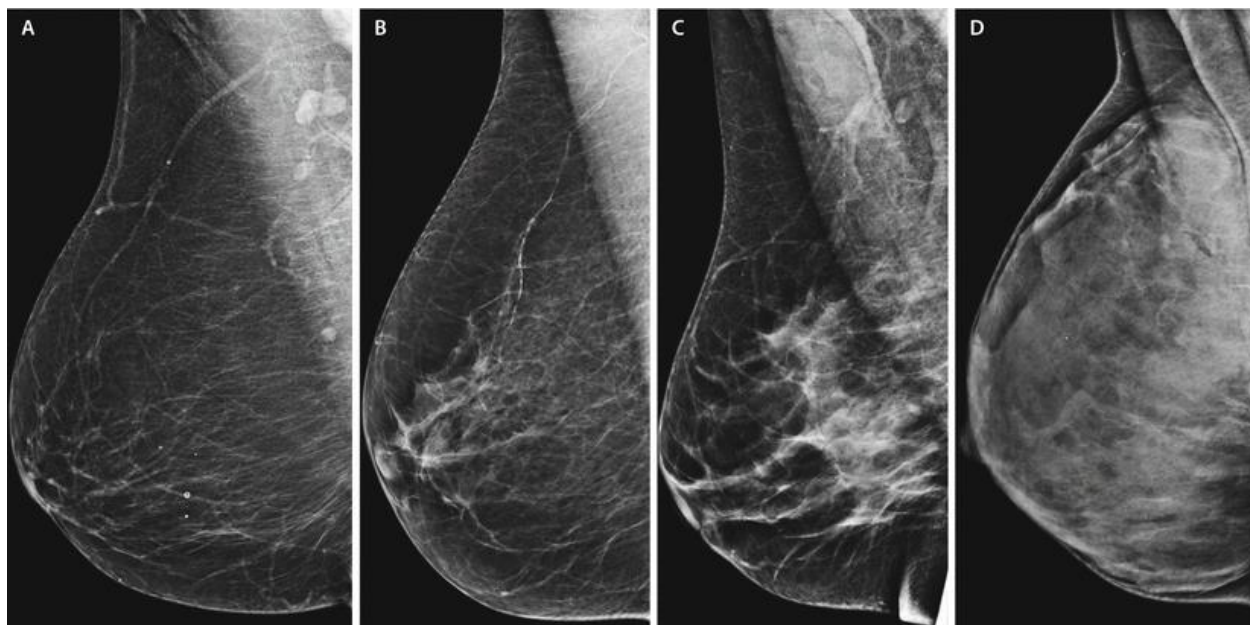
Η μαστογραφία χρησιμοποιείται τόσο για τυπικό προληπτικό έλεγχο όσο και για διαγνωστικό έλεγχο με ύπαρξη συμπτωμάτων. Οι ενδείξεις για τη μαστογραφία περιλαμβάνουν [97]:

- ✚ Προληπτικό έλεγχο για καρκίνο του μαστού
- ✚ Αξιολόγηση ασθενών με κλινικά συμπτώματα
- ✚ Καθοδήγηση εικόνας για βιοψία
- ✚ Προεγχειρητική σταδιοποίηση
- ✚ Προεγχειρητικός εντοπισμός
- ✚ Παρακολούθηση της θεραπείας και της γενικότερης αποκατάστασης της υγείας της ασθενούς

Οι μαστοί όπως προαναφέραμε αποτελούνται από ινοαδενικό ιστό και λιπώδη ιστό, η αναλογία αυτών των δύο καθορίζει την πυκνότητά του μαστού. Όσο περισσότερο ινοαδενικό ιστό έχει ο μαστός, τόσο μεγαλύτερη είναι η πυκνότητα του και τόσο πιθανότερο είναι οι μικρές βλάβες να αποκρύπτονται από το υπόβαθρο και να μειώνεται η ευαισθησία [97].

Η πυκνότητα του μαστού ταξινομείται σε τέσσερις κατηγορίες σύμφωνα με το σύστημα BI-RADS (Breast Imaging Reporting and Data System) (βλ. Εικόνα 11) [97], [98]:

- A. Οι μαστοί είναι σχεδόν εξ ολοκλήρου λιπώδεις.
- B. Υπάρχουν διάσπαρτες περιοχές πυκνού ινοαδενικού ιστού.
- C. Οι μαστοί είναι ετερογενώς πυκνοί, γεγονός που μπορεί να αποκρύψει μικρές μάζες.
- D. Οι μαστοί είναι εξαιρετικά πυκνοί, γεγονός που μειώνει την ευαισθησία της μαστογραφίας.

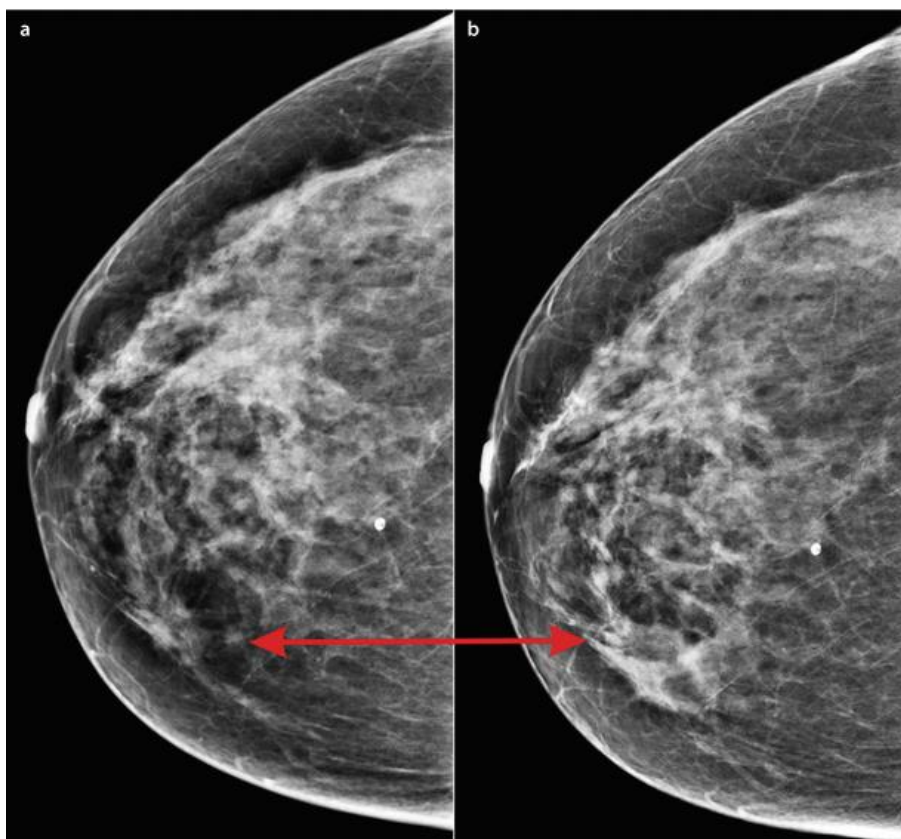


Εικόνα 11: Οι 4 κατηγορίες πυκνότητας των μαστών όπως απεικονίζονται στην εικόνα της Μαστογραφίας [97]

Περίπου το 10% των γυναικών έχουν εντελώς λιπώδεις μαστούς (κατηγορία A), το 40% των γυναικών έχουν διάσπαρτες περιοχές πυκνού ινοαδενικού ιστού (κατηγορία B), το 40% έχουν ετερογενώς πυκνούς μαστούς (κατηγορία C) και το 10% των μαστών είναι εξαιρετικά πυκνοί (κατηγορία D). Η πυκνότητα τείνει να μειώνεται με την ηλικία, επομένως η μαστογραφία είναι πιο ευαίσθητη σε γυναίκες άνω των 40 ετών, αν και υπάρχουν γυναίκες με πυκνούς μαστούς ακόμη και σε προχωρημένη ηλικία καθώς και νεαρές γυναίκες με λιπώδεις μαστούς [97].

Η ολοκληρωμένη αξιολόγηση μιας μαστογραφικής μελέτης περιλαμβάνει την αξιολόγηση και των δύο όψεων και των δύο μαστών για την ανίχνευση πιθανών ανωμαλιών, ιδιαίτερη προσοχή επίσης δίνεται στη μασχάλη και το δέρμα. Οι δύο όψεις είναι ζωτικής σημασίας για τη διαφοροποίηση των καλοήθων αλλοιώσεων από τις πραγματικές βλάβες. Εάν μια ασύμμετρη πυκνότητα είναι ορατή μόνο στη μία προβολή, πιθανότατα προκαλείται από υπέρθεση φυσιολογικού ιστού του μαστού καθώς οι πραγματικές βλάβες εμφανίζονται συνήθως και στις δύο προβολές [97].

Η σύγκριση νέων και προηγούμενων μαστογραφιών είναι κρίσιμης σημασίας για την ανίχνευση παθολογικών ευρημάτων, ιδίως στο πλαίσιο του προσυμπτωματικού ελέγχου, όπου μια ανεπαίσθητη αλλαγή της εστιακής πυκνότητας του μαστού μπορεί να είναι το πρώτο και μοναδικό σημάδι κακοήθειας (βλ. Εικόνα 12). Στις καλοήθεις μεταβολές, η σύγκριση με προηγούμενες εικόνες επιβεβαιώνει τον σταθερό χαρακτήρα και το μέγεθος των μαζών, των παραμορφώσεων ή των μικροασβεστώσεων / μικροαποτιτανώσεων, ιδίως εάν είναι μη τυπικής καλοήθους μορφολογίας. Στις περιπτώσεις αυτές, η ασθενής μπορεί να αποφύγει τις περιττές πρόσθετες απεικονίσεις ή βιοψίες [97].



Εικόνα 12: Ανίχνευση κακοήθειας - Μια αναπτυσσόμενη πυκνότητα ως το μόνο σημάδι παθολογίας, ορατό μόνο κατά τη σύγκριση της τρέχουσας εικόνας b με την προηγούμενη εικόνα a

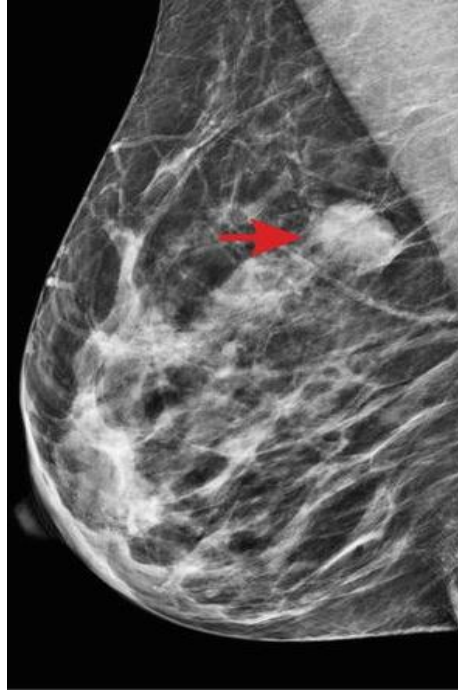
Οι παράγοντες που βελτιώνουν την απόδοση της μαστογραφίας είναι οι εξής [97]:

- ✚ Ψηφιακή μαστογραφία πλήρους πεδίου
- ✚ Σωστή τοποθέτηση του μαστού
- ✚ Χαμηλή πυκνότητα του μαστού
- ✚ Σύγκριση με προηγούμενες εικόνες

Τα μη φυσιολογικά ευρήματα στη μαστογραφία μπορεί να περιλαμβάνουν [97]:

- ✚ Μάζες (που χαρακτηρίζονται από το σχήμα, τα περιθώρια και την πυκνότητα)
- ✚ Αρχιτεκτονικές παραμορφώσεις στην αρχιτεκτονική του μαστού
- ✚ Ασυμμετρίες
- ✚ Μικροαποτιτανώσεις / Μικροασβεστώσεις (αξιολογούνται από τη μορφολογία και την κατανομή τους)

Οι τυπικές κακοήθεις βλάβες εμφανίζονται στη μαστογραφία ως ακανόνιστες μάζες με ακανόνιστα ή δυσδιάκριτα όρια και πυκνότητα ίση ή μεγαλύτερη από το παρέγχυμα του μαστού (Εικόνα 13). Η στρέβλωση του παρεγχύματος ή η υποχώρηση του μπορεί να αποτελεί ένδειξη κακοήθειας ακόμη και αν δεν είναι ορατή μια διακριτή μάζα. Μπορεί επίσης να παρατηρηθούν συνοδά χαρακτηριστικά, όπως αλλαγές στο δέρμα και τη θηλή (πάχυνση, σύσπαση), εισβολή στο θωρακικό τοίχωμα ή μασχαλιαία λεμφαδενοπάθεια [97].



Εικόνα 13: Τυπική εμφάνιση κακοήθους βλάβης με δυσδιάκριτα όρια και μεγαλύτερη πυκνότητα από το φυσιολογικό παρέγχυμα [97]

Οι καλοήθεις βλάβες είναι συνήθως ωοειδείς ή στρογγυλές με ομαλά περιθώρια και μεταβλητή πυκνότητα. Οι αλλοιώσεις πολύ χαμηλής πυκνότητας που περιέχουν λίπος είναι σχεδόν πάντα καλοήθεις. Η μαστογραφία από μόνη της δεν μπορεί να διακρίνει μεταξύ συμπαγών και κυστικών αλλοιώσεων, καθώς εμφανίζονται παρόμοιες, συνεπώς, συνήθως απαιτείται πρόσθετη αξιολόγηση με υπερηχογράφημα [97].

Οι καλοήθεις μικροασβεστώσεις στο μαστό αποτελούν συχνό εύρημα. Τα καλοήθη χαρακτηριστικά περιλαμβάνουν δερματική ή αγγειακή εντόπιση, στρογγυλό σχήμα, μεγάλη χονδροειδής μορφολογία που μοιάζει με «ποπ κορν» και είναι χαρακτηριστική για τα ασβεστοποιημένα ινοαδενώματα, μεγάλο σχήμα που μοιάζει με ράβδο και διάχυτο ομοιογενές μοτίβο. Η άμορφη δυστροφική ασβεστοποίηση μετά από τραύμα ή χειρουργική επέμβαση και οι αποτιτανώσεις ιζημάτων ασβεστίου σε μικροκυστίδια έχουν επίσης τυπική εμφάνιση και είναι καλοήθους προέλευσης [97].

Την εξέλιξη της ψηφιακής μαστογραφίας αποτελεί η ψηφιακή τομοσύνθεση μαστού (Digital breast tomosynthesis - DBT) που είναι μια σχετικά πρόσφατη μέθοδος απεικόνισης η οποία παρέχει τρισδιάστατη εικόνα του μαστού. Πολλαπλές εικόνες χαμηλής δόσης λαμβάνονται κατά τη διάρκεια της κίνησης του σωλήνα ακτίνων X σε τόξο γύρω από το μαστό. Μια τρισδιάστατη εικόνα του μαστού ανακατασκευάζεται από τις ληφθείσες εικόνες και προβάλλεται σε πολλαπλές παράλληλες τομές πάχους 1 mm. Αυτό επιτρέπει την καλύτερη απεικόνιση των παραμορφώσεων και των μαζών. Λόγω του αυξημένου αριθμού εικόνων, ο χρόνος ανάγνωσης για μια μελέτη DBT είναι μεγαλύτερος από ό,τι για τη μαστογραφία. Η δόση

ακτινοβολίας από τη DBT (όταν χρησιμοποιείται ως αυτόνομη μέθοδος) δεν είναι υψηλότερη από εκείνη της ψηφιακής μαστογραφίας [97].

Η ταξινόμηση των απεικονιστικών αποτελεσμάτων του μαστού γίνεται με το σύστημα κατηγοριοποίησης BI-RADS (Breast Imaging-Reporting Data System) [98]. Πρόκειται για ένα εργαλείο εκτίμησης κινδύνου και διασφάλισης ποιότητας των αποτελεσμάτων. Το σύστημα BI-RADS είναι, ουσιαστικά, μια κλίμακα που συνοψίζει τα συμπεράσματα της μαστογραφίας, του υπερηχογραφήματος μαστού ή της μαγνητικής μαστογραφίας. Διευκολύνει την παρακολούθηση των αποτελεσμάτων και μειώνει την σύγχυση στην ερμηνεία τους. Διακρίνεται σε επτά κατηγορίες, όπως παρουσιάζονται παρακάτω [98].

- ✚ BI-RADS 0: Ανεπαρκής εξέταση. Απαιτούνται επιπρόσθετες απεικονιστικές εκτιμήσεις και εξετάσεις ή προηγούμενες μαστογραφίες για σύγκριση.
- ✚ BI-RADS 1: Αρνητική. Συμμετρική εικόνα μαστών, χωρίς κάποια μάζα, δομική παραμόρφωση ή ύποπτες μικροαποτιτανώσεις.
- ✚ BI-RADS 2: Καλοήγη ευρήματα. Υπάρχει 0% πιθανότητα κάποιας κακοήθειας. Η αξιολόγηση είναι φυσιολογική και τα ευρήματα μπορεί να είναι κύστες, λιπώματα, καλοήθειας αποτιτανώσεις, τυπικά ινοαδενώματα κ.ά..
- ✚ BI-RADS 3: Πιθανώς καλοήγη ευρήματα. Υπάρχει θεωρητικά μια αλλοίωση μη ψηλαφητή, με μικρότερη του 2% πιθανότητα κακοήθειας. Συστήνεται επανέλεγχος σε 6 μήνες. Αν τα ευρήματα δεν δείξουν καμία αλλαγή με την πάροδο 24 μηνών, η τελική εκτίμηση αλλάζει σε BI-RADS 2 (καλοήθης) και δεν απαιτείται περαιτέρω παρακολούθηση. Εάν μια BI-RADS 3 αλλοίωση εμφανίσει οποιαδήποτε αλλαγή κατά τη διάρκεια της παρακολούθησης, θα αλλάξει σε μια BI-RADS 4 ή 5 και τότε τα κατάλληλα μέτρα πρέπει να ληφθούν.
- ✚ BI-RADS 4: Ευρήματα ύποπτα για κακοήθεια. Υπάρχει έως και 94% πιθανότητα καρκίνου. Η κατηγορία αυτή χωρίζεται σε 3 υποκατηγορίες, στην BI-RADS 4a (2 - 9%, μικρή πιθανότητα κακοήθειας), BI-RADS 4b (10 - 49%, μέτρια πιθανότητα κακοήθειας) και BI-RADS 4c (50 - 94%, μεγάλη πιθανότητα κακοήθειας). Είναι αναγκαία, τουλάχιστον, η επιβεβαίωση της ιστολογικής διάγνωσης, με βιοψία μαστού με κόπτουσα βελόνα.
- ✚ BI-RADS 5: Ευρήματα με μεγάλη υποδήλωση κακοήθειας. Η πιθανότητα κακοήθειας είναι πάνω από 95% και απαιτείται χειρουργική επέμβαση, καθώς και λήψη κατάλληλων μέτρων.

- ✚ BI-RADS 6: Γνωστή και αποδεδειγμένη με βιοψία κακοήθεια. Αφορά διαγνωσμένη κακοήθεια και χρησιμοποιείται κυρίως για ασθενείς, οι οποίες έχουν υποβληθεί σε προεγχειρητική χημειοθεραπεία.

Στις περιπτώσεις όπου υπάρχουν πολλαπλά ευρήματα, ως αποτέλεσμα της εξέτασης της μαστογραφίας επιλέγεται η μέγιστη BI-RADS κατηγορία.

3.1.1.2 Υπέρηχος μαστών

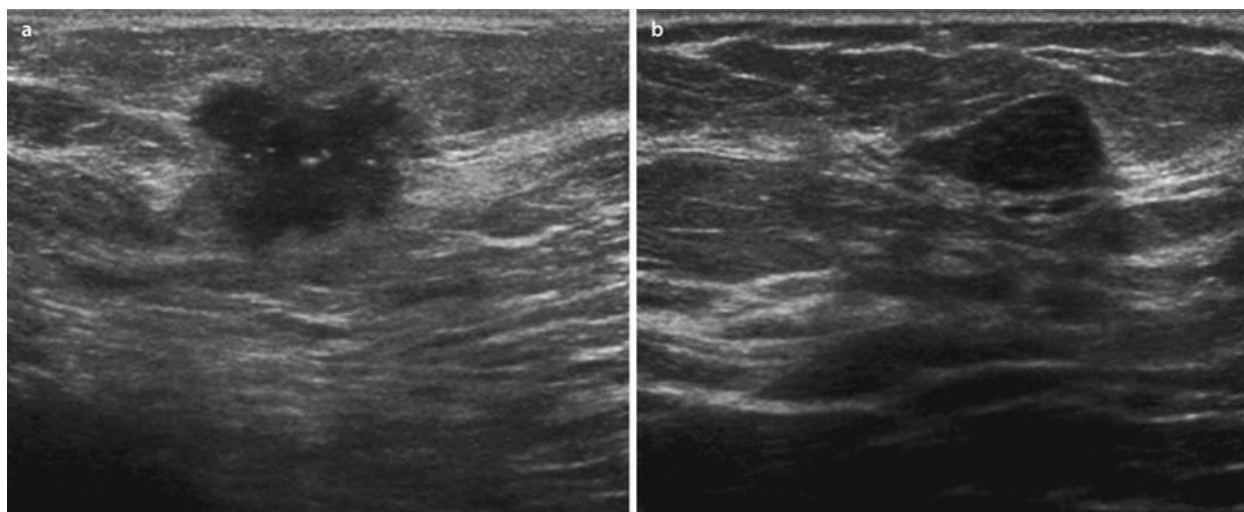
Ο υπέρηχος (ή υπερηχογράφημα) είναι μια μέθοδος που χρησιμοποιεί ακουστικά κύματα υψηλής συχνότητας και η εικόνα συντίθεται σε πραγματικό χρόνο με βάση τα διαφορετικά ανακλαστικά χαρακτηριστικά των ιστών. Για την απεικόνιση του μαστού συνιστώνται γραμμικοί μεταρροπείς με ελάχιστη συχνότητα 10 - 12 MHz. Η μέθοδος δεν ενέχει πρακτικά κανέναν κίνδυνο για τους ιστούς και μπορεί να χρησιμοποιηθεί με ασφάλεια σε όλες τις ομάδες ασθενών, επίσης κατά τη διάρκεια της εγκυμοσύνης και του θηλασμού [97].

Το υπερηχογράφημα του μαστού χρησιμοποιείται ως αυτόνομη διαγνωστική τεχνική ή σε συνδυασμό με άλλες απεικονιστικές μεθόδους. Οι ενδείξεις για υπερηχογράφημα μαστού περιλαμβάνουν [97]:

- ✚ Συμπληρωματική μέθοδος στη μαστογραφία προσυμπτωματικού ελέγχου
- ✚ Επιβεβαίωση και χαρακτηρισμός ανωμαλιών που διαπιστώνονται με άλλες μεθόδους
- ✚ Αξιολόγηση ψηλαφητών μαζών και άλλων συμπτωμάτων που σχετίζονται με τον μαστό
- ✚ Αξιοποίηση πληροφοριών για τη σταδιοποίηση σε γυναίκες με καρκίνο του μαστού
- ✚ Αξιολόγηση των εμφυτευμάτων μαστού
- ✚ Καθοδήγηση επεμβατικών διαδικασιών στο μαστό και τη μασχάλη
- ✚ Αξιολόγηση νεαρών, εγκύων και θηλαζουσών ασθενών με κλινικά συμπτώματα

Η ποιότητα της εξέτασης είναι μεταβλητή και επηρεάζεται από διάφορους παράγοντες όπως είναι οι τεχνικές παράμετροι του μηχανήματος, η ρύθμιση της εικόνας και η εμπειρία του χρήστη. Η ορατότητα της βλάβης στον υπέρηχο εξαρτάται από τη φύση και το μέγεθός της. Ορισμένες βλάβες μπορεί να έχουν πολύ ευδιάκριτη μορφολογία στον υπέρηχο, όπως οι κύστεις, και ορισμένες βλάβες μπορεί να έχουν πολύ ανεπαίσθητη εμφάνιση (ινοαδενώματα ή ορισμένες κακοήθεις βλάβες). Επίσης, η πλειονότητα των μικροασβεστώσεων δεν είναι καθόλου ορατές στο υπερηχογράφημα [97].

Η αξιοποίηση της εξέτασης του υπερήχου μαστών, σύμφωνα με το Αμερικανικό Κολλέγιο Ραδιολογίας, είναι απαραίτητη για το χαρακτηρισμό της ύποπτης μάζας χρησιμοποιώντας μορφολογικά χαρακτηριστικά όπως είναι το σχήμα, το περιθώριο και ο προσανατολισμός. Πολύ σημαντικό στοιχείο θεωρείται η επίδραση της βλάβης στον περιβάλλοντα ιστό, δηλαδή η ενδεχόμενη αρχιτεκτονική διαταραχή του μαστού. Οι κατηγορίες χαρακτηριστικών, όπως είναι τα χαρακτηριστικά αγγείωσης, σκληρότητας και η ηχογένεια, και οι τεχνικές, όπως το έγχρωμο ή δυναμικό Doppler και η ελαστογραφία, μπορεί να συνεισφέρουν πολύτιμες πληροφορίες για την εξαγωγή των κατάλληλων συμπερασμάτων ως προς την εξεταζόμενη μάζα. Με τη βοήθεια του Υπερήχου γίνεται αναγνώριση των ειδικών ευρημάτων όπως είναι οι απλές και πολλαπλές κύστεις, οι ομαδοποιημένες μικροκύστεις και οι ενδομαστικοί λεμφαδένες, γεγονός που απλοποιεί ακόμα περισσότερο την ερμηνεία της εκάστοτε περίπτωσης [98]. Στην Εικόνα 14 απεικονίζονται δύο στιγμιότυπα από την εξέταση του υπερήχου, με κακοήθεις και καλοήθεις αλλοιώσεις αντίστοιχα.



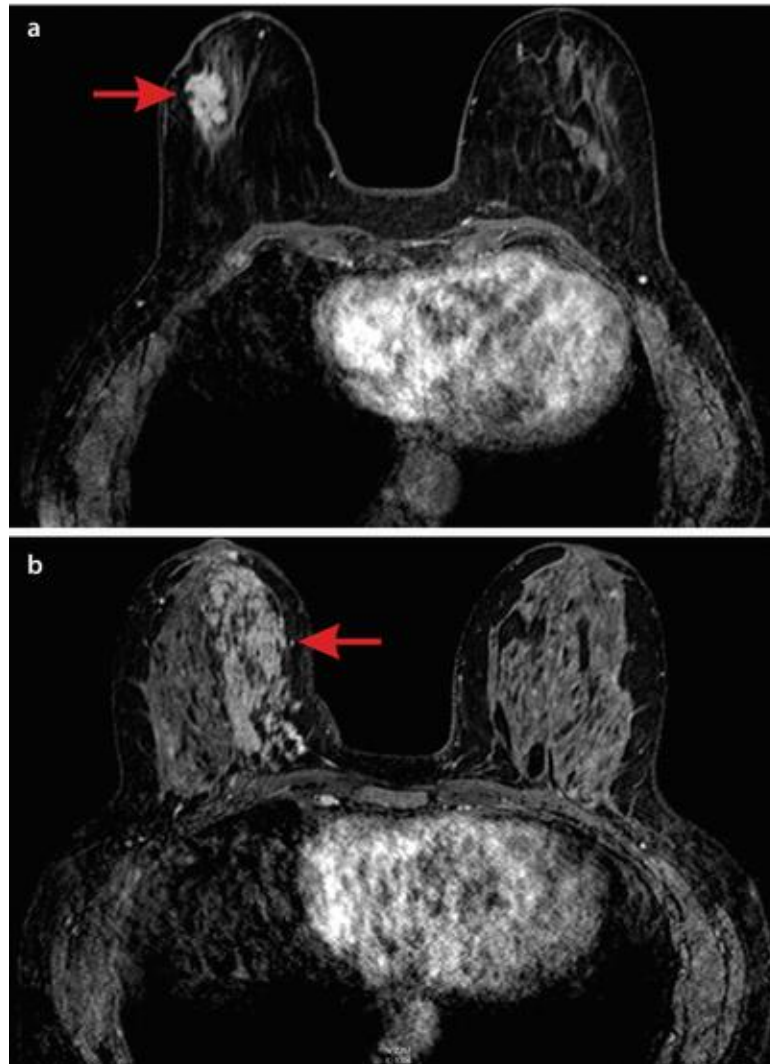
Εικόνα 14: Βλάβες μαστού στο υπερηχογράφημα – a) κακοήθης όγκος με δυσδιάκριτα όρια, έντονα υποηχογενής, με φωτεινές κουκκίδες που αντιπροσωπεύουν μικροασβεστώσεις- b) καλοήθης αλλοίωση - ινοαδένωμα - με ομαλά περιγράμματα, παράλληλο προσανατολισμό και ομογενή δομή [97]

Όσον αφορά σε σημαντικά, δηλαδή «ύποπτα» ευρήματα, το μέγεθος της βλάβης θα πρέπει να δίνεται σε δύο τουλάχιστον διαστάσεις, ιδανικά σε τρεις, ειδικά αν ο όγκος της βλάβης πρόκειται να συγκριθεί με μία ή περισσότερες προηγούμενες εξετάσεις. Δεν είναι απαραίτητο να αναφέρονται οι διαστάσεις των μικρών απλών κύστεων. Στην περίπτωση πολλαπλών κύστεων, και στους δύο μαστούς, αρκεί η πληροφορία της θέσης και των διαστάσεων της μεγαλύτερης κύστης στον εκάστοτε μαστό προκειμένου να εξαχθούν ασφαλή συμπεράσματα από τους ειδικούς [98].

3.1.1.3 Μαγνητική Τομογραφία Μαστών

Η απεικόνιση του μαστού με μαγνητικό συντονισμό, η Μαγνητική Τομογραφία Μαστών (Magnetic Resonance Imaging - MRI), χρησιμοποιείται όλο και περισσότερο για τον χαρακτηρισμό και τη διάγνωση

της παθολογίας του μαστού. Η μέθοδος χαρακτηρίζει τους ιστούς με βάση τη μορφολογία και τη σύνθεσή τους, τα οποία οδηγούν σε μεταβλητά σήματα στις διάφορες ακολουθίες μαγνητικής τομογραφίας. Η MRI χρησιμοποιεί την εφαρμογή σκιαγραφικών μέσων για την ανίχνευση της παθολογίας και αξιολογεί τα βιολογικά χαρακτηριστικά των βλαβών. Η ερμηνεία είναι πολύπλοκη και απαιτεί δεξιότητα, εμπειρία και χρόνο [97]. Στην Εικόνα 15 απεικονίζεται ο καρκίνος του μαστού στη μαγνητική τομογραφία όπου οι κακοήθεις βλάβες (βέλη) εμφανίζονται ως μαζική βλάβη (α) και ως εκτεταμένη μη μαζική ενίσχυση (b).



Εικόνα 15: Καρκίνος του μαστού στη μαγνητική τομογραφία. Κακοήθεις βλάβες (βέλη) που εμφανίζονται ως μαζική βλάβη (α) και εκτεταμένη μη μαζική ενίσχυση (Non-Mass-Like Enhancement - NMLE) (b) [97]

Η MRI δρα συμπληρωματικά με τις προαναφερθείσες διαγνωστικές εξετάσεις και αυξάνει τη διαγνωστική απόδοση στις περιπτώσεις των πυκνών μαστών. Η χρήση της MRI στη διαδικασία διάγνωσης καρκίνου του μαστού έχει αυξηθεί σημαντικά τις τελευταίες δύο δεκαετίες. Οι ενδείξεις για τη χρήση της MRI περιλαμβάνουν τόσο τη διαλογή όσο και τη διάγνωση όπως φαίνεται στις παρακάτω περιπτώσεις [97]:

- ✦ Προληπτικός έλεγχος γυναικών με υψηλό κίνδυνο εμφάνισης καρκίνου του μαστού κατά τη διάρκεια της ζωής τους.
- ✦ Προεγχειρητική σταδιοποίηση - αξιολόγηση της έκτασης της νόσου και ανίχνευση πιθανής πολυεστιακής ή πολυκεντρικής ή αντιπλευρικής νόσου σε ασθενείς με αποδεδειγμένο καρκίνο του μαστού.
- ✦ Προθεραπευτική αξιολόγηση της τοπικής έκτασης της νόσου - αξιολόγηση της πιθανής συμμετοχής του θωρακικού τοιχώματος, των αγγείων ή των νευρικών δεσμίδων για την εκτίμηση της χειρουργικής δυνατότητας.
- ✦ Παρακολούθηση της πρωτογενούς συστηματικής θεραπείας - η μαγνητική τομογραφία μπορεί να αξιολογήσει τις πρώιμες αλλαγές στον όγκο που προκαλούνται από τη χημειοθεραπεία για να καταδειχθεί η επίδραση της θεραπείας. Είναι επίσης πολύτιμη κατά την έναρξη της θεραπείας για τον καθορισμό του προτύπου της νόσου στο μαστό.
- ✦ Εκτίμηση της υπολειπόμενης νόσου μετά την πρωτογενή συστηματική θεραπεία.
- ✦ Ασαφή ευρήματα από τις συμβατικές μεθόδους - διφορούμενα ευρήματα στη μαστογραφία ή στο υπερηχογράφημα, ασυμφωνία με τα κλινικά ευρήματα ή διαφοροποίηση της ουλής από την υποτροπή - η μαγνητική τομογραφία συνιστάται μόνο σε περιπτώσεις που δεν είναι δυνατή η βιοψία.
- ✦ Ανίχνευση κρυφού πρωτοπαθούς όγκου σε ασθενείς με μεταστατική προσβολή μασχαλιαίων λεμφαδένων και αρνητική μαστογραφία και υπερηχογράφημα.
- ✦ Αξιολόγηση των εμφυτευμάτων μαστού - χρησιμοποιείται ειδικό πρωτόκολλο για τη σιλικόνη και το υγρό.

3.1.2 Βιοψία Μαστού

Στις περιπτώσεις που ο ειδικός ιατρός θεωρήσει ύποπτα τα ευρήματα των ανωτέρω εξετάσεων διάγνωσης του καρκίνου του μαστού, παραπέμπει τη γυναίκα σε βιοψία. Η βιοψία ανήκει στις επεμβατικές μεθόδους διάγνωσης και πρόκειται για τη λήψη ενός μικρού δείγματος ιστού ή υγρού από έναν ύποπτο όγκο του μαστού. Ο ιστός που λαμβάνεται ελέγχεται στη συνέχεια στο μικροσκόπιο, ώστε να εκτιμηθεί το είδος των κυττάρων του. Από τα αποτελέσματα, δηλαδή την ύπαρξη κυτταρικών ανωμαλιών ή άλλων ασυνήθιστων αλλαγών, διευκολύνεται ο σχεδιασμός για τα επόμενα διαγνωστικά και θεραπευτικά βήματα και σε κάποιες περιπτώσεις προκύπτει η τελική διάγνωση. Ανάλογα με τα χαρακτηριστικά της ύποπτης ανωμαλίας, προτείνονται διάφορες, προεγχειρητικές διαδικασίες λήψης ιστού για βιοψία μαστού, με συχνότερες τις παρακάτω.

Βιοψία με λεπτή βελόνα αναρρόφησης (Fine Needle Aspiration - FNA)

Η αναρρόφηση με λεπτή βελόνα είναι ένας τύπος βιοψίας που πραγματοποιείται με μια μικρή βελόνα για τη λήψη δειγμάτων ιστού και υγρού από συμπαγείς ή κυστικές βλάβες του μαστού. Τα κύτταρα που λαμβάνονται τοποθετούνται σε ένα γυάλινο πλακάκι και εξετάζονται στο μικροσκόπιο από ειδικό κυτταρολόγο. Η FNA αποτελεί μία από τις πολλές διαφορετικές μεθόδους διάγνωσης των μαζών του μαστού εκτός της επίσημης εκτομής [6]. Η FNA προτιμάται ιδιαίτερα όταν πρέπει να εκτιμηθεί η διήθηση των λεμφαδένων. Παρόλα αυτά λόγω της μικρής ποσότητας υλικού δεν είναι δυνατή η πλήρης ιστολογική ανάλυση, αλλά μόνο η κυτταρολογική [99]. Η FNA ενδείκνυται για ασθενείς με μαστογραφική ανωμαλία ή ψηλαφητές αλλοιώσεις του μαστού. Η FNA παραμένει αμφιλεγόμενη για τη διάγνωση ύποπτων κακοήθων αλλοιώσεων λόγω της σχετικά μικρής ποσότητας ιστού του μαστού που λαμβάνεται και του υψηλού ποσοστού μη διαγνωστικών ή ανεπαρκών δειγμάτων. Σε αυτή την περίπτωση, διενεργείται η λεγόμενη «τριπλή δοκιμασία» με συνδυασμό της φυσικής εξέτασης, των απεικονιστικών μελετών και της κυτταρολογίας για να τεθεί αυτή η διάγνωση ώστε να αυξηθεί η διαγνωστική ακρίβεια. Η FNA και η βιοψία με κόπτουσα βελόνα (Core Needle Biopsy – CNB) μπορούν να πραγματοποιηθούν με τη βοήθεια απεικονιστικής καθοδήγησης (συγκεκριμένα υπερήχων) ή να γίνουν «με ελεύθερο χέρι» στην περίπτωση ψηλαφητών αλλοιώσεων του μαστού. Η υποβοηθούμενη με κενό βιοψία εκτελείται με τη βοήθεια υπερηχογραφήματος, μαστογραφίας ή μαγνητικής τομογραφίας. Υπάρχει διχογνωμία σχετικά με το ποια είναι η προτιμώμενη μέθοδος για τη διάγνωση του καρκίνου του μαστού, καθώς η καθεμία έχει πλεονεκτήματα και αδυναμίες και η ενσωμάτωσή τους στην κλινική πρακτική ποικίλλει ανάλογα με τον εκάστοτε χειρουργό μαστού [6].

Σύμφωνα με τις πληροφορίες της ιατρικής κοινότητας η εν λόγω εξέταση παρά το γεγονός ότι είναι σχετικά οικονομικά αποδοτική, ενέχει ορισμένους κινδύνους που αποτρέπουν την χρήση της ως αναπόσπαστου μέρους της διαγνωστικής διαδικασίας, όπως είναι η μαστογραφία και ο υπέρηχος. Συγκεκριμένα, παρόλο που η FNA φαίνεται να είναι μια απλή διαδικασία, μπορεί να εμφανιστούν διάφορες επιπλοκές που περιλαμβάνουν τα ακόλουθα [6]:

- ✚ Μώλωπες
- ✚ Αιμάτωμα
- ✚ Μόλυνση
- ✚ Πνευμοθώρακα εάν η βελόνα προωθηθεί βαθιά στην περιοχή του θώρακα
- ✚ Πόνος

Επιπλέον, σημειώνεται η ανάγκη ύπαρξης έμπειρου κυτταροπαθολόγου για τη διάγνωση των αποτελεσμάτων [6].

Τα πλεονεκτήματα της FNA βιοψίας έναντι της βιοψίας με κόπτουσα βελόνα (CNB - Core Needle biopsy) είναι τα εξής:

- ✚ Αποδοτική ως προς το κόστος (η αποδοτικότητα μειώνεται όταν υπάρχουν ασαφή αποτελέσματα/ανάγκη για επαναληπτικές βιοψίες)
- ✚ Λιγότερο επεμβατική/λιγότερο επώδυνη λόγω του μικρότερου μεγέθους της βελόνας
- ✚ Μπορεί να εκτελεστεί στην κλινική ή στο γραφείο
- ✚ Κατάλληλη για ασθενείς που υποβάλλονται σε αντιπηκτική αγωγή με χαμηλό κίνδυνο αιματώματος/αιμορραγίας
- ✚ Μπορεί να χρησιμοποιηθεί για δειγματοληψία μασχαλιαίου ιστού

Τα μειονεκτήματα της FNA βιοψίας έναντι της βιοψίας με κόπτουσα βελόνα είναι τα εξής:

- ✚ Χαμηλότερη διαγνωστική απόδοση σε σχέση με τη βιοψία με κόπτουσα βελόνα
- ✚ 4% έως 13% των δειγμάτων που λαμβάνονται κρίνονται ανεπαρκή ή μη διαγνωστικά
- ✚ Πιο δύσκολη η διάκριση μεταξύ διηθητικού και in situ καρκίνου
- ✚ Μπορεί να απαιτείται πιο έμπειρος κυτταροπαθολόγος για τη διάγνωση
- ✚ Τα μη διαγνωστικά αποτελέσματα καθιστούν αναγκαία την επανάληψη της βιοψίας (συνήθως με CNB)

Παρά τα πλεονεκτήματα που εμφανίζει η μέθοδος βιοψίας FNA, καθώς είναι οικονομικά αποδοτική, ελάχιστα επεμβατική και μπορεί να χρησιμοποιηθεί για τη δειγματοληψία του μασχαλιαίου ιστού, επιλέξαμε να μη τη συμπεριλάβουμε στην εν λόγω μελέτη λόγω των ανωτέρω μειονεκτημάτων της ως προς τη διαγνωστική διαδικασία.

Βιοψία Μαστού με κόπτουσα βελόνα (Core Needle Biopsy -CNB):

Η μέθοδος αυτή είναι μια πολύ διαδεδομένη τεχνική βιοψίας μαστού. Στην CNB χρησιμοποιείται μεγαλύτερη βελόνα, σε σχέση με την FNA, με αποτέλεσμα να λαμβάνονται μεγαλύτερα δείγματα ιστού. Πιο συγκεκριμένα, η βελόνα αυτή εισάγεται στο κέντρο της ύποπτης περιοχής, αφαιρώντας πολύ μικρά τμήματα ιστού από το κέντρο της βλάβης, γι' αυτό και η διαδικασία αυτή λέγεται βιοψία του πυρήνα. Η εισαγωγή της βελόνας γίνεται με καθοδήγηση κάποιου απεικονιστικού μέσου. Με αυτόν τον τρόπο είναι δυνατή μια λεπτομερής ιστολογική διάγνωση, που συμπεριλαμβάνει βιολογικά χαρακτηριστικά του όγκου και δείκτες πρόβλεψης, όπως η ορμονοεξάρτηση και ο δείκτης πολλαπλασιασμού Ki67. Σαν τεχνική είναι πιο χρονοβόρα και επίπονη, σε σχέση με την FNA, ενώ απαιτείται τοπική αναισθησία. Η CNB έχει υψηλότερα ποσοστά ευαισθησίας και ειδικότητας από την FNA. Τα πλεονεκτήματα της βιοψίας αυτής

είναι η αποφυγή ενός αχρειαστού χειρουργείου και ολικής αναισθησίας, εφόσον προκύψει καλοήθης ιστολογική διάγνωση [100].

Διαδερμικές στερεοτακτικές τεχνικές

Για κάθε ύποπτη βλάβη που εντοπίζεται στη μαστογραφία ή την τομοσύνθεση και είναι ορατή με υπερήχους συνήθως συνίσταται βιοψία υπό υπερηχογραφική καθοδήγηση. Η βιοψία υπό υπερηχογραφική καθοδήγηση είναι γενικά πιο άνετη για την ασθενή (η οποία βρίσκεται σε ύπτια θέση), δεν εκθέτει την ασθενή ή το μαστό σε πρόσθετη ακτινοβολία, είναι λιγότερο δαπανηρή και είναι συχνά πιο εύκολα προσβάσιμη. Οι μαστογραφικά εντοπισμένες βλάβες που δεν είναι ορατές με υπερήχους, συνηθέστερα οι αποτιτανώσεις και σπανιότερα οι μάζες, ασυμμετρίες και αρχιτεκτονικές παραμορφώσεις, υποβάλλονται σε στερεοτακτική βιοψία με αναρρόφηση (Mammotome). Οι αλλοιώσεις που απεικονίζονται μόνο ή καλύτερα στην τομοσύνθεση μπορούν να υποβληθούν σε στερεοτακτική βιοψία υπό την καθοδήγηση της τομοσύνθεσης [101].

Οι στερεοτακτικές βιοψίες μαστού έχουν καταστεί απαραίτητες και αποτελούν κατάλληλη πρακτική για τις ασθενείς στις οποίες η μαστογραφία ή η τομοσύνθεση αποκαλύπτει αλλοιώσεις του μαστού που υποδηλώνουν κακοήθεια. Διατίθενται πλέον διάφορα συστήματα στερεοτακτικής βιοψίας και τύποι βελόνων, τα οποία επιτρέπουν ακριβέστερη δειγματοληψία των βλαβών καθώς και επιτυχή βιοψία βλαβών σε δύσκολες θέσεις σε ασθενείς κάθε σωματότυπου [101].

Η χρήση βελόνας βιοψίας με υποβοήθηση κενού (Vacuum-Assisted Biopsy - VAB) αποτελεί πλέον την πρότυπη βελόνα για τη στερεοτακτική βιοψία. Η συσκευή βιοψίας με υποβοήθηση κενού (Vacuum-Assisted Biopsy Device - VABD) εισήχθη αρχικά από τους Burbank και Parker ως διαγνωστικό εργαλείο, αλλά αργότερα ο ρόλος της VABD επεκτάθηκε σε δυνητική θεραπευτική αντιμετώπιση μικρών μαστογραφικών βλαβών και καλοήθων παθήσεων του μαστού [102]. Η πρώτη τεκμηριωμένη χρήση της θεραπευτικής χειρουργικής επέμβασης μαστού με VABD υπό την καθοδήγηση υπερήχων αποδίδεται στους Zannis et al. το 1998 [103]. Η εν λόγω βιοψία μπορεί να γίνει με διάφορες μορφές εικόνας αν και οι δύο συνηθέστερες είναι το υπερηχογράφημα και η μαστογραφία. Απαιτείται από τον χειριστή να είναι εξοικειωμένος με την υπερηχογραφική/στερεοτακτική απεικόνιση του μαστού σε πραγματικό χρόνο.

Οι VAB βελόνες τροφοδοτούνται με αναρρόφηση και διαθέτουν περιστρεφόμενο κόφτη. Μπορούν να λάβουν πολλαπλά δείγματα χωρίς την ανάγκη αφαίρεσης της βελόνας από τη βλάβη. Συγκεκριμένα, στη στερεοτακτική βιοψία, μέσω ενός καθετήρα, με τη χρήση κενού, γίνεται αναρρόφηση του ιστού του μαστού και τοποθετείται στο θάλαμο δειγματοληψίας της συσκευής. Έχει αποδειχθεί ότι οι βελόνες VAB είναι πιο ακριβείς και επιτρέπουν τη δειγματοληψία ιστού μεγαλύτερου όγκου σε σχέση με τις βελόνες αυτοματοποιημένης βιοψίας. Επιπλέον, η VAB προσφέρει υψηλότερο ποσοστό ανάκτησης των

αποτιτανώσεων και χαμηλότερο ποσοστό σφαλμάτων με αποτέλεσμα μειωμένα ποσοστά επαναληπτικής βιοψίας. Μετά τη στερεοτακτική βιοψία, είναι επιτακτική ανάγκη να συγκριθούν τα αποτελέσματα της απεικόνισης και της παθολογικής ανάλυσης προκειμένου να υπάρξει ακτινολογική και παθολογική συμφωνία και να προσδιοριστεί κατά πόσον είναι απαραίτητη η χειρουργική εκτομή ή κάποια άλλη πρόσθετη παρέμβαση [101].

Μια από τις σημαντικότερες εξελίξεις των τελευταίων ετών είναι η εισαγωγή της διαδερμικής βιοψίας του μαστού στη διαγνωστική διαδικασία. Οι διαδερμικές βιοψίες μαστού επιτρέπουν την ιστολογική μελέτη μιας βλάβης με χαμηλότερο κόστος και λιγότερη νοσηρότητα από τη χειρουργική βιοψία. Διάφορα συστήματα, μετρητές και μέθοδοι προσέγγισης επιτρέπουν στους χειρουργούς να προβούν σε βιοψία σχεδόν οποιουδήποτε τύπου βλάβης στο μαστό. Οι επεμβατικές διαδικασίες στο μαστό δεν περιορίζονται στη λήψη υλικού για ιστολογική διάγνωση ή τοποθέτηση δεικτών για τη χειρουργική καθοδήγηση: η διαδερμική εκτομή ή εκτομή με ραδιοσυχνότητα είναι πολλά υποσχόμενες τεχνικές που βρίσκονται σε στάδιο ανάπτυξης επί του παρόντος [104].

3.1.3 Γενετική Εξέταση

Όσον αφορά στα γενετικά τεστ για την ανίχνευση μεταλλάξεων που ενοχοποιούνται για τον καρκίνο του μαστού, λόγω του μικρού ποσοστού γυναικών (5 - 10 %) που φαίνεται να φέρουν μεταλλάξεις στο σύνολο των νοσούντων από καρκίνο του μαστού αλλά και του μεγάλου κόστους της εν λόγω εξέτασης, η καθολική εφαρμογή της γενετικής εξέτασης αποτελεί μια αναποτελεσματική στρατηγική για τα συστήματα υγείας [5]. Ο γενετικός έλεγχος θα πρέπει να είναι διαθέσιμος σε όλες τις ασθενείς με προσωπικό ιστορικό καρκίνου του μαστού. Ο έλεγχος αυτός θα πρέπει να περιλαμβάνει τα BRCA1/BRCA2 και PALB2, αλλά και συνδυασμό τους με άλλα γονίδια ανάλογα με το κλινικό σενάριο και το οικογενειακό ιστορικό. Για τις ασθενείς με νεοδιαγνωσθέντα καρκίνο του μαστού, ο εντοπισμός μιας μετάλλαξης μπορεί να επηρεάσει την τοπική θεραπεία (χειρουργική επέμβαση και ενδεχομένως ακτινοβολία) και τη συστηματική θεραπεία. Επιπλέον, στις περιπτώσεις που η γενετική εξέταση επιφέρει θετικά αποτελέσματα και εντοπιστούν μεταλλάξεις στα γονίδια που ενοχοποιούνται για τον καρκίνο του μαστού, τα μέλη της οικογένειας, που θα κρίνουν οι ειδικοί, θα πρέπει να υποβληθούν σε κατάλληλες αντίστοιχες εξετάσεις [105]. Λόγω του ότι υπάρχει μεγάλη ποικιλία διαθέσιμων πάνελ, με διαφορετικά γονίδια σε διαφορετικούς πίνακες, παρατηρείται έλλειψη συνεννόησης μεταξύ των ειδικών σχετικά με το ποια γονίδια πρέπει να ελέγχονται στα διάφορα κλινικά σενάρια. Υπάρχει επίσης, διακύμανση στις απόψεις όσον αφορά την κατανόηση του κινδύνου και την κατάλληλη κλινική διαχείριση των μεταλλάξεων σε ορισμένα γονίδια από μέρους των ειδικών [105].

Κάθε ασθενής που εξετάζεται από χειρουργό μαστού, η οποία υποβλήθηκε σε γενετικό έλεγχο στο παρελθόν και δεν εντοπίστηκε καμία παθογόνος μετάλλαξη, θα πρέπει να επανεκτιμάται και να εξετάζεται το ενδεχόμενο επικαιροποίησης του ελέγχου. Ειδικότερα, για μια ασθενή που είχε αρνητική εξέταση για μετάλλαξη των γονιδίων BRCA1 και 2, η οποία προέρχεται από οικογένεια χωρίς παθογόνες μεταλλάξεις, θα πρέπει να εξετάζεται το ενδεχόμενο συμπληρωματικής εξέτασης. Αυτό επισημαίνεται καθώς οι γενετικές εξετάσεις που διενεργήθηκαν πριν από το 2014 πιθανότατα δεν θα είχαν συμπεριλάβει το PALB2 ή άλλα δυνητικά συναφή γονίδια και μπορεί να μην είχαν συμπεριλάβει επίσης έλεγχο για μεγάλες γονιδιωματικές αναδιατάξεις σε BRCA1 ή BRCA2. Σύμφωνα με το NCCN (National Comprehensive Cancer Network) της Αμερικής, ο γενετικός έλεγχος θα πρέπει να είναι διαθέσιμος ακόμα και σε ασθενείς χωρίς ιστορικό καρκίνου του μαστού οι οποίες πληρούν τις κατευθυντήριες γραμμές της NCCN. Οι γυναίκες που δεν έχουν νοσήσει από καρκίνο του μαστού θα πρέπει να ενημερώνονται ότι η εξέταση μίας προσβεβλημένης με την ασθένεια συγγενούς τους πρώτα, όποτε είναι δυνατόν, είναι πιο κατατοπιστική από την υποβολή των ίδιων σε γενετική εξέταση [105]. Όταν δεν είναι εφικτό να εξεταστεί πρώτα ο προσβεβλημένος συγγενής, τότε θα πρέπει να εξετάζεται το μη προσβεβλημένο μέλος της οικογένειας, εάν ενδιαφέρεται, αλλά θα πρέπει να ακολουθείται προσεκτική συμβουλευτική διαδικασία πριν από την εξέταση, ώστε να εξηγηθεί η περιορισμένη αξία των «μη κατατοπιστικών αρνητικών» αποτελεσμάτων. Είναι επίσης λογικό να γίνει εξέταση σε πολυγονιδιακό πάνελ εάν το οικογενειακό ιστορικό είναι ελλιπές (π.χ. περίπτωση υιοθεσίας, ο ασθενής δεν είναι βέβαιος για τον ακριβή τύπο καρκίνου που προσβάλλει μέλη της οικογένειας) ή στο οικογενειακό ιστορικό εντοπίζονται άλλοι καρκίνοι. Οι μεταλλάξεις αβέβαιης σημασίας είναι αλληλουχίες DNA που δεν είναι κλινικά αξιοποιήσιμες. Αυτός ο τύπος αποτελέσματος πρέπει να θεωρείται ασαφής και η αντιμετώπιση της ασθενούς θα πρέπει να γίνει με βάση τους παράγοντες κινδύνου της και όχι να επηρεάζεται από αυτό το αποτέλεσμα [105].

3.1.4 Κλινική Εξέταση και Αυτοεξέταση Μαστού

Σύμφωνα με τις κατευθυντήριες γραμμές του NCCN, το σημείο εκκίνησης για τον έλεγχο και την αξιολόγηση των ανωμαλιών του μαστού είναι η καταγραφή του πλήρους ιατρικού ιστορικού, ακολουθούμενο από κλινική εξέταση του μαστού (Clinical Breast Examination - CBE). Οι μαστοί πρέπει να επιθεωρούνται, από τους ειδικούς, με τις ασθενείς σε όρθια και ύπτια θέση. Ασθενείς μπορούν επίσης να τοποθετηθούν σε οποιαδήποτε θέση χρειάζεται ώστε να αναδειχθεί οποιοδήποτε νέο σχήμα ή αλλαγές στο περίγραμμα του μαστού. Η CBE πρέπει να περιλαμβάνει την ψηλάφηση ολόκληρου του μαστού, της μασχαλιαίας περιοχής καθώς και του λαιμού για την ύπαρξη λεμφαδένων. Συμπτώματα ή θετικά ευρήματα στη CBE μπορεί να περιλαμβάνουν ένα ψηλαφητό εξόγκωμα ή μάζα, ασύμμετρη πάχυνση, εκκρίσεις στη θηλή σε περίπτωση απουσίας ψηλαφητής μάζας και δερματικές αλλαγές όπως είναι το πορτοκαλί χρώμα στο δέρμα, το ερύθημα, το εξάνθημα της θηλής και το έκζεμα [106].

Όσον αφορά στην αυτοεξέταση του μαστού μαστού (Breast Self-Examination - BSE), αν και η μηνιαία αυτοεξέταση προτεινόταν για πολλά χρόνια, πρόσφατα αναγνωρίστηκε ότι έχει περιορισμένη αξία στην ανίχνευση του πρώιμου καρκίνου του μαστού. Μεγάλες τυχαιοποιημένες κλινικές δοκιμές απέτυχαν να αποδείξουν ότι η BSE οδήγησε σε μείωση της θνησιμότητας λόγω καρκίνου του μαστού ή της συνολικής θνησιμότητας σε πληθυσμούς μέσου κινδύνου. Οργανισμοί όπως η Αμερικανική Αντικαρκινική Εταιρεία (American Cancer Society - ACS), το Αμερικανικό Κολέγιο Μαιευτικής και Γυναικολογίας και η Ομάδα Εργασίας Προληπτικών Υπηρεσιών των ΗΠΑ (US Preventative Service Task Force - USPSTF) δεν έχουν βρει επαρκή στοιχεία για να συστήσουν υπέρ ή κατά της διενέργειας BSE. Η έμφαση στη μηνιαία BSE έχει αντικατασταθεί από την έννοια της «αυτογνωσίας των μαστών», κατά την οποία οι γυναίκες αποκτούν γνώσεις σχετικά με τη φυσιολογική εμφάνιση και την αίσθηση των μαστών τους, χωρίς όμως συγκεκριμένο πρόγραμμα ή τεχνική εξέτασης. Ωστόσο, οι T. Bevers et al. [107], συμμαρίζονται την άποψη που εξακολουθούν να έχουν πολλοί ότι η BSE έχει αξία. Οι γιατροί θα πρέπει να εκπαιδεύουν τις γυναίκες ώστε να ενισχύουν την ευαισθητοποίησή τους και να αναφέρουν οποιοσδήποτε δερματικές αλλαγές, λακκούβες, εκκρίσεις από τη θηλή ή παρουσία νέων εξογκωμάτων εντοπίσουν. Ο περιορισμός της χρήσης της BSE ως εργαλείου διαλογής σχετίζεται με τον αυξημένο αριθμό βιοψιών που διενεργούνται για καλοήθεις παθήσεις του μαστού, οι οποίες έχουν οικονομικό και συναισθηματικό κόστος [107].

Οι γυναίκες που επιλέγουν να πραγματοποιήσουν BSE θα πρέπει να το κάνουν όταν οι μαστοί τους είναι λιγότερο πιθανό να είναι ευαίσθητοι ή πρησμένοι, συχνά λίγες ημέρες αμέσως μετά την έμμηνο ρύση. Η εξέταση είναι καλύτερο να γίνεται τόσο σε ύπτια όσο και σε όρθια θέση χρησιμοποιώντας τις άκρες των τριών μεσαίων δακτύλων. Για την εξέταση του μαστού χρησιμοποιούνται διαφορετικά επίπεδα πίεσης (ελαφριά, μεσαία, σταθερή). Αν και μπορούν να χρησιμοποιηθούν διάφορες τεχνικές, οι γυναίκες πρέπει να είναι συνεπείς στη χρησιμοποιούμενη τεχνική που χρησιμοποιούν. Μία από τις ευκολότερες τεχνικές περιλαμβάνει την ψηλάφηση των μαστών με δεξιόστροφο τρόπο και εναλλακτικά από τη θηλή προς τα έξω, μπορεί επίσης να χρησιμοποιηθεί ένα κάθετο μοτίβο (πάνω-κάτω) [107].

3.2 Ανασκόπηση και Σύγκριση της Απόδοσης των Διαγνωστικών Εξετάσεων για τον Καρκίνο του Μαστού

Μετά την περιγραφή των βασικότερων και πιο διαδεδομένων εξετάσεων προληπτικού ελέγχου και διάγνωσης για τον καρκίνο του μαστού, είναι απαραίτητη η εκτίμηση της διαγνωστικής απόδοσής τους. Αυτό σημαίνει ότι θα αναλυθεί ποιοτικά και ποσοτικά η χρησιμότητά τους στην έγκαιρη διάγνωση καρκίνου του μαστού και κατ' επέκταση στη μείωση θνητότητας εξαιτίας της ασθένειας, χάρη στην εκάστοτε τεχνική. Συγκεκριμένα, στη συνέχεια παρατίθενται τα αποτελέσματα ερευνών που μελετούν την απόδοση των διαθέσιμων διαγνωστικών μέσων που περιγράφηκαν στην υποενοότητα 3.1, όπου ελέγχονται

και συγκρίνονται ορισμένοι γνωστοί στατιστικοί δείκτες, όπως είναι η ακρίβεια (accuracy), η ευαισθησία (sensitivity) και η ειδικότητα (specificity).

3.2.1 Ποιοτική Εκτίμηση της Διαγνωστικής Απόδοσης των BSE και CBE

Η αξιολόγηση της BSE και της κλινικής εξέτασης CBE του μαστού ως μέσων διάγνωσης καρκίνου στον μαστό, κατά κύριο λόγο δεν περιλαμβάνει στατιστικά μέτρα απόδοσης. Αυτό οφείλεται κυρίως στη μικρή βιβλιογραφία που υπάρχει, με αποτέλεσμα να μην υπάρχει εκτενής και ποσοτική τεκμηρίωση των δεδομένων. Ωστόσο, έχουν προκύψει αρκετά συμπεράσματα, όσον αφορά την αξία τους ως επιπρόσθετων ή βοηθητικών τεχνικών ανίχνευσης όγκων στο μαστό [108].

Ορισμένες έρευνες που έχουν διεξαχθεί, έχουν διαπιστώσει, αρχικά, συχνότερη και ευκολότερη ανίχνευση μικρού όγκου, μέσω αυτών των πρακτικών, στις ηλικίες 35 - 39, σε σύγκριση με την ηλικιακή ομάδα 60 - 74. Όταν εφαρμόζεται τακτικά η μέθοδος BSE, οι γυναίκες είναι πιο πιθανό να ανακαλύψουν έναν όγκο μόνες τους. Το μέγεθος των όγκων που εντοπίζονται και με τις δύο πρακτικές, μάλιστα, μπορεί να είναι αρκετά μικρό, ώστε να θεωρηθεί πρώιμη ανίχνευση. Δηλαδή είναι καρκίνοι σε πρώιμο στάδιο σε σχέση με καρκίνους που διαγιγνώσκονται σε γυναίκες που δεν διενεργούν ούτε BSE ούτε CBE. Επιπλέον, η πλειονότητα των εν λόγω διαπιστώσεων συμβαίνει κατά τη διάρκεια καθημερινών δραστηριοτήτων, όπως είναι το μπάνιο [109]. Ταυτόχρονα, όμως, διαπιστώνεται ότι οι περισσότεροι όγκοι που εντοπίζονται αποδεικνύονται καλοήθεις. Η απευθείας διάγνωση κακοήθειας δεν φαίνεται να είναι εφικτή μόνο μέσω των μεθόδων αυτών. Η CBE έχει δείξει ασυνεπή στοιχεία σχετικά με την αποτελεσματική διάγνωση του καρκίνου του μαστού [110]. Παράγοντες που επηρεάζουν την ακρίβεια μιας κλινικής εξέτασης μαστών είναι το μέγεθος του όγκου, η ηλικία, η πυκνότητα των μαστών, αλλά φυσικά και η εμπειρία του ιατρού [111]. Η πραγματική αξία της CBE διαφαίνεται όταν συνοδεύει μία μαστογραφία, στην οποία λειτουργεί επικουρικά ανεβάζοντας τον δείκτη ευαισθησίας της διάγνωσης [93].

Οι δύο κυριότερες μελέτες που έγιναν με στόχο την σύνδεση των συχνών BSE και CBE με την θνητότητα, εξαιτίας του καρκίνου στον μαστό, ήταν ξεκάθαρες. Δεν υπάρχει καμία ένδειξη ότι οι τακτικές αυτές οδηγούν σε μείωση των θανάτων. Επιπλέον, οι δύο έρευνες, μία στην Ρωσία και μία στην Κίνα, έδειξαν ότι η συχνή αυτοεξέταση οδηγεί σε πολλαπλάσιες ιατρικές επισκέψεις, επιπλέον διαγνωστικές εξετάσεις, αύξηση του άγχους, υπερβολικό φόβο της ασθένειας και ψυχολογική πίεση, εξαιτίας συχνών ψευδώς θετικών ευρημάτων [111],[112].

Συμπερασματικά, λόγω του διχασμού στην επιστημονική κοινότητα, η σύσταση της BSE και της CBE ως διαγνωστικού εργαλείου, χωρίς συνοδεία απεικονιστικής εξέτασης, θεωρήθηκε αμφιλεγόμενη. Η συνεισφορά τους στην έγκαιρη διάγνωση δεν έχει αποδειχθεί με ξεκάθαρα στοιχεία και οι έρευνες είναι ελλιπείς. Αντίθετα, προτάθηκε η προώθηση μίας ευρύτερης νοοτροπίας εξοικείωσης των γυναικών με το

στήθος τους, ώστε οποιαδήποτε αλλαγή να γίνεται άμεσα αντιληπτή από τις ίδιες. Σε αντίθεση με την αυτοεξέταση, αυτή η προσέγγιση δεν περιλαμβάνει τις αυστηρές οδηγίες ενός κλινικού περιβάλλοντος, ελαττώνοντας την πρόκληση επιπρόσθετου άγχους [109].

Σύμφωνα με τους A. Miller et al. [113], η αποτελεσματικότητα του προσυμπτωματικού ελέγχου με CBE ή/και BSE εξετάζεται με βάση έμμεσα στοιχεία από τυχαίοποιημένες μελέτες προσυμπτωματικού ελέγχου του μαστού και από μελέτες παρατήρησης. Σε χώρες όπου ο καρκίνος του μαστού διαγιγνώσκεται σε προχωρημένο στάδιο, ο προσυμπτωματικός έλεγχος με CBE, και με τη διδασκαλία της BSE ως αναπόσπαστο στοιχείο, θα είναι πιθανώς αποτελεσματικός στη μείωση της θνητότητας από καρκίνο του μαστού. Ωστόσο, σε ιατρικά προηγμένες χώρες όπου παρέχεται επαρκής θεραπεία, καμία μέθοδος προσυμπτωματικού ελέγχου δεν είναι πιθανό να είναι επαρκώς επωφελής ώστε να αντισταθμίσει τις βλάβες που προκαλούν τα ψευδώς θετικά αποτελέσματα και η υπερδιάγνωση [113].

Σε κοινωνίες όπου η συμμετοχή στον προσυμπτωματικό έλεγχο για καρκίνου του μαστού και η ευαισθητοποίηση γύρω από τη νόσο είναι χαμηλή, η διάγνωση συνήθως γίνεται όταν ο καρκίνος βρίσκεται σε πιο προχωρημένα στάδια. Επίσης, για τις χώρες όπου ο καρκίνος του μαστού παρατηρείται σε μικρότερες ηλικίες, η CBE εξακολουθεί να διατηρεί τη σημασία της. Με την επαρκή ευαισθησία της, η CBE συγκαταλέγεται μεταξύ των σημαντικών διαγνωστικών μεθόδων που αποτρέπουν την παράβλεψη του καρκίνου του μαστού, ιδίως σε περιπτώσεις όπου οι πόροι υγείας είναι περιορισμένοι και η πρόσβαση στη μαστογραφία και το υπερηχογράφημα του μαστού είναι επίσης περιορισμένη [114].

3.2.2 Η Διαγνωστική Απόδοση της Μαστογραφίας

Σύμφωνα με την επιστημονική μελέτη των P. Steyerova et al. [97], η μαστογραφία παρουσιάζει υψηλότερη ευαισθησία για το ενδοτορικό καρκίνωμα (ductal carcinoma) (79,1 - 80%) από ό,τι για το λοβιακό καρκίνωμα (67,2 - 72,1%). Στους φορείς της μετάλλαξης BRCA, οι όγκοι εκδηλώνονται λιγότερο συχνά ως αγκαθωτές μάζες ή παραμορφώσεις και οι βλάβες είναι συχνά καλά περιγεγραμμένες και μιμούνται την καλοήγη νόσο, γεγονός που έχει ως αποτέλεσμα τη χαμηλότερη ευαισθησία της μαστογραφίας σε γυναίκες υψηλού κινδύνου. Ορισμένες κακοήθεις βλάβες εκδηλώνονται κυρίως ως μικροασβεστώσεις. Αυτές μπορεί να συνδέονται ή να μην συνδέονται με μάζες. Οι μικροασβεστώσεις υποδηλώνουν σε μεγάλο βαθμό κακοήθεια εάν είναι άμορφες, λεπτές πλειομορφικές, λεπτές γραμμικές με/χωρίς διακλαδώσεις ή χονδροειδείς και ετερογενείς [97].

Η πυκνότητα του μαστού εξαρτάται από την αναλογία του ινοαδενικού ιστού και του λιπώδους ιστού, με την έντονη ύπαρξη του πρώτου να σημαίνει μεγαλύτερη πυκνότητα. Με την αύξηση της ηλικίας μειώνεται και η πυκνότητα και έτσι η μαστογραφία παρουσιάζει μεγαλύτερη ευαισθησία, δηλαδή υψηλότερη

απόδοση, σε γυναίκες άνω των 40 ετών, χωρίς φυσικά αυτό να είναι απόλυτο [97]. Καταλήγουμε, με αυτόν τον τρόπο να παρατηρούμε ευαισθησία που μπορεί να αγγίζει ακόμα και το 98% σε μαστούς με αυξημένο λιπώδη ιστό (κατηγορίας ACR a). Αντίθετα, στις περιπτώσεις πυκνών μαστών, στο πλαίσιο του γενικού πληθυσμού, η ευαισθησία διατηρεί έναν χαμηλό μέσο όρο της τάξης του 30 - 40% [96].

Οι William E. Barlow et al. [115] αξιολόγησαν κατά πόσον η ηλικία, η πυκνότητα, ο όγκος του μαστού που αναφέρει η ίδια η ασθενής και η προηγούμενη μαστογραφία επηρεάζουν την απόδοση της διαγνωστικής μαστογραφίας. Από τον Ιανουάριο του 1996 έως τον Μάρτιο του 1998, συνέλεξαν προοπτικά δεδομένα διαγνωστικής μαστογραφίας από γυναίκες ηλικίας 25 - 89 ετών χωρίς προηγούμενο καρκίνο του μαστού και τα συνέδεσαν με δεδομένα αποτελεσμάτων του καρκίνου μαστού από έξι Μητρώα μαστογραφιών που συμμετείχαν στην Κοινοπραξία επιτήρησης του καρκίνου του μαστού (Breast Cancer Surveillance Consortium). Οι William E. Barlow et al. [115] χρησιμοποίησαν την τελική μαστογραφική αξιολόγηση στο τέλος της απεικονιστικής εξέτασης για να προσδιορίσουν το ποσοστό της μη φυσιολογικής μαστογραφικής εξέτασης, τη θετική προγνωστική αξία (Positive Predictive Values - PPV), την ευαισθησία, την ειδικότητα και την περιοχή κάτω από την καμπύλη ROC. Χρησιμοποίησαν την ηλικία, την πυκνότητα του μαστού, την προηγούμενη μαστογραφία και τον όγκο του μαστού που ανέφερε η ίδια η ασθενής, από κοινού ως παράγοντες πρόβλεψης της απόδοσης. Όλες οι στατιστικές δοκιμές ήταν αμφίπλευρες. Από τις 41.427 διαγνωστικές μαστογραφίες, οι 6.279 (15,2%) κρίθηκαν παθολογικές. Η συνολική PPV ήταν 21,8%, η ευαισθησία ήταν 85,8% και η ειδικότητα ήταν 87,7%. Η πολυπαραγοντική ανάλυση έδειξε ότι η ευαισθησία και η ειδικότητα γενικά μειώθηκαν καθώς αυξανόταν η πυκνότητα του μαστού, ότι το στοιχείο της προηγούμενης μαστογραφίας μείωσε την ευαισθησία (OR = 0,52, 95% CI = 0,36 - 0,74) αλλά αύξησε την ειδικότητα (OR = 1,43, 95% CI = 1,31 - 1,57) και ότι η ύπαρξη του αναφερόμενου από την ασθενή όγκου του μαστού αύξησε την ευαισθησία (OR = 1,64, 95% CI = 1,13 - 2,38) αλλά μείωσε την ειδικότητα (OR = 0,54, 95% CI = 0,49 - 0,59). Διαπιστώθηκε ότι η υψηλότερη πυκνότητα του μαστού και η προηγούμενη μαστογραφία σχετίζονταν αρνητικά με την ακρίβεια. Το συμπέρασμα της εν λόγω ερευνητικής ομάδας ήταν ότι η διαγνωστική μαστογραφία σε γυναίκες με σημάδια ή συμπτώματα καρκίνου του μαστού παρουσιάζει υψηλότερη ευαισθησία και χαμηλότερη ειδικότητα από ό,τι η μαστογραφία προσυμπτωματικού ελέγχου ενώ η υψηλότερη πυκνότητα του μαστού και η προηγούμενη μαστογραφία φαίνεται να μειώνουν την απόδοση [115].

Οι γυναίκες με οικογενειακό ιστορικό καρκίνου του μαστού έχουν υψηλότερη συχνότητα εμφάνισης της νόσου νωρίτερα στη ζωή τους, επομένως ο προσυμπτωματικός έλεγχος συστήνεται από την ηλικία των 35 ετών. Σε αυτές τις περιπτώσεις, συνιστώνται επίσης επιπρόσθετες απεικονιστικές μέθοδοι εκτός από τη μαστογραφία καθώς οι γυναίκες νεότερης ηλικίας έχουν συνήθως πυκνό ιστό μαστού όπου η μαστογραφία έχει χαμηλή ευαισθησία, επομένως τα παθολογικά χαρακτηριστικά μπορεί να αγνοηθούν. Επιπλέον, έχει

παρατηρηθεί ότι σε γυναίκες υψηλού κινδύνου φορείς μεταλλάξεων, όπως στο γονίδιο BRCA1, τα απεικονιστικά χαρακτηριστικά των του καρκίνου του μαστού είναι συνήθως λιγότερο συγκεκριμένα, καθώς τείνουν να μοιάζουν με καλοήθεις βλάβες και η μαστογραφία μπορεί να έχει περιορισμένη αξία.

Με ενδιαφέρον παρατηρείται ότι τα συμπεράσματα μελετών, που αφορούν γυναίκες υψηλής επικινδυνότητας, δεν ακολουθούν το ίδιο μοτίβο. Στις περιπτώσεις αποκλειστικής εφαρμογής της μαστογραφίας παρατηρείται πολύ χαμηλή ευαισθησία στο νεαρό ηλικιακό γκρουπ των 25 - 30 χρόνων, με υψηλό κίνδυνο εμφάνισης καρκίνου του μαστού, με τις τιμές της μαστογραφίας να κυμαίνονται από 14% έως 59%. Αυτό οδηγεί σε μη έγκαιρη διάγνωση των 2/3 των καρκίνων. Η χαμηλή αυτή απόδοση της μαστογραφίας στην παραπάνω υποομάδα γυναικών, συγκριτικά με τον γενικό πληθυσμό, δικαιολογείται απόλυτα από την πυκνή σύσταση του μαστού των νεαρών γυναικών. Άλλη μια εξήγηση, όμως, βρίσκεται και στο γεγονός ότι όταν το καρκίνωμα οφείλεται στην ύπαρξη μιας BRCA μετάλλαξης, ο όγκος συχνά μιμείται στην όψη καλοήθεις μάζες, οδηγώντας σε χαμηλότερη ευαισθησία [96], [97].

Έρευνες τόνισαν, επίσης, την καλύτερη ευαισθησία της μαστογραφίας έναντι της MRI εξέτασης στην ανίχνευση του LCIS καρκίνου. Υστερεί, ωστόσο στις περιπτώσεις διηθητικού καρκίνου του μαστού με ευαισθησία που δεν ξεπερνά το 40% [96], [97]. Η εξέλιξη της μαστογραφίας, όπως προαναφέραμε, η DBT έχει επιδείξει αυξημένη ευαισθησία (90,7% έναντι 85,2%) σε σύγκριση με την ψηφιακή μαστογραφία και μειωμένα ψευδώς θετικά αποτελέσματα. Ωστόσο, η αξιολόγηση των μικροασβετώσεων μπορεί να είναι περιορισμένη [97].

3.2.3 Η Διαγνωστική Απόδοση του Υπερήχου Μαστών

Πολλές έρευνες υποστηρίζουν ότι ο υπέρηχος είναι πιο αποτελεσματική τεχνική από την μαστογραφία σε νεαρές γυναίκες με πυκνό μαστό, καθώς εμφανίζει καλύτερα ποσοστά διαγνωστικής ευαισθησίας [97]. Το υπερηχογράφημα χρησιμοποιείται επιπλέον για τον εντοπισμό διηθημένων λεμφαδένων. Η ευαισθησία της διάγνωσης, στην περίπτωση αυτή, υπολογίζεται να είναι 61,4% και η ειδικότητα 82%. Η ποιότητα ενός υπερηχογραφήματος, δυστυχώς, επηρεάζεται πολύ εύκολα από εξωτερικούς παράγοντες. Πλήττεται από θορύβους, μη υπολογίσιμες παραμέτρους αλλά και τον ειδικό που το χειρίζεται. Επιπρόσθετα, λόγω της χαμηλής ειδικότητάς του μπορεί να οδηγήσει σε αύξηση των βιοψιών για καλοήθεις όγκους. Όλα αυτά το καθιστούν λιγότερο αξιόπιστο ως μεμονωμένο διαγνωστικό μέσο καρκίνου στον μαστό [96], [97].

Η ερευνητική ομάδα των A. Vourtsis et al. [96], στη μελέτη τους διαπίστωσαν ότι ο υπέρηχος μαστού αυξάνει το ποσοστό ανίχνευσης καρκίνου κατά 2,7 - 4,6 ανά 1.000 σε γυναίκες με πυκνούς μαστούς. Είναι σημαντικό ότι οι καρκίνοι που εντοπίζονται με το US έχει παρατηρηθεί ότι είναι ιδιαίτερα μικροί διηθητικοί καρκίνοι με αρνητικούς λεμφαδένες. Ωστόσο, η εξέταση με υπερηχογράφημα μαστού σε γυναίκες υψηλού

οικογενειακού κινδύνου, σε συνδυασμό με τη μαγνητική τομογραφία, έχει αποδειχθεί ότι έχει περιορισμένη αξία για την ανίχνευση της νόσου [96].

Παρά το γεγονός ότι ο υπέρηχος αποτελεί το συμπληρωματικό εργαλείο της μαστογραφίας για την ανίχνευση του καρκίνου του μαστού, υπάρχουν κάποιες περιπτώσεις που τα αποτελέσματά του δεν επιτρέπουν στους ειδικούς να εξάγουν ασφαλή συμπεράσματα. Σύμφωνα με εκτεταμένη μελέτη του Harmine M. Zonderland [116], ο υπέρηχος μπορεί να εμφανίσει ψευδώς αρνητικά ή ψευδώς θετικά αποτελέσματα. Συγκεκριμένα, ο υπέρηχος δεν δύναται να ανιχνεύσει την παρουσία αποτιτανώσεων. Επίσης, τόσο ο υπέρηχος όσο και η μαστογραφία δεν μπορούν να ανιχνεύσουν τους διάχυτα αναπτυσσόμενους καρκίνους που δεν έχουν κάποια διακριτή μάζα, γεγονός που οδηγεί σε ψευδώς αρνητικά αποτελέσματα. Επιπλέον κάποιες φορές παρατηρείται αλληλοεπικάλυψη καλοηθών και κακοήθων αλλοιώσεων, γεγονός που μπορεί να οδηγήσει είτε σε ψευδώς αρνητικά είτε σε ψευδώς θετικά αποτελέσματα εάν αυτά δεν αξιολογηθούν σχολαστικά από έμπειρους ακτινολόγους. Το γεγονός αυτό αποτελεί μειονέκτημα για την επιλογή του υπερήχου ως μοναδική μέθοδος διαλογής.

Η εξέταση του υπερήχου μπορεί να διαφέρει ως προς την ακρίβεια διάγνωσης ανάλογα με τη μορφολογία του καρκίνου του μαστού. Για παράδειγμα οι αστεροειδείς αλλοιώσεις σε λιπώδεις μαστούς και οι αλλοιώσεις που συνοδεύονται από μικροαποτιτανώσεις συχνά μπερδεύουν την αναγνώριση από τον υπέρηχο. Οι ακτινολόγοι θα πρέπει να γνωρίζουν ότι το αρνητικό αποτέλεσμα του υπερήχου δεν μπορεί να αποτελέσει αξιόπιστη διαβεβαίωση ότι δεν υπάρχει επιβλαβής αλλοίωση, γι' αυτό το λόγο ο υπέρηχος δε μπορεί να αποτελέσει το βασικό διαγνωστικό μέσο για την ανίχνευση του καρκίνου του μαστού αλλά δρα ευεργετικά συμπληρωματικά ως προς τη μαστογραφία. Ο υπέρηχος θα πρέπει να χρησιμοποιείται προκειμένου να ενισχύσει τη διάγνωση με τις πληροφορίες που παρέχει για τη μορφολογία και άλλα χαρακτηριστικά της συγκεκριμένης περιοχής για την οποία καλείται να κάνει συμπληρωματική διάγνωση. Ενώ σε περιπτώσεις που τα μαστογραφικά ευρήματα δεν μπορούν να συσχετιστούν με τα ευρήματα του υπερήχου, πρέπει να λαμβάνονται υπόψιν τα μαστογραφικά αποτελέσματα, προκειμένου να αποφευχθεί μια λανθασμένη διάγνωση [116].

3.2.4 Η Διαγνωστική Απόδοση της Μαγνητικής Τομογραφίας Μαστών

Η μαγνητική τομογραφία μπορεί να χρησιμοποιηθεί ως συμπληρωματική εξέταση τόσο για τη μαστογραφία όσο και για τον υπέρηχο για τον προσυμπτωματικό έλεγχο του καρκίνου του μαστού, αυξάνοντας τα ποσοστά ανίχνευσης. Στην τρέχουσα κλινική πρακτική, για τη διαχείριση του καρκίνου του μαστού, η μαγνητική τομογραφία δεν χρησιμοποιείται για τον βασικό διαγνωστικό έλεγχο αλλά κυρίως για τον καθορισμό του μεγέθους του όγκου ή/και για την ανίχνευση άλλων περιοχών με όγκους. Συχνά, η MRI χρησιμοποιείται ως μέσο παρακολούθησης μιας διάγνωσης υψηλού κινδύνου μετά τη θεραπεία. Η

πιθανή χρήση της για τη διαλογή νέων και υψηλού κινδύνου ασθενών αξίζει να διερευνηθεί περαιτέρω [4].

Η απόδοση στην ανίχνευση καρκίνου του μαστού σε ασθενείς υψηλού κινδύνου εμφάνισης της νόσου βελτιώνεται, βάσει σχετικών ερευνών, με την εφαρμογή της απεικονιστικής μεθόδου MRI, σε σχέση με την μαστογραφία. Η MRI μαστών παρέχει την υψηλότερη διαγνωστική ευαισθησία μεταξύ των τριών απεικονιστικών τεχνικών (MRI, μαστογραφίας και υπερήχου), κατά την εξέταση γυναικών υψηλού κινδύνου. Τα στοιχεία υποδηλώνουν ότι ένα ποσοστό καρκίνων της τάξης του 30% δεν θα είχε διαγνωστεί εξαρχής, αν δεν είχε εφαρμοστεί πολυμεθοδική προσέγγιση που να συμπεριλαμβάνει MRI [96]. Η ολική ευαισθησία της MRI στην κατηγορία ασθενών υψηλού κινδύνου κυμαίνεται στις διάφορες μελέτες μεταξύ 71 - 100%. Μια συγκριτική ανασκόπηση 11 μελετών έδειξε ότι η ευαισθησία της MRI ως μεμονωμένη εξέταση ήταν 77% σε γυναίκες με πυκνούς μαστούς [96], [97].

Το βασικότερο πλεονέκτημα της MRI είναι το ότι δεν επηρεάζεται από την πυκνότητα του μαστού. Η αξία της, λοιπόν, σε γυναίκες υψηλής επικινδυνότητας, διαφαίνεται στην ικανότητα ανίχνευσης όγκων που παρομοιάζουν καλοήθεις μάζες, αφού φαίνεται να ξεπερνά τα δύο σημαντικά αυτά εμπόδια. Όπως και η μαστογραφία, έτσι και η MRI παρουσιάζει αποκλίσεις στην απόδοση μεταξύ διαφορετικών ηλικιακών ομάδων. Συγκεκριμένα, σε γυναίκες άνω των 50 ετών τα αποτελέσματα της διάγνωσης ήταν καλύτερα, συγκριτικά με αυτά των νεότερων γυναικών. Επιπρόσθετα, σε περιπτώσεις διηθητικού καρκίνου, η ευαισθησία της μαγνητικής υπολογίστηκε να είναι 71,1% σχεδόν διπλάσια από της μαστογραφίας στην αντίστοιχη περίπτωση [96], [97].

3.2.5 Η Διαγνωστική Απόδοση της Βιοψίας Μαστών

Ορισμένες μελέτες αναφέρουν ότι η ευαισθησία της βιοψίας αναρρόφησης, FNA, για κακοήθεια είναι περίπου 64% όταν έχει ληφθεί ένα δείγμα, 91% όταν έχουν ληφθεί δείγματα και αντίστοιχα, η ειδικότητα κυμαίνεται γύρω στο 56% (ανεπαρκή ή μη ικανοποιητικά κυτταρολογικά παρασκευάσματα) [6]. Σύμφωνα με έρευνα που πραγματοποίησαν οι P.P.Frankel et al. [117], η ευαισθησία, η ειδικότητα, η θετική και αρνητική προγνωστική αξία και η ακρίβεια ήταν υψηλότερες για τη βιοψία με κόπτουσα βελόνα CNB από ό,τι για την FNA, ανεξάρτητα από το μέγεθος της βλάβης του μαστού. Συγκεκριμένα, η διαγνωστική ακρίβεια ήταν 97,5% για την CNB και 77,5% για την FNA [117]. Οι εν λόγω ερευνητές τονίζουν ότι παρά το γεγονός ότι η CNB έδειξε υψηλότερα ποσοστά στους δείκτες διάγνωσης που προαναφέρθηκαν, από την FNA, για ψηλαφητούς και μη ψηλαφητούς όγκους του μαστού, η FNA παραμένει χρήσιμη για την ελάχιστη επεμβατική διάγνωση των βλαβών του μαστού, ιδίως όταν τα αποτελέσματά της αναλύονται μαζί με την κλινική και την απεικονιστική εξέταση [117].

3.2.6 Η Διαγνωστική Απόδοση του Συνδυασμού των Εξετάσεων

Συνδυασμός Μαστογραφίας με Υπέρηχο μαστών

Πολλές έρευνες καταδεικνύουν τη μείζονος σημασίας συνεισφορά του Υπερήχου μαστών στη διαγνωστική διαδικασία ανίχνευσης του καρκίνου του μαστού ως συμπληρωματικού, ως προς τη μαστογραφία, μέσου. Σύμφωνα με την ερευνητική ομάδα του K.M Kelly [118], η μαστογραφία, που αποτελεί τη βασική μέθοδο διάγνωσης του καρκίνου του μαστού, παραλείπει πολλούς καρκίνους, ειδικά σε γυναίκες με πυκνό στήθος. Η εν λόγω ερευνητική ομάδα, συνέκρινε τη διαγνωστική απόδοση της μαστογραφίας μεμονωμένα με το συνδυασμό των εξετάσεων της μαστογραφίας και του υπερήχου μαστών, σε γυναίκες με πυκνό στήθος και/ή με αυξημένο κίνδυνο καρκίνου του μαστού. Η μελέτη διεξήχθη σε 4.419 γυναίκες που έκαναν τη μαστογραφία ρουτίνας και υπερήχο μαστών, και αξιολογήθηκαν οι καρκίνοι που προέκυψαν κατά τη διάρκεια της μελέτης και αξιολογήθηκε η πορεία των ασθενών σε διάρκεια ενός χρόνου. Για τις υπό μελέτη περιπτώσεις, υπολογίστηκε η ευαισθησία, η ειδικότητα και η PPV για τη σύσταση βιοψίας, για τη μαστογραφία μεμονωμένα και για τον συνδυασμό Μαστογραφίας και Υπερήχου μαστών. Σύμφωνα με την ερευνητική ομάδα των K. M Kelly et al. [118], η ευαισθησία μόνο για τη μαστογραφία υπολογίστηκε γύρω στο 40% και αυξήθηκε στο 81% με την προσθήκη της εξέτασης του υπερήχου Μαστών στη διαγνωστική διαδικασία.

Συνδυασμός Μαστογραφίας με Υπέρηχο μαστών και MRI

Μέχρι στιγμής έχει αναφερθεί ότι ο συνδυασμός μαστογραφίας και υπερήχων αυξάνει την ποσοστά ανίχνευσης του καρκίνου του μαστού σε σύγκριση με τη χρήση μόνο της μαστογραφίας. Υπάρχουν επίσης ισχυρά στοιχεία σχετικά με την προσθήκη του ελέγχου με MRI στην εκτίμηση του τύπου του όγκου, αν και η πυκνότητα του μαστού και το οικογενειακό ιστορικό αποτελούν σημαντικούς παράγοντες που μεταβάλλουν τη διαγνωστική απόδοση αυτών των απεικονιστικών τεχνικών [4]. Δύο συστηματικές ανασκοπήσεις που διερευνούν τη διαγνωστική απόδοση των τριών προαναφερόμενων μεθόδων εντοπίστηκαν στη δημοσιευμένη βιβλιογραφία το 2004 [119] και το 2007 [120].

Οι Irwig et al.[119] έδειξαν ότι η προσθήκη υπερήχων ή/και MRI στη μαστογραφία μπορεί να αυξήσει την ευαισθησία του προσυμπτωματικού ελέγχου, αλλά μπορεί επίσης να οδηγήσει σε ψευδώς θετικά ευρήματα. Οι Lord et al. [120] αξιολόγησαν την αποτελεσματικότητα της προσθήκης της MRI στη μαστογραφία, με ή χωρίς υπερηχογράφημα μαστού και τα ευρήματα κατέληξαν στο ότι ο συνδυασμός της MRI με τη μαστογραφία αυξάνει την κλίμακα της ψευδώς θετικής διάγνωσης στις ασθενείς. Επιπλέον οι Hadadi et al. [121], διαπίστωσαν ότι ο υπέρηχος αυξάνει την ευαισθησία στη διάγνωση, σε γυναίκες με πυκνούς και μη πυκνούς μαστούς, εάν χρησιμοποιηθεί συμπληρωματικά με τη μαστογραφία ωστόσο, η πρόσθετη χρήση MRI με υπέρηχο αυξάνει την ευαισθησία, ενώ μειώνει την ειδικότητα.

Σε μελέτες που συμπεριέλαβαν αποκλειστικά γυναίκες υψηλού κινδύνου, η απόδοση της MRI ανήλθε στο 94% όταν συνδυάστηκε με μαστογραφία, όταν μάλιστα η απόδοση της μαστογραφίας μόνο ήταν 34% [96], [97].

Συνδυασμός Μαστογραφίας με Υπέρηχο (HHUS, ABUS), DBT και CE - MAMMO

Οι I. Hadadi et al. [121], στην μελέτη τους συνέκριναν τη διαγνωστική απόδοση της μαστογραφίας έναντι του συνδυασμού της με συμπληρωματικές μεθόδους απεικόνισης, συμπεριλαμβανομένων του χειροκίνητου υπερήχου (HHUS), του αυτόματου υπερήχου (Automated Breast Ultrasound - ABUS), της DBT, της μαστογραφίας με ενισχυμένη αντίθεση (Contrast-Enhanced Mammography – CE - MAMMO) και της MRI σε γυναίκες με μη πυκνούς και πυκνούς μαστούς. Η εν λόγω ερευνητική ομάδα διεξήγαγε αναζήτηση στις βάσεις δεδομένων Medline, Embase, PubMed, CINAHL, Scopus και Web of Science έως τον Οκτώβριο του 2019, για τη διενέργεια μετα-ανάλυσης των δημοσιευμένων μελετών. Οι I. Hadadi et al. [121], διαπίστωσαν ότι στους πυκνούς μαστούς, η προσθήκη συμπληρωματικών απεικονιστικών μεθόδων αύξησε σημαντικά τα ποσοστά ανίχνευσης καρκίνου: HHUS (σχετικός κίνδυνος [RR] = 1,49 - 95%, CI, 1,19 - 1,86), ABUS (RR = 1,44, 95% CI, 1,16 - 1,78), DBT (RR = 1,38, 95% CI, 1,14 - 1,67), CE - MAMMO (RR = 1,37, 95% CI, 1,12 - 1,69) και MRI (RR = 2,16, 95% CI, 1,81 - 2,58). Στους μη πυκνούς μαστούς, η HHUS και η MRI έδειξαν σημαντική αύξηση των ποσοστών ανίχνευσης καρκίνου σε αντίθεση με την DBT: HHUS (RR = 1,14, 95% CI, 1,01 - 1,29), MRI (RR = 1,78, 95% CI, 1,14 - 2,77) και DBT (RR = 1,09, 95% CI, 1,13 - 1,75). Συμπερασματικά, η προσθήκη συμπληρωματικών απεικονιστικών μεθόδων στην μαστογραφία αυξάνει τα ποσοστά ανίχνευσης καρκίνου σε γυναίκες με πυκνούς και μη πυκνούς μαστούς. Το υπερηχογράφημα και η MRI αυξάνουν τα ποσοστά των ποσοστών ανάκλησης (recall rate) η αλλιώς αληθώς θετικών περιπτώσεων (True Positive Rate – TPR) ανεξάρτητα από τη πυκνότητα του μαστού ενώ η MRI οδηγεί σε υψηλότερες τιμές τόσο για τα ποσοστά ανίχνευσης του καρκίνου όσο και για τα ποσοστά ανάκλησης [121].

Συμπερασματικά, οι συμπληρωματικές μέθοδοι απεικόνισης του μαστού αυξάνουν τα ποσοστά ανίχνευσης του καρκίνου σε γυναίκες με πυκνούς και μη πυκνούς μαστούς. Η συμπληρωματική MRI έχει υψηλότερο ποσοστό ανίχνευσης του καρκίνου ενώ η συμπληρωματική DBT ενήργησε καλύτερα στη μείωση των ψευδώς θετικών αποτελεσμάτων, το υπερηχογράφημα και η MRI αύξησαν τα ψευδώς θετικά αποτελέσματα ανεξάρτητα από τη σύνθεση του μαστού (πυκνοί και μη πυκνοί μαστοί). Σημειώνεται ότι η DBT δεν έχει δοκιμαστεί σε γυναίκες με υψηλό κίνδυνο ανάπτυξης καρκίνου του μαστού, όπως οι περισσότερες μελέτες που αφορούν σε υπερήχους και MRI [121].

3.2.7 Πίνακες Ανασκόπησης της Απόδοσης των Διαγνωστικών Δοκιμασιών και των Συνδυασμών τους

Παρακάτω παρουσιάζονται συγκεντρωτικά σε πίνακες τα αποτελέσματα της ανασκόπησης σχετικά με την απόδοση (υπό όρους ευαισθησίας και ειδικότητας) των διαφόρων σχετικών διαγνωστικών δοκιμασιών σε διάφορα σενάρια προληπτικού ελέγχου για τον καρκίνο του μαστού, σύμφωνα με τις μελέτες και μετα-αναλύσεις που παρουσιάστηκαν στην προηγούμενη ενότητα. Σκοπός των πινάκων ανασκόπησης είναι να παρουσιαστούν συγκεντρωμένα τα επιμέρους στοιχεία προκειμένου να είναι ευκολότερη η συνοπτική τους επισκόπηση.

Στον παρακάτω πίνακα (Πίνακας 2) περιλαμβάνεται η ευαισθησία και η ειδικότητα της μαστογραφίας στις διάφορες ομάδες ασθενών ανάλογα με την ηλικιακή ομάδα που ανήκουν και την πυκνότητα των μαστών (κλίμακα A - D κατά BI-RADS σε ποσοστό % του συνόλου) τους όπως παρουσιάστηκαν από την ερευνητική ομάδα των P. Steyerova et al. [97].

Πίνακας 2: Ευαισθησία και Ειδικότητα μαστογραφίας ανάλογα με την ηλικία και την πυκνότητα μαστού [97]

Ηλικία	A	B	C	D	Ευαισθησία (%)	Ειδικότητα
40-49	3	23	57	17	65,6 - 69,7	90,7 - 90,9
50-59	6	37	49	8	72,9 - 73,8	91,6 - 92,3
60-69	12	43	41	4	73,3	93
70-79	13	51	32	4	81,4	94,1
≥	13	46	34	7	86,1	94,3
Εμφυτεύματα Στήθους					45	97,7

Όπως διαπιστώνουμε στον Πίνακα 2, η ευαισθησία της μαστογραφίας κυμαίνεται από 45 έως 86,1% σε όλα τα διαφορετικά ηλικιακά γκρουπ και τις πυκνότητες μαστών, ενώ η ειδικότητα αντίστοιχα κυμαίνεται από 90,7 έως 97,9%. Είναι αξιοσημείωτο το γεγονός ότι παρατηρούμε την ευαισθησία και την ειδικότητα να ακολουθούν την αντίστροφη πορεία, δηλαδή όσο αυξάνεται η ευαισθησία μειώνεται η ειδικότητα.

Η ίδια ερευνητική ομάδα [97], μελέτησε την απόδοση του υπερηχογραφήματος μαστού και των τεχνικών βιοψίας FNA και CNB για την διάγνωση διηθημένων λεμφαδένων όπως παρουσιάζεται στον Πίνακα 3:

Πίνακας 3: Σύγκριση διαγνωστικής απόδοσης μεταξύ υπερηχογραφήματος μαστών και δύο μεθόδων βιοψίας (FNA & CNB) για τη διάγνωση των διηθημένων λεμφαδένων [97]

	Υπερηχογράφημα Μαστών (US)	FNA	CNB
Ευαισθησία %	61,4	72,2	83,3
Ειδικότητα %	82	100	100

Όπως βλέπουμε, η βιοψία, και με τις δύο τεχνικές, υπερτερεί του υπερηχογραφήματος, ωστόσο η μέτρια προς καλή απόδοση του υπερήχου μπορεί να αποδειχθεί σωτήρια όταν δεν υπάρχει δυνατότητα άλλης εξέτασης. Σύμφωνα με τους P. Steyerova et al. [97], όταν η πρόσβαση σε καλύτερες απεικονιστικές εξετάσεις εμποδίζεται είτε λόγω κόστους, είτε λόγω διαθεσιμότητας, το υπερηχογράφημα μαστού έρχεται να καλύψει τα κενά, ως ένα οικονομικό, εύχρηστο και, υπό προϋποθέσεις, αποτελεσματικό εργαλείο πρόβλεψης της διήθησης των λεμφαδένων.

Στη συνέχεια, στον Πίνακα 4, αποτυπώνεται η διαγνωστική αξία του συνδυασμού της μαστογραφίας με την ψηφιακή τομογραφία (DBT), τον υπέρηχο (US) και τη μαγνητική τομογραφία (MRI) σε περιπτώσεις γυναικών με πολύ πυκνούς μαστούς και αρνητική μαστογραφία [97]:

Πίνακας 4: Απόδοση απεικονιστικών μεθόδων όταν συνδυάζονται με μαστογραφία σε πυκνούς μαστούς (γυναίκες με αρνητική μαστογραφία) [97]

	+DBT	+US (χειροκίνητος)	+US (αυτόματος)	+MRI
Ευαισθησία %	91 - 93	80 - 83	67,7	75 - 100
Ειδικότητα %	69 - 71	86,4 – 94,5	91, 6	78,1 – 93,2

Από τον παραπάνω πίνακα, είναι εμφανής η αύξηση των δεικτών απόδοσης στην ανίχνευση καρκίνου του μαστού, κυρίως με τη χρήση της DBT και της MRI. Συνεπώς, μία πολυμεθοδική προσέγγιση για την διάγνωση του καρκίνου του μαστού φαίνεται να παρέχει πιο έγκυρα συμπεράσματα.

Συνοψίζοντας, η ερευνητική ομάδα των N. Aristokli et al. [4] δημοσίευσαν μια εξαιρετικά πρόσφατη ανασκόπηση (2022), όπου συγκρίνονται η μαστογραφία (MAMMO), η μαγνητική τομογραφία (MRI), το υπερηχογράφημα (US) και οι συνδυασμοί τους ως προς τη διαγνωστική τους απόδοση στην ανίχνευση του καρκίνου του μαστού, ανάλογα με τον τύπο του όγκου, την πυκνότητα του μαστού και το ιστορικό της γυναίκας.

Σύμφωνα με τους N. Aristokli et al. [4], μετά από μια εκτεταμένη μετα – ανάλυση άλλων ερευνών που διεξήγαγαν, η δημοσιευμένη βιβλιογραφία κατέδειξε ότι η χρήση κάθε μεθόδου ξεχωριστά έδειξε συνολική ευαισθησία και ειδικότητα για τη μαγνητική τομογραφία (MRI) στο 94,6% (εύρος 85,7% - 100%) και 74,2% (εύρος 25% - 100%) αντίστοιχα, ενώ για τη μαστογραφία (MAMMO) ήταν 54,5% (εύρος 27% - 86,8%) και 85,5% (εύρος 62,9% - 98,8%) αντίστοιχα, και για το υπερηχογράφημα (US) ήταν 67,2% (εύρος 26,9% - 87,5%) και 76,8% (εύρος 18,8% - 96,9%) αντίστοιχα (βλ. Πίνακας 5). Όταν συνδυάστηκαν τα αποτελέσματα και των τριών τεχνικών, προέκυψε ευαισθησία 97,7% (εύρος 95%- 100%) και ειδικότητα 63,3% (εύρος 37,1% - 87,5%). Επιπλέον, η μαστογραφία ενισχυμένης αντίθεσης (Contrast-enhanced) CE-MAMMO και η μαγνητική τομογραφία ενισχυμένης αντίθεσης CE-MRI απεικόνισαν συνολική ευαισθησία 90,5% (εύρος 80,9% - 100%) και 91,5% (εύρος 89,1% - 93,8%) αντίστοιχα, και συνολική ειδικότητα 52,6% (εύρος 15% - 76,1%) και 64,7% (εύρος 43,7% - 85,7%). Εν κατακλείδι, τα διαγνωστικά αποτελέσματα για όλους τους συνδυασμούς μεθόδων φαίνονται στον Πίνακας 5.

Πίνακας 5: Συνολική ευαισθησία και ειδικότητα για τη διαγνωστική ικανότητα των απεικονιστικών μεθόδων μέσω των μελετών που περιλαμβάνονται στην ανασκόπηση των N. Aristokli et al [4].

Απεικονιστικό Σύστημα	Αριθμός Μελετών	Ευαισθησία (%)	Αριθμός Μελετών	Ειδικότητα(%)
MRI	22	94,6%	16	74,2%
MAMMO	17	54,5%	12	85,5%
US	14	62,2%	10	76,8%
CE-MRI	2	91,5%	2	64,7%
CE-MAMMO	3	90,5%	3	52,6%
CE-MAMMO + US	1	90,5%	1	61,4%
MRI + MAMMO	3	95,8%	3	70,1%
MRI + US	2	92,3%	2	76,8%
MAMMO + US	8	72,2%	6	87,8%
MRI + MAMMO + US	3	97,7%	3	63,3%

Συμπερασματικά, όπως βλέπουμε στον παραπάνω πίνακα (Πίνακας 5) ο συνδυασμός των τεχνικών αυξάνει τη διαγνωστική απόδοση σε σύγκριση με τη μεμονωμένη χρήση τους. Συγκεκριμένα, ως μεμονωμένες μέθοδοι, η υψηλότερη ευαισθησία έχει παρατηρηθεί για τη μαγνητική τομογραφία και η χαμηλότερη ευαισθησία για τη μαστογραφία, ανεξάρτητα από τον τύπο του μαστού, την πυκνότητα και το ιστορικό. Η

ευαισθησία αυξάνεται ακόμη περισσότερο από τον συνδυασμό US & MRI ή MAMMO & MRI ή MRI & MAMMO & US. Η ειδικότητα φαίνεται να επηρεάζεται από το μέγεθος, τον τύπο του όγκου και το ιστορικό της ασθενούς, ωστόσο με βάση την πυκνότητα του μαστού, η υψηλότερη ειδικότητα παρατηρήθηκε μόνο από την χρήση US [4].

3.3 Συζήτηση - Συμπεράσματα

Η έγκαιρη διάγνωση είναι ύψιστης σημασίας για τον περιορισμό του καρκίνου του μαστού και αποτελεί την σωστή βάση, ώστε να επιτευχθεί η βελτίωση της πρόγνωσης και του ποσοστού επιβίωσης των γυναικών. Όσο νωρίτερα εντοπίζεται η νόσος, τόσο καλύτερα είναι τα προγνωστικά επιβίωσης της ασθενούς, ενώ μπορεί να αποφευχθούν και οι πιο δύσκολης μορφής θεραπείες. Κάθε περίπτωση της νόσου πρέπει να αντιμετωπίζεται ξεχωριστά στο πλαίσιο της εξατομικευμένης ιατρικής, έτσι ώστε να επιτυγχάνονται τα βέλτιστα αποτελέσματα.

Οι διαγνωστικές μέθοδοι για την ανίχνευση του καρκίνου του μαστού εφαρμόζονται άλλοτε χωριστά και άλλοτε ως συνδυασμός. Όταν ένα εύρημα κρίνεται ασαφές από μία μέθοδο, οι άλλες μέθοδοι μπορούν να παρέχουν συμπληρωματικές πρόσθετες πληροφορίες. Η βασική διαγνωστική μέθοδος, δηλαδή η εξέταση αναφοράς, είναι η μαστογραφία. Ο πιο συχνός συνδυασμός είναι η προσθήκη υπερήχων στη μαστογραφία είτε για την περαιτέρω αξιολόγηση ενός ευρήματος που εντοπίζεται στη μαστογραφία, είτε σε ασθενείς με κλινικά ευρήματα όπου η μαστογραφία είναι ασαφής, είτε όταν ένα παθολογικό εύρημα μπορεί να αποκρύπτεται ή να μην αποτυπώνεται στην εικόνα της μαστογραφίας (πυκνή περιοχή του μαστού, κλινικό εύρημα σε περιφερειακή περιοχή του μαστού ή στη μασχάλη κ.λπ.). Εάν στη μαστογραφία ή στο υπερηχογράφημα διαπιστωθεί μια ανωμαλία που υποδηλώνει κακοήθεια, συνήθως προτιμάται η βιοψία ιστού σε σχέση με την περαιτέρω χρήση προηγμένων απεικονιστικών μεθόδων (μαγνητική τομογραφία) [97].

Συνεπώς, η μαστογραφία και το υπερηχογράφημα μαστών αποτελούν τις πιο ευρέως χρησιμοποιούμενες μεθόδους στην απεικόνιση του μαστού και αποτελούν τα βασικότερα εργαλεία για την έγκαιρη διάγνωση της νόσου του καρκίνου του μαστού. Η MRI είναι μια εξειδικευμένη απεικονιστική μέθοδος που επιλέγεται σε περιπτώσεις πολύ πυκνών μαστών προκειμένου να αυξήσει την απόδοση της διάγνωσης. Η μαστογραφία αποτελεί την κύρια διαγνωστική μέθοδο για την ανίχνευση του καρκίνου του μαστού, κατά συνέπεια είναι κρίσιμης σημασίας η δυνατότητα της αξιολόγησης των μαστογραφικών ευρημάτων με όρους ποσοστού κινδύνου ανάπτυξης της νόσου. Ως εκ τούτου, το σύστημα BI-RADS, που αναλύθηκε στην υποενότητα 3.1.1.1 χρησιμοποιείται πλέον ευρέως για την ερμηνεία των ευρημάτων της μαστογραφίας, επιτρέποντας ομοιομορφία στις εκθέσεις και βελτίωση της επικοινωνίας μεταξύ των ακτινολόγων και των κλινικών ιατρών. Η τελευταία έκδοση περιλαμβάνει μια ενότητα για το

υπερηχογράφημα μαστού και τη μαγνητική τομογραφία μαστού, αντανακλώντας τη σταδιακά εκτεταμένη χρήση αυτών των απεικονιστικών μεθόδων στην απεικόνιση του μαστού [104].

Οι διαγνωστικές μέθοδοι που περιεγράφηκαν ανωτέρω επέδειξαν διαφορετικά ποσοστά διαγνωστικής απόδοσης σε καταστάσεις όπου τα δεδομένα ήταν διαφορετικά. Είναι γεγονός ότι η απόδοση της εξέτασης της μαστογραφίας, στην εύρεση καρκινικών αλλοιώσεων, έχει βελτιωθεί σημαντικά με την συνεχή εξέλιξη των τεχνολογικών μέσων. Όπως αναφέρθηκε, ωστόσο, επηρεάζεται από πολλούς παράγοντες, ανάμεσά τους η πυκνότητα του μαστού και ο τύπος καρκίνου που καλείται να εντοπίσει και διαγνώσει. Πιο αναλυτικά, σε πληθώρα δημοσιευμένων μελετών και μετά-αναλύσεων, η μαστογραφία φαίνεται να υστερεί στην παραγωγή ξεκάθαρων συμπερασμάτων για τις γυναίκες που ανήκαν στην κατηγορία υψηλού κινδύνου ανάπτυξης καρκίνου του μαστού και για τις γυναίκες με πυκνούς μαστούς. Συγκεκριμένα, η διαγνωστική απόδοση της μαστογραφίας αποδεικνύεται ασθενέστερη σε γυναίκες με πυκνό μαστό από ό,τι σε εκείνες με μη πυκνό [97].

Μετά τη μαστογραφία, ο υπέρηχος είναι αναμφίβολα η πιο ευρέως χρησιμοποιούμενη τεχνική στην απεικόνιση του μαστού και είναι η μόνη απεικονιστική μέθοδος που χρησιμοποιείται σε νεαρές ασθενείς. Αποτελεί τη συμπληρωματική μέθοδο της μαστογραφίας στη διαγνωστική διαδικασία και παρουσιάζει αυξημένη ευαισθησία αλλά μειωμένη ειδικότητα σε σχέση με τη μαστογραφία. Επιπλέον, η μαγνητική τομογραφία μαστών παρέχει την υψηλότερη ευαισθησία (94,6%) ανάμεσα στις τρεις απεικονιστικές μεθόδους, βοηθώντας να εξαχθούν ασφαλέστερα συμπεράσματα για την πραγματική κατάσταση των γυναικών, ειδικά όσων κινδυνεύουν από καρκίνο του μαστού [4].

Σύμφωνα με τα αποτελέσματα που παρουσιάζονται στον Πίνακα 5, ο συνδυασμός της μαστογραφίας και του υπερήχου παρουσιάζει την υψηλότερη τιμή στην ειδικότητα (87,8%) ενώ η ευαισθησία αυξάνεται στο 72,2% σε σχέση με την ευαισθησία που εμφανίζει μόνη της η μαστογραφία (54,5%). Από τα αποτελέσματα του πίνακα διαπιστώνουμε ότι ο συνδυασμός και των τριών απεικονιστικών μεθόδων, της μαστογραφίας, του υπερήχου και της μαγνητικής τομογραφίας, αυξάνει σημαντικά την ευαισθησία της διάγνωσης (97,7%) μειώνοντας ωστόσο την ειδικότητα (63,3%) [4]. Παρά το γεγονός ότι η χρήση της MRI αυξάνει σημαντικά την τιμή της ευαισθησίας της διάγνωσης, η MRI εμφανίζει λιγότερη ειδικότητα από τις άλλες μεθόδους απεικόνισης [122]. Ως εκ τούτου, οι ενδείξεις για τη συστηματική χρήση της μαγνητικής τομογραφίας είναι ένα θέμα διαμάχης για την επιστημονική κοινότητα, καθώς μπορεί να καθυστερήσει τη διαγνωστική διαδικασία. Η MRI μπορεί να οδηγήσει σε περιττές βιοψίες και κατά συνέπεια σε υπερθεραπεία (αυξημένα ποσοστά πλευρικής και διμερούς μαστεκτομής). Λαμβάνοντας επιπλέον υπόψιν το κόστος της εξειδικευμένης εξέτασης της μαγνητικής τομογραφίας, συμπεραίνουμε ότι η συστηματική συμπερίληψή της στη διαγνωστική διαδικασία θα επέφερε σημαντική οικονομική επιβάρυνση για τα συστήματα υγείας.

Συνεπώς, η MRI δεν είναι κατάλληλη για συστηματική χρήση κατά τη διάρκεια της διαγνωστικής διαδικασίας, ωστόσο υπάρχουν συγκεκριμένες περιπτώσεις όπου η διαγνωστική της αξία είναι αναμφισβήτητη [93].

Κεφάλαιο 4^ο

Στατιστικά Μοντέλα Εκτίμησης

Κινδύνου του Καρκίνου του Μαστού και

Ευφυή Συστήματα Υποστήριξης

Κλινικής Απόφασης

4.1 Στατιστικά Μοντέλα Εκτίμησης Κινδύνου

Σύμφωνα με τη διεθνή βιβλιογραφία έχει αναπτυχθεί πληθώρα μοντέλων προκειμένου να ποσοτικοποιήσουν τη συνδυαστική επίδραση διαφόρων παραγόντων κινδύνου και να προβλέψουν είτε τον κίνδυνο εμφάνισης καρκίνου του μαστού, είτε τον κίνδυνο ύπαρξης μια γενετικής μετάλλαξης υψηλού κινδύνου, ειδικά στα γονίδια BRCA1 και BRCA2, ή τον κίνδυνο και των δύο. Αυτά τα μοντέλα εκτίμησης κινδύνου καρκίνου του μαστού μπορούν να διαχωριστούν σε εκείνα που χρησιμοποιούν κυρίως ορμονικούς και περιβαλλοντικούς παράγοντες και εκείνα που επικεντρώνονται περισσότερο στον κληρονομικό κίνδυνο. Στη συνέχεια αναλύονται τα σημαντικότερα υφιστάμενα μοντέλα πρόβλεψης κινδύνου για τον καρκίνο του μαστού, συμπεριλαμβανομένων των παραγόντων κινδύνου που ενσωματώνουν, τη βασική μεθοδολογία για την ανάπτυξή τους, τις πληροφορίες που παρέχει το καθένα, τα πλεονεκτήματα και τους περιορισμούς τους, τις σχετικές μελέτες επικύρωσης και τον τρόπο πρόσβασης στο κάθε ένα από αυτά για κλινικούς ή ερευνητικούς σκοπούς [123].

Μοντέλα που προβλέπουν τον Κίνδυνο Ανάπτυξης Καρκίνου του Μαστού

Ορμονικά/Περιβαλλοντικά Μοντέλα

Μοντέλο Gail: [AKA: NCI model, Breast Cancer Risk Assessment Tool (BCRAT)] Το μοντέλο Gail είναι ένα από τα πρώιμα μοντέλα κινδύνου για τον καρκίνο του μαστού, που αναπτύχθηκε με τη χρήση δεδομένων περιπτώσεων - ελέγχου από το πρόγραμμα Επίδειξης Ανίχνευσης Καρκίνου του Μαστού (Breast Cancer Detection Demonstration Project - BCDDP) που περιγράφηκε για πρώτη φορά το 1989

[124]. Το μοντέλο προορίζεται για γυναίκες 35 ετών και άνω. Το μοντέλο Gail συνδυάζει δύο στοιχεία: έναν προοπτικό υπολογισμό του απόλυτου κινδύνου και την τροποποίηση του σχετικού κινδύνου μέσω ενός πολυμεταβλητού μοντέλου λογιστικής παλινδρόμησης που περιλαμβάνει παράγοντες κινδύνου. Αυτοί είναι κυρίως ορμονικοί (ηλικία κατά την εμμηναρχί, ηλικία κατά τον πρώτο τοκετό) και παθολογικοί (αριθμός προηγούμενων βιοψιών μαστού, προσωπικό ιστορικό άτυπης υπερπλασίας), αλλά περιλαμβάνουν επίσης την αναγνώριση του κληρονομικού κινδύνου, ενσωματώνοντας τον αριθμό των θηλυκών συγγενών πρώτου βαθμού με καρκίνο του μαστού. Το πιο πρόσφατο μοντέλο Gail προβλέπει τον κίνδυνο ανάπτυξης διηθητικού καρκίνου του μαστού με την πάροδο του χρόνου, έως και την ηλικία των 90 ετών. Σύμφωνα με την συστηματική έρευνα και μετα-ανάλυση των X.Wang et al. [125] το 2018, η συγκεντρωτική ευαισθησία και ειδικότητα του Gail μοντέλου 1 ήταν 0,63 (95% CI 0,27–0,89) και 0,91 (95% CI 0,87–0,94) και του μοντέλου Gail 2 ήταν 0,35 (95% CI 0,17–0,59) και 0,86 (95% CI 0,76–0,92) αντίστοιχα [125].

Μοντέλο Care: Το μοντέλο Care που αναπτύχθηκε το 2007, είναι η επέκταση του μοντέλου Gail στις Αφροαμερικανίδες. Το βασικό κίνητρο για την ανάπτυξή του ήταν η ανησυχία ότι το αρχικό μοντέλο Gail που είχε αναπτυχθεί για μια μεγάλη ομάδα κυρίως λευκών γυναικών, θα μπορούσε ενδεχομένως να υποεκτιμήσει τον κίνδυνο καρκίνου του μαστού στον πληθυσμό των Αφροαμερικανίδων γυναικών. Το μοντέλο αναπτύχθηκε με τη χρήση μεθοδολογίας παρόμοιας με το μοντέλο Gail και χρησιμοποιεί τους ίδιους παράγοντες κινδύνου. Αναπτύχθηκε με τη χρήση δεδομένων ελέγχου - περιπτώσεων από τη μελέτη Women's Contraceptive and Reproductive Experiences (CARE), η οποία περιελάμβανε 1607 περιπτώσεις γυναικών που διαγνώστηκαν με διηθητικό καρκίνο του μαστού και 1647 περιπτώσεις ελέγχου [126].

Προεμμηνοπαυσιακά/μεταεμμηνοπαυσιακά μοντέλα Barlow/BCSC: Ενσωματώνοντας τη μαστογραφική πυκνότητα, το μοντέλο Barlow (αρχικά Breast Cancer Surveillance Consortium – BCSC, αν και η ονομασία αυτή έχει αντικατασταθεί από το μοντέλο Tice, που περιγράφεται αργότερα) περιγράφηκε το 2006. Το μοντέλο Barlow προορίζεται για γυναίκες που υποβάλλονται σε μαστογραφία και εκτιμά τον κίνδυνο εμφάνισης είτε διηθητικού καρκίνου του μαστού ή DCIS εντός ενός έτους από τη διαλογή. Δεν μπορεί να εφαρμοστεί σε γυναίκες με προηγούμενη διάγνωση καρκίνου του μαστού ή DCIS ή σε γυναίκες με προηγούμενη αύξηση του μαστού ή μαστεκτομή. Περιλαμβάνει δύο μοντέλα: ένα προεμμηνοπαυσιακό και ένα μετεμμηνοπαυσιακό. Χρησιμοποιήθηκε η λογιστική παλινδρόμηση για τον εντοπισμό των παραγόντων κινδύνου και την ενσωμάτωσή τους σε κάθε μοντέλο. Εξετάστηκαν πολλοί παράγοντες, αλλά μόνο όσοι βρέθηκαν στατιστικά σημαντικοί ενσωματώθηκαν στα δύο τελικά μοντέλα. Για τις προεμμηνοπαυσιακές γυναίκες, επιλέχθηκαν ως παράγοντες κινδύνου: η ηλικία, η πυκνότητα του μαστού (κατά κατηγορία BI-RADS), ο αριθμός των συγγενών πρώτου βαθμού με καρκίνο του μαστού και οι προηγούμενες επεμβάσεις στο μαστό. Για τις μετεμμηνοπαυσιακές γυναίκες, οι παράγοντες κινδύνου περιελάμβαναν δημογραφικούς παράγοντες (ηλικία, εθνικότητα (ισπανόφωνη ή μη ισπανόφωνη), φυλή) BMI, ορμονικούς παράγοντες

(ηλικία της γέννησης του πρώτου παιδιού, χρήση ορμονικής θεραπείας, τύπος εμμηνόπαυσης), κληρονομικοί παράγοντες (αριθμός πρώτου βαθμού συγγενών με καρκίνο του μαστού), μαστογραφικά ευρήματα (πυκνότητα του μαστού, προηγούμενα μαστογραφικά ευρήματα), και προηγούμενες επεμβάσεις στο μαστό. Το μοντέλο μπορεί να εντοπίζει τις γυναίκες υψηλού κινδύνου καλύτερα από το μοντέλο Gail, αν και η προγνωστική ακρίβεια παραμένει μέτρια. Το c - statistic ήταν 0,631 (95% CI , 0,618 - 0,644) για τις προεμμηνοπαυσιακές γυναίκες και 0,624 (95% CI, 0,619 - 0,630) για τις μετεμμηνοπαυσιακές γυναίκες. Συμπερασματικά, η πυκνότητα του μαστού αποτελεί ισχυρό πρόσθετο παράγοντα κινδύνου για τον καρκίνο του μαστού [127].

Μοντέλο Tice: Το 2008, οι Tice et al. [128], αφού προηγουμένως είχε σημειωθεί ότι η υψηλότερη πυκνότητα του μαστού είχε συνδεθεί με αυξημένο κίνδυνο καρκίνου του μαστού, ανέπτυξαν ένα μοντέλο που χρησιμοποιεί αναλογικά κινδύνους για την πρόβλεψη του 5ετούς κινδύνου διηθητικού καρκίνου του μαστού σε γυναίκες που δεν είχαν διαγνωστεί προηγουμένως με καρκίνο μαστού. Το μοντέλο Tice [ή αλλιώς: το μοντέλο για τον καρκίνο του μαστού Breast Cancer Surveillance Consortium (BCSC)] ενσωματώνει την κατηγοριοποίηση της πυκνότητας του μαστού BI-RADS στο μοντέλο κινδύνου [129], έχοντας δοκιμάσει πιλοτικά ένα παρόμοιο μοντέλο με την ενσωμάτωση της πυκνότητας του μαστού στο μοντέλο Gail τρία χρόνια πριν [130]. Εκτός από την πυκνότητα του μαστού, το μοντέλο Tice χρησιμοποιεί το ιστορικό των καρκίνων του μαστού σε συγγενείς πρώτου βαθμού, το προσωπικό ιστορικό βιοψίας μαστού, τη φυλή και την εθνικότητα. Το μοντέλο αναπτύχθηκε χρησιμοποιώντας τη μελέτη κοόρτης BCSC. Όταν επικυρώθηκε το μοντέλο χρησιμοποιώντας τον πληθυσμό της μελέτης BCSC, το c-statistic ήταν 0,66 (95% CI 0,65-0,67) και το μοντέλο επέδειξε καλό odds ratio 1,03 (95% CI 0,99-1,06), αν και επέδειξε ελαφρά υποεκτίμηση του κινδύνου στις νεότερες γυναίκες ηλικίας 40-44 ετών (0,94, 95% CI του 0,89-1,00), στις γυναίκες από την Ασία (0,95, 95% CI του 0,81-1,12), και τις ισπανόφωνες γυναίκες (0,94, 95% CI 0,85-1,04). Το μοντέλο Tice/BCSC επικυρώθηκε περαιτέρω με μια ξεχωριστή κοόρτη από το Mayo Mammography Health Study (MMHS), κατά την οποία ενσωματώθηκε η βαθμολογία πολυγονιδιακού κινδύνου (Polygenic Risk Score - PRS), αξιολογήθηκε και κατέδειξε καλή διακριτική ακρίβεια στην MMHS τόσο με όσο και χωρίς την προσθήκη του PRS και διαπιστώθηκε αύξηση του c – statistic από 0,66 σε 0,69 [131].

Μοντέλο Bodian: Το εν λόγω μοντέλο αρχικά περιγράφηκε το 1996, και εκτιμά τον κίνδυνο ανάπτυξης διεισδυτικού ή in situ καρκίνου του μαστού με την πάροδο του χρόνου, εξετάζει τα υποκείμενα σε βάθος 5ετίας με βάση την παρουσία του LCIS σε συνδυασμό με την ενσωμάτωση του σχετικού κινδύνου που απορρέει από άλλες ορμονικούς, κληρονομικούς και παθολογικούς παράγοντες κινδύνου [132]. Οι σχετικοί κίνδυνοι υπολογίστηκαν χρησιμοποιώντας τυποποιημένες αναλογίες νοσηρότητας και την πιθανότητα ανάπτυξης καρκίνου που εκτιμήθηκε με τη χρήση της ανάλυσης Kaplan-Meier. . Οι παράγοντες

κινδύνου που αξιολογήθηκαν περιλάμβαναν την ηλικία, το ιστορικό των καρκίνων του μαστού σε συγγενείς πρώτου βαθμού, τις εγκυμοσύνες, την ηλικία του πρώτου τοκετού, την έκταση του LCIS, τον αριθμό των διαγνώσεων με LCIS, και την παρουσία πρόσθετης καλοήθους νόσου του μαστού [συμπεριλαμβανομένης της άτυπης υπερπλασίας των πόρων (Atypical Ductal Hyperplasia - ADH), της αδένωσης και των κύστεων]. Η μελέτη στην οποία βασίζεται το μοντέλο αποτελούνταν από 236 ασθενείς με LCIS που διαγνώστηκαν μεταξύ 29 και 82 ετών και παρακολούθηθηκαν για περίπου 18 έτη.

Μοντέλο BBD-BC (Benign breast disease to breast cancer): Το μοντέλο εκτίμησης κινδύνου BBD-BC, που μερικές φορές αναφέρεται ως μοντέλο Pankratz Hartmann [133], περιγράφηκε για πρώτη φορά το 2015. Αυτό το μοντέλο είναι κατάλληλο για ασθενείς που έχουν υποβληθεί σε βιοψίες για καλοήγη νόσο του μαστού, ιδίως για νόσο που προσδίδει αυξημένο κίνδυνο εμφάνισης καρκίνου του μαστού. Το μοντέλο χρησιμοποιεί σχετικούς κινδύνους για την εκτίμηση ενός εξαρτώμενου από την ηλικία λόγου κινδύνου του διηθητικού ή in situ καρκίνου του μαστού, υπολογίζοντας τον 5ετή και 10ετή κίνδυνο από τη στιγμή διεξαγωγής της βιοψίας του μαστού, καθώς και τον κίνδυνο εμφάνισης καρκίνου του μαστού κατά τη διάρκεια της ζωής της γυναίκας έως την ηλικία των 90 ετών [134]. Μεταβλητές που χρησιμοποιήθηκαν στο μοντέλο περιλαμβάνουν: τα ιστολογικά ευρήματα (συνολική ιστολογική εντύπωση, αριθμός εστιών άτυπης υπερπλασίας, έκταση της λοβιακής αναδίπλωσης, παρουσία ακτινωτής ουλής, σκληρυντική αδένωση, τους ορμονικούς παράγοντες (ηλικία πρώτου τοκετού σε συνδυασμό με το συνολικό αριθμό τοκετών ως συνδυαστική μεταβλητή), και κληρονομικούς παράγοντες (οικογενειακό ιστορικό καρκίνου του μαστού). Το οικογενειακό ιστορικό ταξινομείται ως ισχυρός παράγοντας κινδύνου εάν υπάρχει ιστορικό καρκίνου του μαστού σε συγγενή πρώτου βαθμού κάτω των 50 ετών ή σε δύο ή περισσότερους συγγενείς με τον ένα να είναι πρώτου βαθμού, ως πιο αδύναμος παράγοντας εάν υπάρχει ιστορικό σε οποιονδήποτε συγγενή μικρότερου βαθμού, και ως μηδενικός παράγοντας εάν δεν υπάρχει οικογενειακό ιστορικό με καρκίνο του μαστού. Το μοντέλο βασίζεται στην κοόρτη Mayo benign breast disease (BBD), η οποία αποτελείται από 9.854 γυναίκες ηλικίας 18-85 ετών που διεξήγαγαν βιοψίες για κάτι που τελικά διαπιστώθηκε ότι ήταν καλοήθης νόσος του μαστού. Για την ανάπτυξη του μοντέλου επιλέχθηκαν τυχαία 377 γυναίκες που είχαν αναπτύξει καρκίνο του μαστού ως δείγμα μαζί με 734 περιπτώσεις ελέγχου.

Μοντέλο Rosner-Colditz: Οι Rosner και Colditz περιέγραψαν για πρώτη φορά το μοντέλο τους το 1996, καθιστώντας το ένα από τα παλαιότερα μοντέλα κινδύνου. Το αρχικό μοντέλο βασίστηκε σε δεδομένα από το Nurses' Health Study που συλλέχθηκαν από το 1976 έως το 1990. Τα δεδομένα συλλέχθηκαν προοπτικά μέσω ερωτηματολογίων, αν και αναλύθηκαν αναδρομικά για την ανάπτυξη του μοντέλου, το οποίο αρχικά περιλάμβανε μόνο αναπαραγωγικούς παράγοντες κινδύνου [135]. Μια διευρυμένη έκδοση του μοντέλου δημοσιεύθηκε το 2000, βασισμένη σε δεδομένα από το 1980 έως το 1994, περιλάμβανε πρόσθετους παράγοντες κινδύνου όπως είναι η προηγούμενη καλοήθης νόσος του μαστού, το οικογενειακό ιστορικό

καρκίνου του μαστού για συγγενείς πρώτου βαθμού, η χρήση ορμονών μετά την εμμηνόπαυση, ο BMI, το ύψος και η κατανάλωση αλκοόλ [136]. Το Rosner-Colditz προβλέπει τον κίνδυνο ανάπτυξης διηθητικού καρκίνου του μαστού με την πάροδο του χρόνου έως την ηλικία των 70 ετών σε γυναίκες χωρίς προηγούμενη διάγνωση καρκίνου του μαστού χρησιμοποιώντας τους ορμονικούς, κληρονομικούς, παθολογικούς και άλλους κλινικούς παράγοντες που περιγράφονται ανωτέρω.

Το μοντέλο επικυρώθηκε χρησιμοποιώντας ένα ανεξάρτητο σύνολο δεδομένων από σχετική μελέτη που διεξήχθη στην Καλιφόρνια (California Teachers Study - CTS) [137]. Όπως φάνηκε το μοντέλο απέδωσε εξίσου καλά όταν εφαρμόστηκε στην προαναφερθείσα ανεξάρτητη μελέτη επικύρωσης όπως και όταν εφαρμόστηκε απευθείας στα δεδομένα της μελέτης Nurses' Health Study. Το μοντέλο αυτό δεν είναι επί του παρόντος διαθέσιμο μέσω ειδικού λογισμικού ή διαδικτυακού πακέτου, αλλά μπορεί να υπολογιστεί με τη χρήση τυποποιημένων στατιστικών πακέτων. Όπως απέδειξαν οι A.B Rosner et al.[137], το μοντέλο Rosner-Colditz αποδίδει σταθερά καλά όταν εφαρμόζεται σε ένα ανεξάρτητο σύνολο δεδομένων. Οι επιδόσεις είναι καλύτερες στην πρόβλεψη της επίπτωσης μεταξύ των γυναικών 47-69 ετών και σε χρονικό διάστημα 5 ετών. Οι τιμές της περιοχής (area under the ROC curve) AUC υπερβαίνουν εκείνες του Gail κατά 3-5 % όταν και τα δύο εφαρμόζονται στο ανεξάρτητο σύνολο δεδομένων επικύρωσης. Συγκεκριμένα, οι τιμές AUC για το μοντέλο Rosner-Colditz είναι 0,608 στο NHS και 0,609 στο CTS και για το Gail είναι 0,569 και 0,572 αντίστοιχα. Τα μοντέλα μπορούν να βελτιωθούν περαιτέρω με την προσθήκη της πυκνότητας του μαστού ή άλλων δεικτών κινδύνου.

Μοντέλο Tworoger: Το μοντέλο Tworoger είναι μια τροποποίηση των μοντέλων Gail και Rosner-Colditz που αναπτύχθηκαν το 2014 που ενσωματώνει τη μέτρηση των ενδογενών ορμονών μαζί με τις μεταβλητές που ήδη χρησιμοποιούνται από αυτά τα μοντέλα για την πρόβλεψη του κινδύνου διηθητικού καρκίνου του μαστού [124], [135], [138]. Η τροποποίηση αναπτύχθηκε και δοκιμάστηκε σε μετεμμηνοπαυσιακές γυναίκες που δεν ελάμβαναν επί του παρόντος εξωγενείς ορμόνες στη μελέτη Nurses' Health Study, μια κοόρτη που είχε προηγουμένως χρησιμοποιηθεί για την επικύρωση του μοντέλου Gail. Για να προσδιοριστεί η επίδραση της ενσωμάτωσης των επιπέδων των ορμονών σε κάθε μοντέλο, υπολογίστηκε το στατιστικό μέτρο c (concordance statistic) για την πρόβλεψη διηθητικού καρκίνου του μαστού εντός 5 ετών συγκρίθηκαν τόσο με όσο και χωρίς την τροποποίηση Tworoger. Σύμφωνα με τους S. Tworoger et al. [138], η AUC για το μοντέλο Gail (μόνο του) ήταν 0,549 και για το μοντέλο Rosner-Colditz αντίστοιχα ήταν 0,589. Για το μοντέλο Gail, η συμπερίληψη της οιστρονής, της προλακτίνης και των επιπέδων τεστοστερόνης αύξησε την AUC κατά 5,9 μονάδες, δηλαδή επετεύχθη AUC 0,60 ($p = 0,003$) γεγονός που υποδηλώνει ότι η ενσωμάτωση αυτών των τριών επιπέδων ορμονών βελτιώνει σημαντικά τη διακριτική ικανότητα του μοντέλου. Για το μοντέλο Rosner-Colditz, η χρήση όλων των επιπέδων ορμονών βελτίωσε την AUC κατά 3,3 μονάδες (νέα AUC 0,62) [138].

Μοντέλο Petracci: Το μοντέλο Petracci, που περιγράφηκε από τους Petracci et al. το 2011 [139], αναπτύχθηκε από μια ιταλική μελέτη κοόρτης μέρος μιας πολυκεντρικής μελέτης ελέγχου περιπτώσεων που εξέταζε την επίδραση της διατροφής στον διηθητικό καρκίνο του μαστού και συνεπώς προέβλεπε τον κίνδυνο εμφάνισης διηθητικού καρκίνου του μαστού καρκίνου [140]. Η εν λόγω ερευνητική ομάδα ταξινόμησε τους παράγοντες κινδύνου στο μοντέλο σε «μη τροποποιήσιμους», οι οποίοι περιλάμβαναν: την ηλικία κατά την εμμηναρχή, την ηλικία του πρώτου τοκετού, τον αριθμό των προηγούμενων βιοψιών μαστού, τον αριθμό των συγγενών πρώτου βαθμού με καρκίνο του μαστού, την ενδεχόμενη επαγγελματική σωματική δραστηριότητα και το μορφωτικό επίπεδο, και στους "τροποποιήσιμους" παράγοντες κινδύνου: την κατανάλωση αλκοόλ, τη σωματική δραστηριότητα στον ελεύθερο χρόνο και τον BMI [139]. Το μοντέλο επικυρώθηκε σε μια ξεχωριστή ιταλική μελέτη κοόρτης μέρος της μελέτης Florence - European Prospective Investigation into Cancer and Nutrition (EPIC) και βρέθηκε να έχει καλή βαθμονόμηση με λόγο Ο:Ε 1,10 (95% CI 0,96-1,26), αλλά φτωχή διακριτική ικανότητα με το c-statistic (AUC) να είναι 0,62 (95% CI 0,555-0,689) για γυναίκες κάτω των 50 ετών και 0,57 (95% CI 0,519-0,614) για τις γυναίκες 50 ετών και άνω [139].

Κληρονομικά Μοντέλα

Μοντέλο Claus: Το μοντέλο Claus (η αλλιώς Yale University Model) περιγράφηκε για πρώτη φορά από τους Claus et al. το 1991 [141], πριν από την ανακάλυψη των γονιδιακών τύπων BRCA1 και BRCA2. Το μοντέλο προορίζεται να χρησιμοποιηθεί σε γυναίκες με τουλάχιστον μία γυναίκα συγγενή πρώτου ή δεύτερου βαθμού με καρκίνο του μαστού. Χρησιμοποιεί κληρονομικές μεταβλητές για την πρόβλεψη του κινδύνου ανάπτυξης για καρκίνο του μαστού κατά τη διάρκεια της ζωής της γυναίκας, τόσο για διηθητικό καρκίνο όσο και για DCIS, και αρχικά αξιολογεί τον αριθμό των συγγενών πρώτου (όχι δεύτερου) βαθμού με καρκίνο του μαστού ή καρκίνο των ωοθηκών, λαμβάνοντας υπόψη την ηλικία τους κατά τη διάγνωση. Αυτό το μοντέλο στη συνέχεια επικαιροποιήθηκε για να λάβει υπόψη του συγγενείς πρώτου βαθμού με καρκίνο των ωοθηκών και για να ενσωματώσει συγγενείς δευτέρου βαθμού [141]. Το μοντέλο βασίστηκε σε μια μελέτη κοόρτης γυναικών από το Cancer and Steroid Hormone Study (CASH), μια μελέτη ελέγχου - περιπτώσεων που διεξήχθη από τα Κέντρα Ελέγχου Ασθενειών (Centers for Disease - CDC), με τα άτομα να αντιστοιχίζονται με βάση τη γεωγραφική τοποθεσία και την ηλικία (εύρος 5 ετών). Το μοντέλο κατέληξε στο συμπέρασμα της ύπαρξης ενός επικρατούντος γονότυπου που αυξάνει σημαντικά τον κίνδυνο καρκίνου κατά τη διάρκεια της ζωής του ατόμου, αναγνωρίζοντας παράλληλα ότι η πλειονότητα των περιπτώσεων καρκίνου του μαστού δεν συνδέεται με γονιδιακή μετάλλαξη υψηλού κινδύνου. Στη συγκριτική μελέτη των E.Amir et al. [142] η AUC για το μοντέλο Claus ήταν 0,716.

Μοντέλα πρόβλεψης Γενετικών Μεταλλάξεων

Κληρονομικά Μοντέλα

Μοντέλο Myriad I: Αναφέρεται επίσης ως Shattuck-Eidens [143], και είναι ένα πρώιμο μοντέλο για τον προσδιορισμό του κινδύνου της επιβλαβούς μετάλλαξης BRCA1 για τις γυναίκες με προσωπικό ή οικογενειακό ιστορικό καρκίνου του μαστού χρησιμοποιώντας κληρονομικές μεταβλητές. Το Myriad I δεν μπορεί να εφαρμοστεί σε γυναίκες χωρίς προσωπικό ή οικογενειακό ιστορικό καρκίνου του μαστού ή των ωοθηκών. Για την ανάπτυξη του εν λόγω μοντέλου χρησιμοποιήθηκαν 798 γυναίκες που είχαν ήδη είχαν οικογενειακό ιστορικό υψηλού κινδύνου με βάση την εμφάνιση καρκίνου του μαστού σε νεαρή ηλικία ή πολλαπλών περιπτώσεων καρκίνου του μαστού και/ή καρκίνου των ωοθηκών στο οικογενειακό τους ιστορικό, αλλά χωρίς τεκμηριωμένες περιπτώσεις γενετικών μεταλλάξεων. Τα δείγματά τους αναλύθηκαν με ειδικές μεθόδους όπου και εντοπίστηκαν επιβλαβείς μεταλλάξεις του BRCA1. Πραγματοποιήθηκε ανάλυση λογιστικής παλινδρόμησης για τον προσδιορισμό των παραγόντων κινδύνου που προβλέπουν τις μεταλλάξεις BRCA1. Οι παράγοντες κινδύνου περιλάμβαναν: την ηλικία πρώτης διάγνωσης του καρκίνου του μαστού ή των ωοθηκών, την εβραϊκή εθνικότητα Ashkenazi, τους αμφοτερόπλευρους καρκίνους του μαστού, τον αριθμό συγγενών (οποιοδήποτε βαθμού) που έχουν προσβληθεί από καρκίνο του μαστού, τον αριθμό συγγενών (οποιοδήποτε βαθμού) που έχουν προσβληθεί από καρκίνο των ωοθηκών, τον αριθμό συγγενών (οποιοδήποτε βαθμού) που έχουν προσβληθεί τόσο από καρκίνο του μαστού όσο και από καρκίνο των ωοθηκών. Ως μοντέλο πρόβλεψης γενετικών μεταλλάξεων, αυτό το μοντέλο προβλέπει μόνο τις επιβλαβείς μεταλλάξεις BRCA1. Στη συνέχεια αναπτύχθηκαν άλλα μοντέλα, συμπεριλαμβανομένου του μοντέλου MYRIAD II, που προβλέπουν επίσης την πιθανότητα της μετάλλαξης BRCA2.

Μοντέλο Myriad II: Το μοντέλο Myriad II, που αναφέρεται επίσης ως μοντέλο Frank, δημοσιεύτηκε για πρώτη φορά το 1998 [144] και προορίζεται για χρήση σε γυναίκες με προσωπικό ή οικογενειακό ιστορικό καρκίνου του μαστού. Το μοντέλο χρησιμοποιεί λογιστική παλινδρόμηση για την εκτίμηση της πιθανότητας ύπαρξης BRCA1 ή BRCA2 μετάλλαξης και χρησιμοποιεί ως παράγοντες: το προσωπικό ιστορικό καρκίνου του μαστού, την ηλικία εμφάνισης του καρκίνου του μαστού που κατηγοριοποιείται σε μικρότερη ή μεγαλύτερη των 50 ετών, ιστορικό καρκίνου των ωοθηκών, ιστορικό ανδρικού καρκίνου του μαστού, ιστορικό συνδυασμού καρκίνου του μαστού και των ωοθηκών, οικογενειακό ιστορικό καρκίνου του μαστού, συμπεριλαμβανομένης της ηλικίας κατά την διάγνωση και ιστορικό καρκίνου των ωοθηκών σε οποιαδήποτε ηλικία σε συγγενείς πρώτου ή δεύτερου βαθμού. Ο γενετικός έλεγχος πραγματοποιήθηκε από τα εργαστήρια Myriad Genetic Laboratories σε γυναίκες οι οποίες είτε είχαν διαγνωστεί με καρκίνο του μαστού πριν από την ηλικία των 50 ετών, είτε με καρκίνο των ωοθηκών, ή είχαν γυναίκα συγγενή

πρώτου ή δεύτερου βαθμού με διάγνωση καρκίνου του μαστού σε ηλικία κάτω των 50 ετών ή αντίστοιχα διάγνωση καρκίνου των ωοθηκών. Το μοντέλο δοκιμάστηκε και βελτιώθηκε σε ένα πρόσθετο σύνολο 10.000 γυναικών. Οι εκτιμήσεις είναι διαθέσιμες για γυναίκες Ashkenazim εβραϊκής καταγωγής και μη. Στις μελέτες επικύρωσης, το μοντέλο Myriad II παρουσίασε ευαισθησία και ειδικότητα πρόβλεψης για ύπαρξη μετάλλαξης BRCA 0,71 και 0,63, αντίστοιχα. Άλλες συγκριτικές μελέτες βρήκαν ομοίως ότι το Myriad II έχει συγκρίσιμες ευαισθησίες και ειδικότητες με άλλα μοντέλα που προβλέπουν την ύπαρξη της μετάλλαξης BRCA.

Μοντέλα Penn και Penn II: Το μοντέλο Couch (AKA: UPenn ή μοντέλο Penn) περιγράφηκε για πρώτη φορά από τους Couch et al. το 1997 [145], και προορίζεται για ασθενείς με οικογενειακό ιστορικό καρκίνου του μαστού. Το μοντέλο εκτιμά τον πιθανότητα της μετάλλαξης BRCA1 τόσο σε άτομα όσο και σε μέλη της οικογένειας τους με βάση τη σχέση τους με ένα προσβεβλημένο από καρκίνο μέλος της οικογένειας. Το μοντέλο Couch προήλθε από ανάλυση λογιστικής παλινδρόμησης που βασίστηκε σε 263 οικογένειες κυρίως καυκάσιων γυναικών με οικογενειακό ιστορικό καρκίνου του μαστού, με ή χωρίς καρκίνο των ωοθηκών. Οι γυναίκες με εβραϊκή καταγωγή Ashkenazi αναλύθηκαν ξεχωριστά χρησιμοποιώντας ένα διαφορετικό σύνολο υπολογισμένων πιθανοτήτων. Το μοντέλο του Couch (Penn) εξελίχθηκε στη συνέχεια σε μοντέλο Penn II, το οποίο προβλέπει την πιθανότητα ύπαρξης της μετάλλαξης BRCA1 ή BRCA2 σε ένα άτομο ή μέλος της οικογένειας του. Αυτό το μοντέλο προέκυψε από λογιστική ανάλυση παλινδρόμησης 861 γενεαλογικών δέντρων από την Ευρώπη και τη Βόρεια Αμερική και επεκτάθηκε προκειμένου να συμπεριλάβει επιπλέον το ιστορικό, συμπεριλαμβανομένης της παρουσίας αμφοτερόπλευρων καρκίνων του μαστού, η οποία αποκλείστηκε από το μοντέλο Couch, ο ανδρικός καρκίνος του μαστού, και το οικογενειακό ιστορικό καρκίνου παγκρέατος και προστάτη [146].

Σύστημα βαθμολόγησης του Μάντσεστερ (Manchester Scoring System - MSS): Το MSS αναπτύχθηκε το 2004 από τους Evans et al. ως ένα απλό σύστημα βαθμολόγησης αρχικά βασισμένο αποκλειστικά στην αξιολόγηση του οικογενειακού ιστορικού (MSS-2004) [147] στη συνέχεια επικαιροποιήθηκε το 2005, με την εκτέλεση περισσότερων εκτεταμένων γενετικών ελέγχων σε μεγαλύτερη πληθυσμιακή ομάδα και πάλι το 2009 με την ίδια διαδικασία (MSS-2009) [148]. Το MSS είναι ένα υπολογιστικό εργαλείο για την υποστήριξη της σύστασης για γενετικό έλεγχο (συγκεκριμένα, για τις μεταλλάξεις BRCA1 και BRCA2), που βασίζεται στην αξιολόγηση του οικογενειακού ιστορικού και των παθολογικών δεδομένων. Για κάθε άτομο της οικογένειας (άνδρας ή γυναίκα) αποδίδεται μια βαθμολογία σχετικά με τον κίνδυνο καρκίνου του μαστού, καρκίνου των ωοθηκών, καθώς και καρκίνου του προστάτη και του παγκρέατος, τόσο με βάση το μητρικό όσο και το πατρικό γενεαλογικό δέντρο και το άθροισμα των βαθμολογιών στα δύο δέντρα υπολογίζεται. Εάν ένα από αυτά τα αθροίσματα υπερβαίνει ένα δεδομένο όριο, που αντιστοιχεί σε ποσοστό 10-20% πιθανότητα μετάλλαξης, συνιστάται γενετική εξέταση. Επίσης, οι ηλικίες κατά τη διάγνωση

θεωρούνται ότι προσδίδουν μια πιο ακριβή πιθανότητα εντοπισμού της BRCA1 ή BRCA2 μετάλλαξης [147].

Εργαλείο αξιολόγησης οικογενειακού ιστορικού (Family history assessment tool - FHAT): Το FHAT είναι ένα αλγόριθμος βαθμολόγησης, που αναπτύχθηκε από τους Gilpin et al. [148] για να βοηθήσει τους γιατρούς στον καθορισμό των γυναικών που χρήζουν γενετικής συμβουλευτικής, μεταξύ εκείνων που δεν έχουν ακόμη διαγνωστεί με καρκίνο του μαστού και/ή των ωοθηκών, χρησιμοποιώντας πληροφορίες για καρκίνους στο πλαίσιο του οικογενειακού περιβάλλοντος. Ο στόχος ήταν να αναπτυχθεί ένα μοντέλο για την επιλογή εκείνων των γυναικών που βρίσκονταν σε περίπου διπλάσιο ποσοστό από τον πληθυσμιακό κίνδυνο εμφάνισης καρκίνου του μαστού ή των ωοθηκών. Οι στόχοι δεν περιλάμβαναν την αξιολόγηση του κατά πόσον η οικογένεια θα πληρούσε τα ερευνητικά κριτήρια για γενετικές εξετάσεις ή την εκτίμηση της πραγματικής πιθανότητας εμφάνισης καρκίνου ή εύρεσης της BRCA1 ή BRCA2 μετάλλαξης.

Μοντέλο Bellcross: Το μοντέλο που αναπτύχθηκε από τους Bellcross et al. [149] (AKA: B-RST™ Genetic Referral Screening Tool) σχεδιάστηκε ως απάντηση στην USPSTF (U.S. Preventive Services Task Force) του 2005, σύμφωνα με την οποία οι γυναίκες με οικογενειακό ιστορικό που υποδηλώνει αυξημένο κίνδυνο καρκίνου του μαστού ή των ωοθηκών λόγω επιβλαβών μεταλλάξεων σε BRCA1 ή BRCA2, παραπέμπονται για γενετική συμβουλευτική και ενδεχομένως σε εξέταση BRCA. Η ειδική ομάδα κατέληξε στο συμπέρασμα ότι υπήρχε εξαιρετική ανάγκη για απλά εργαλεία που μπορούν να είναι εύκολα να χρησιμοποιηθούν στο γραφείο ενός γιατρού πρωτοβάθμιας περίθαλψης προκειμένου να προσδιοριστεί εάν οι ασθενείς πρέπει να παραπεμφθούν σε γενετική συμβουλευτική ή όχι [89].

Το B-RST™ είναι ένα απλό ερωτηματολόγιο που έχει σχεδιαστεί για τον εντοπισμό των γυναικών που μπορεί να διατρέχουν κίνδυνο ανάπτυξης καρκίνου του μαστού ή των ωοθηκών λόγω μεταλλάξεων των BRCA1 ή BRCA2. Το ερωτηματολόγιο περιέχει τις ακόλουθες ερωτήσεις που αφορούν στο οικογενειακό τους ιστορικό: (1) ιστορικό καρκίνου του μαστού ή των ωοθηκών πριν από την ηλικία των 50 ετών στις ίδιες, στη μητέρα, στην αδελφή, στην κόρη, σε γιαγιάδες ή θείες, σε δύο ή περισσότερες γυναίκες συγγενείς με καρκίνο του μαστού μετά την ηλικία των 50 ετών που είναι στην ίδια πλευρά της οικογένειας (2) καρκίνο του μαστού σε οποιονδήποτε άνδρα συγγενή και (3) εβραϊκή καταγωγή Ashkenazi. Ένα απλό κατώφλι (δύο ή περισσότερες θετικές απαντήσεις στα παραπάνω) χρησιμοποιείται για τον εντοπισμό των "θετικών" ασθενών που πρέπει να παραπεμφθούν για ενδελεχή γενετική συμβουλευτική.

Λογισμικό COS: Το λογισμικό COS (Case Only Study) σχεδιάστηκε από τους Berrino et al. με βάση το BRCAPRO (περιγράφεται παρακάτω) έκδοση 4.0 για χρήση σε μια ευρωπαϊκή μελέτη που εξέταζε την αλληλεπίδραση της γενετικής και της διατροφής σε περιπτώσεις καρκίνου του μαστού με νεαρή ηλικία εμφάνισης [150]. Έχει σχεδιαστεί κυρίως για χώρες με αναφερόμενες περιπτώσεις καρκίνου του μαστού

που κυμαίνονται από 50 έως 100 ανά 100.000 γυναίκες ετησίως. Ομοίως με το BRCAPRO, το λογισμικό COS χρησιμοποιείται για την πρόβλεψη του κινδύνου ύπαρξης της μετάλλαξης BRCA1 ή BRCA2, αν και αυτό το μοντέλο λαμβάνει υπόψη και ένα υποθετικό τρίτο γονίδιο BRCA για τις περιπτώσεις ισχυρού οικογενειακού ιστορικού με αρνητική εξέταση μετάλλαξης BRCA1 ή BRCA2. Το λογισμικό δεν προβλέπει τον κίνδυνο εμφάνισης καρκίνου του μαστού ή των ωοθηκών, αλλά βασίζεται στο μοντέλο BRCAPRO. Το COS προσαρμόζει το BRCAPRO χρησιμοποιώντας συγκεκριμένες παραδοχές ανά χώρα όσον αφορά τις, ανάλογα την ηλικία, σποραδικές και κληρονομικές περιπτώσεις καρκίνου του μαστού. Στο μοντέλο ενσωματώνονται δεδομένα σχετικά με συγγενείς τρίτου και τέταρτου βαθμού και αποκλείονται τα άτομα που υποβάλλονται σε προφυλακτική μαστεκτομή ή ωοθηκεκτομή.

Μοντέλα που προβλέπουν τον Κίνδυνο Ανάπτυξης Καρκίνου του Μαστού και τις Γενετικές Μεταλλάξεις

Κληρονομικά Μοντέλα

Μοντέλο BRCAPRO: Το μοντέλο BRCAPRO περιγράφηκε αρχικά από τον Berry et al. το 1997 ως ένα μοντέλο πρόβλεψης του κινδύνου ύπαρξης της επιβλαβούς μετάλλαξης BRCA1 [151] και στη συνέχεια επεκτάθηκε ώστε να συμπεριλάβει και τη μετάλλαξη BRCA2 [152]. Το μοντέλο απευθύνεται σε άτομα, άνδρες και γυναίκες, τόσο με όσο και χωρίς οικογενειακό ιστορικό καρκίνου του μαστού ή των ωοθηκών (ή και των δύο), και μπορεί να χρησιμοποιηθεί για την εκτίμηση είτε του εξατομικευμένου είτε του οικογενειακού κινδύνου. Το BRCAPRO προβλέπει την πιθανότητα ύπαρξης επιβλαβούς μετάλλαξης στα BRCA1 ή BRCA2 γονίδια, την πιθανότητα εμφάνισης διηθητικού καρκίνου του μαστού ή καρκίνου των ωοθηκών (για μη διαγνωσμένα άτομα) και, ως πρόσφατη προσθήκη [153], την πιθανότητα ανάπτυξης αμφίπλευρου καρκίνου του μαστού (για ήδη διαγνωσμένα άτομα). Οι προβλέψεις της ύπαρξης γενετικής μετάλλαξης υπολογίζονται με βάση το οικογενειακό ιστορικό και τις δημοσιευμένες εκτιμήσεις του επιπολασμού των BRCA1 και BRCA2, των ποσοστών καρκίνου του μαστού στο γενικό πληθυσμό, και χρήση του θεωρήματος του Bayes. Το BRCAPRO ενημερώνεται συνεχώς καθώς οι πληροφορίες σχετικά με αυτά τα ποσοστά βελτιώνονται. Στο μοντέλο ενσωματώνονται ειδικοί παράγοντες που αφορούν σε συγγενικά πρόσωπα της κάθε εξεταζόμενης περίπτωσης, συγκεκριμένα περιλαμβάνουν τη σχέση με τον συμβουλευόμενο (εξεταζόμενη περίπτωση), το ενδεχόμενο να είναι δίδυμοι, την τρέχουσα ηλικία του συγγενούς, τη διάγνωση ή μη καρκίνου του μαστού και καρκίνου των ωοθηκών, καθώς και την ηλικία ενδεχόμενης διάγνωσης. Το BRCAPRO ενσωματώνει επίσης παθολογικούς δείκτες (ER, CK14, CK5/6, PR, HER2) για γνωστές περιπτώσεις καρκίνου του μαστού, φυλή, εθνικότητα (Εβραίοι Ashkenazi και μη Ashkenazi, Ιταλοί Ασιάτες, Άλλο), καθώς και πληροφορίες σχετικά με επεμβάσεις όπως μαστεκτομή (διευκρινίζοντας την αμφοτερόπλευρη μαστεκτομή) ή ωοθηκεκτομή [153].

Μοντέλο BOADICEA: Ο αλγόριθμος BOADICEA (Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm) περιγράφηκε για πρώτη φορά το 2002 από τους Antoniou et al. [154], ενώ ένα βελτιωμένο μοντέλο δημοσιεύθηκε το 2004 [155]. Το εν λόγω μοντέλο προορίζεται για γυναίκες με πιθανώς αυξημένο κίνδυνο εμφάνισης καρκίνου του μαστού, και καρκίνου των ωοθηκών λόγω του οικογενειακού τους ιστορικού. Ομοίως με το BRCAPRO, αυτό το μοντέλο προβλέπει τόσο την πιθανότητα ενός ατόμου που φέρει τη μετάλλαξη BRCA1 ή BRCA2 όσο και τον κίνδυνο εμφάνισης καρκίνου του μαστού με την πάροδο του χρόνου [155]. Ο κίνδυνος υπολογίζεται χρησιμοποιώντας το οικογενειακό ιστορικό καρκίνου του μαστού και των ωοθηκών, συμπεριλαμβανομένης της ηλικίας της διάγνωσης, χρησιμοποιώντας το θεώρημα του Bayes. Το αρχικό μοντέλο αναπτύχθηκε με βάση μια μελέτη κοόρτης γυναικών του ABC (Anglican Breast Cancer) που διαγνώστηκαν με καρκίνο του μαστού σε ηλικία κάτω των 55 ετών [154]. Για την επικύρωση των υπολογισμών του μοντέλου, τα προβλεπόμενα αποτελέσματα συγκρίθηκαν με άλλα μοντέλα χρησιμοποιώντας μια προηγουμένως δημοσιευμένη μελέτη κοόρτης. Οι οικογενειακοί σχετικοί κίνδυνοι (familial relative risks - FRR), που αντικατοπτρίζουν τον σχετικό κίνδυνο εμφάνισης καρκίνου του μαστού για έναν συγγενή σε σύγκριση με κάποιον από τον γενικό πληθυσμό, που προέκυψαν από αυτή τη μελέτη, συγκρίθηκαν επίσης με τις εκτιμήσεις των μοντέλων Claus και BRCAPRO. Ο προβλεπόμενος σχετικός κίνδυνος, όπως υπολογίζεται από το BOADICEA συσχετίζεται καλά με τα παρατηρούμενα αποτελέσματα. Ο κίνδυνος για τους αμφίπλευρους καρκίνους του μαστού, όπως προβλέπεται από το αλγόριθμο, επικυρώθηκε στη συνέχεια σε μια αξιολόγηση άλλων ομάδων, συμπεριλαμβανομένου του μητρικού ασθενών καρκίνου του μαστού του Κονέκτικατ (Connecticut Tumor Registry) και του Σουηδικού Οικογενειακού Μητρώου Καρκίνου (Swedish Family Cancer Registry) [155]. Το 2008, το μοντέλο επικαιροποιήθηκε περαιτέρω με την ενσωμάτωση δεδομένων από δύο πρόσθετες ομάδες (την Εθνική Μελέτη Ελέγχου - Περιπτώσεων του Ηνωμένου Βασιλείου και τη μελέτη του Μάντσεστερ), καθώς και με τη συνεκτίμηση του οικογενειακού ιστορικού ανδρικού καρκίνου του μαστού, καρκίνου του προστάτη, και καρκίνου του παγκρέατος [156]. Σύμφωνα με τη συγκριτική μελέτη των E.Amir et al. [142], το BOADICEA παρείχε επίσης την καλύτερη διάκριση μεταξύ των φορέων μετάλλαξης και των μη φορέων και ήταν στατιστικά σημαντικά καλύτερο από όλα τα άλλα μοντέλα εκτός από το BRCAPRO (AUC: BOADICEA = 0,77, BRCAPRO = 0,76, και Myriad II = 0,72).

Μοντέλο Jonker: Βασιζόμενοι σε άλλα μοντέλα που προέβλεπαν τον κίνδυνο να φέρει κάποια τις μεταλλάξεις BRCA1 ή BRCA2 (κυρίως στο μοντέλο 1991-Claus), οι Jonker et al. το 2003 [157] ανέπτυξαν ένα μοντέλο που προοριζόταν για τον γενικό πληθυσμό, και όχι για τις ομάδες υψηλού κινδύνου. Αυτό το μοντέλο λαμβάνει επίσης υπόψη οικογενειακές «ομαδοποιήσεις», που δεν αποδίδονται σε μεταλλάξεις των BRCA1 ή BRCA2, μέσω του προσδιορισμού ενός υποθετικού γονιδίου που ονομάζεται «BRCAu» με σκοπό να ληφθεί υπόψη κάθε πολυγονιδιακή συνιστώσα του οικογενειακού κινδύνου. Κατά συνέπεια,

αναπτύχθηκαν δύο ξεχωριστά μοντέλα: το ένα υποθέτοντας ότι το BRCAu είναι κυρίαρχο και το άλλο υποθέτοντας ότι είναι υπολειπόμενο γονίδιο στο μοτίβο κληρονομικότητας. Η πρόβλεψη του κινδύνου καρκίνου του μαστού διαπιστώθηκε παρόμοια με τη χρήση και των δύο μοντέλων σε εσωτερικές μελέτες εκτίμησης. Οι παράγοντες κινδύνου που ενσωματώθηκαν στο μοντέλο και οι επακόλουθες εκτιμήσεις κινδύνου βασίστηκαν σε δημοσιευμένα δεδομένα σχετικά με τους κινδύνους και την επίπτωση του καρκίνου του μαστού και περιελάμβαναν το ιστορικό όλων των μελών της οικογένειας με μονόπλευρο ή αμφίπλευρο καρκίνο του μαστού, καρκίνο των ωοθηκών, ανδρικό καρκίνο του μαστού, με ενσωμάτωση των ανωτέρω πληροφοριών τόσο για τους προσβεβλημένους όσο και για τους μη προσβεβλημένους από καρκίνο μαστού συγγενείς. Το μοντέλο προβλέπει τον κίνδυνο εμφάνισης καρκίνου του μαστού κατά τη διάρκεια της ζωής του ατόμου ενσωματώνοντας τον υπολογισμένο κίνδυνο της BRCA1, BRCA2 ή BRCAu μετάλλαξης [157].

Μοντέλο Tyrer-Cuzick: Το μοντέλο Tyrer-Cuzick (TC) (AKA: International Breast Cancer Study - IBIS) περιγράφηκε αρχικά από τους Tyrer, Duffy και Cuzick το 2004 και εκτιμά τον κίνδυνο για μια γυναίκα να φέρει μια μετάλλαξη BRCA1 ή BRCA2 και τον κίνδυνο ανάπτυξης καρκίνου του μαστού (διηθητικού ή in situ) με την πάροδο του χρόνου, λαμβάνοντας υπόψη κληρονομικούς, ορμονικούς και παθολογικούς παράγοντες κινδύνου. Οι παράγοντες κινδύνου που χρησιμοποιούνται περιλαμβάνουν: την ηλικία, τον BMI, την ηλικία κατά την εμμηναρχή, την ηλικία κατά τον πρώτο τοκετό, τον συνολικό αριθμό τοκετών, η ηλικία στην έναρξη της εμμηνόπαυσης, τη χρήση ορμονών, αποτελέσματα από βιοψίες μαστού, την παρουσία υπερπλασίας, άτυπης υπερπλασίας (πόρων ή λοβίου) ή LCIS, και κληρονομικές πληροφορίες, που αφορούν στους συγγενείς πρώτου και δεύτερου βαθμού με καρκίνο του μαστού ή/και καρκίνο των ωοθηκών και περιλαμβάνουν πληροφορίες όπως είναι η ηλικίας εμφάνισης του καρκίνου και η παρουσία αμφοτερόπλευρου καρκίνου του μαστού [158]. Το μοντέλο λαμβάνει επίσης υπόψη την εβραϊκή καταγωγή Ashkenazi. Το μοντέλο είναι ένα γενετικό μοντέλο δύο εστιών, το οποίο λαμβάνει υπόψη (1) μία εστία που περιέχει πληροφορίες σχετικά με τα γονίδια BRCA (που περιέχει είτε ένα «φυσιολογικό» [159] αλληλόμορφο, ένα BRCA1 ή ένα αλληλόμορφο BRCA2) και (2) μία εστία για ένα υποθετικό χαμηλής διεισδυτικότητας γονίδιο ευαισθησίας που αντιπροσωπεύει την αύξηση του σχετικού κινδύνου εμφάνισης καρκίνου του μαστού. Το μοντέλο TC επιδεικνύει καλύτερη βαθμονόμηση και ακρίβεια πρόβλεψης από άλλα μοντέλα, όπως το Gail, όταν εφαρμόζεται σε πληθυσμούς υψηλού κινδύνου με βάση το οικογενειακό ιστορικό [158].

4.2 Ευφυή Συστήματα Υποστήριξης Κλινικής Απόφασης για τη Διάγνωση και την Κλινική Διαχείριση του Καρκίνου του Μαστού

4.2.1 Ιατρικά Δεδομένα

Ήδη από τις αρχές της δεκαετίας του '90, με τους αλγόριθμους τεχνητής νοημοσύνης και μηχανικής μάθησης ευρέως διαδεδομένους, έγινε σαφές ότι η επίδραση και η χρησιμότητα τέτοιων μεθόδων στον τομέα της ιατρικής θα ήταν καθοριστικές. Η εκμετάλλευση αυτών των συστημάτων αποδεικνύεται γρήγορα γόνιμη στην διαδικασία κατάστροφης ενός πλάνου αντιμετώπισης από τους ειδικούς της υγείας [160]. Αποφάσεις που μέχρι πρότινος λαμβάνονταν βασιζόμενες στην προσωπική εμπειρία και κρίση του ιατρικού προσωπικού, εμπειρία που απαιτεί χρόνια για να εξελιχθεί και κρίση που λόγω της ανθρώπινης φύσης είναι αναμενόμενο να κάνει σφάλματα, μελετώνται πια μέσω υπολογιστικών συστημάτων που έχουν την δυνατότητα της αποθήκευσης πληθώρας δεδομένων και ιατρικών ηλεκτρονικών φακέλων. Με την εξέλιξη των τεχνολογιών, οι μηχανές επεκτάθηκαν και στην ανάλυση του καρκινικού γονιδιώματος, στην ανακάλυψη νέων βιοδεικτών, ενώ οδήγησαν στη δημιουργία εξειδικευμένων φαρμάκων και θεραπειών μέσω της πιο ολοκληρωμένης κατανόησης του καρκίνου [161]. Συγκεκριμένα για το καρκίνο του μαστού, υπήρξαν αμέσως πολλές εφαρμογές υψηλής απόδοσης που στόχευαν είτε στην ανίχνευση κακοήθους όγκου, είτε στην πρόγνωση της εξέλιξής του, αλλά και στην αποτελεσματικότερη αντιμετώπιση της ασθένειας. Επιπλέον, ο συνδυασμός της τεχνητής νοημοσύνης με μεθόδους μηχανικής μάθησης κατέστησαν δυνατή τόσο την αποφυγή περιττών διαδικασιών, είτε χειρουργικών είτε εξεταστικών, όσο και το να μην γίνεται υπέρ του δέοντος χρήση φαρμακευτικών σκευασμάτων που πιθανώς να έβλαπταν την ασθενή.

Πριν αναλύσουμε τις τεχνικές που έχουν εφαρμοστεί, πρέπει να αναφερθούμε με λεπτομέρεια σε ορισμένα από τα σύνολα δεδομένων που έχουν χρησιμοποιηθεί, προκειμένου να γίνουν οι μελέτες. Αυτά ποικίλουν σε κλινικά, παθολογικά, αποτελέσματα βιοψιών, εικόνες μαστογραφίας, γονιδιακά ή ακόμα και κοινωνικοοικονομικά.

Δύο από τα πιο διαδεδομένα σετ, που συναντάμε ιδιαίτερα συχνά στη βιβλιογραφία, είναι τα Wisconsin Breast Cancer Datasets, το Original (WOBC) και το Diagnostic (WDBC) [162]. Το πρώτο, WOBC, περιλαμβάνει 699 δείγματα από κλινικές περιπτώσεις που είχαν συλλεχθεί από το 1989 έως το 1991 και δωρίστηκε στις 15 Ιουλίου του 1992, στο UCI Machine Learning Repository, το οποίο είναι μία συλλογή βάσεων δεδομένων, ελεύθερα προσβάσιμη από την επιστημονική κοινότητα. Τα δείγματα προέρχονται από βιοψία ανθρώπινου ιστού μαστού, με χρήση λεπτής βελόνας αναρρόφησης (Fine Needle Aspirates – FNA). Επίσης, έχουν μια Ταυτότητα (ID) και δέκα χαρακτηριστικά που λαμβάνουν ακέραιες τιμές μεταξύ του 1 και 10, με το 1 να ανταποκρίνεται σε κανονική κατάσταση και το 10 στην πιο ασυνήθιστη. Το δέκατο

χαρακτηριστικό αναφέρεται στην Κλάση (Class) του δείγματος, δηλαδή αν ο ιστός ανήκει σε καλοήγη-ένδειξη 2- ή κακοήγη-ένδειξη 4- όγκο. Τα υπόλοιπα εννέα είναι Πάχος Όγκου (Clump Thickness), Ομοιομορφία Κυτταρικού Μεγέθους (Uniformity of Cell Size), Ομοιομορφία Κυτταρικού Σχήματος (Uniformity of Cell Shape), Προσκόλληση Κυτταρικών Ορίων (Marginal Adhesion), Μέγεθος Μονών Επιθηλιακών Κυττάρων (Single Epithelial Cell Size), Γυμνοί Πυρήνες (Bare Nuclei), Χρωματίνη (Bland Chromatin), Κανονικοί Πυρήνες (Normal Nucleoli) και Μίτωση (Mitosis) [162]. Από αυτές τις περιπτώσεις 458 είναι καλοήθεις, ενώ οι υπόλοιπες 241 κακοήθεις. Ωστόσο, σε 16 εντοπίζονται ελλείψεις τιμών στο πεδίο «Γυμνοί Πυρήνες» και στις περισσότερες περιπτώσεις παραλείπονται από το σετ δεδομένων [163].

Το δεύτερο, WDBC σετ, δόθηκε στο UCI την 1^η Νοεμβρίου του 1995 και αποτελείται από 569 περιστατικά, κάθε ένα από τα οποία αντιστοιχεί σε ψηφιακή εικόνα FNA βιοψίας μάζας σε μαστό. Διατίθενται 32 γνωρίσματα, αναμεσά τους η Ταυτότητα (ID) και η Διάγνωση (Diagnosis). Τα υπόλοιπα 30 έχουν υπολογιστεί για κάθε κυτταρικό πυρήνα των συλλεγμένων ιστών και περιλαμβάνουν τη Μέση Τιμή, την Τυπική Απόκλιση και τη «χειρότερη» ή Μέγιστη (μέσος όρος των τριών μεγαλύτερων μετρήσεων) τιμή για τις εξής 10 πληροφορίες: την Ακτίνα (Radius), την Υφή (Texture), την Περίμετρο (Perimeter), το Εμβαδό (Area), την Ομαλότητα (Smoothness), την Συμπάγεια (Compactness), την Κοιλότητα (Concavity), τα Κοίλα Σημεία (Concave Points), τη Συμμετρία (Symmetry) και την Κλασματική Διάσταση (Fractal Dimension). Ως καλοήθεις κρίνονται 357 περιπτώσεις και ως κακοήθεις 212, ενώ δεν υπάρχουν χαμένες μετρήσεις ή τιμές σε κανένα πεδίο.

Δεδομένα για έρευνες σχετικά με τον καρκίνο του μαστού έχουν, επιπλέον, αποκτηθεί από το πρόγραμμα SEER (Surveillance, Epidemiology and End Results) του Εθνικού Ινστιτούτου Καρκίνου των ΗΠΑ, το οποίο συγκεντρώνει στατιστικά στοιχεία επιβίωσης και πληροφορίες για περιπτώσεις ασθενών με καρκίνο. Οι περιπτώσεις αυτές χρονολογούνται από το 1973 έως το 2014 και είναι δημόσια διαθέσιμες. Οι παράγοντες που συμπεριλαμβάνονται σε αυτό το σετ φτάνουν τους 133 και εκτείνονται από κοινωνικούς, όπως περιοχή ασθενούς, εθνικότητα και ηλικία, μέχρι παθολογικούς σχετικά με τον όγκο, το στάδιο μετάστασης στους λεμφαδένες, την ορμονική κατάσταση [164]. Κάθε έρευνα, όπως θα δούμε στη συνέχεια, επιλέγει τα χαρακτηριστικά που επιθυμεί για να συμπεριλάβει ως είσοδο.

Άλλο ένα σετ που θα συναντήσουμε και περιέχει κλινικοπαθολογικά χαρακτηριστικά ασθενών είναι το METABRIC (Molecular Taxonomy of Breast Cancer International Consortium) του Τμήματος Μοριακής Ογκολογίας των ΗΠΑ. Στόχος της συλλογής των εν λόγω δεδομένων είναι ο εντοπισμός και η ταξινόμηση καρκινικών όγκων σε περαιτέρω υποκατηγορίες βάσει των μοριακών και γονιδιακών δομών με απώτερο στόχο τη λήψη βέλτιστης απόφασης για θεραπεία. Ωστόσο, εδώ ενσωματώνονται και στοιχεία του γενετικού προφίλ των ασθενών, δηλαδή πληροφορίες για το γονιδίωμα.

4.2.2 Ευφυή Συστήματα Υποστήριξης Κλινικής Απόφασης για τη Διάγνωση και την Κλινική Διαχείριση του Καρκίνου του Μαστού

Η πρόοδος στους τομείς της επιστήμης των υπολογιστών και της Τεχνητής Νοημοσύνης οδήγησε στην ανάπτυξη συστημάτων για την υποστήριξη της κλινικής απόφασης όσον αφορά στη διάγνωση του καρκίνου με βάση εξατομικευμένα δεδομένα ασθενών [166], [167]. Τα CDSS δημιουργήθηκαν με σκοπό τη διαχείριση της βιοϊατρικής γνώσης προκειμένου να βελτιωθεί η κλινική πρακτική και να ενισχυθούν με περαιτέρω γνώση οι αποφάσεις που αφορούν στην Υγεία των ανθρώπων. Τα εργαλεία των εν λόγω συστημάτων αποτελούν τα μαθηματικά μοντέλα βασισμένα στην Πιθανοθεωρία, οι τεχνικές επεξεργασίας βιοϊατρικών σημάτων, οι τεχνικές ανάλυσης ιατρικών δεδομένων και γενικά οι αλγόριθμοι της τεχνητής νοημοσύνης. Τα CDSS καλύπτουν ένα ευρύ φάσμα εφαρμογών, από την υποστήριξη της διάγνωσης, μέχρι την εκτίμηση της πιθανότητας εμφάνισης διαφόρων ασθενειών ή την αποτελεσματικότητα εναλλακτικών θεραπευτικών σχεδίων. Προκειμένου να δημιουργηθούν τα αντίστοιχα μοντέλα, τα CDSS χρησιμοποιούν δεδομένα που αφορούν σε προσωπικά ιατρικά δεδομένα ασθενών (δημογραφικά και επιδημιολογικά δεδομένα), παράγοντες κινδύνου, περιβαλλοντικούς ή γενετικούς παράγοντες και πληροφορίες για την αποτελεσματικότητα των υπάρχοντων θεραπευτικών σχημάτων.

Τα CDSS χρησιμοποιούν εργαλεία και μεθόδους τόσο της Στατιστικής επιστήμης, όπως είναι η ανάλυση παλινδρόμησης, όσο και της Τεχνητής Νοημοσύνης αφού πολλές φορές βασίζονται σε αλγορίθμους μηχανικής μάθησης, τεχνητά νευρωνικά δίκτυα (Artificial Neural Networks - ANN) και τεχνικές αναγνώρισης προτύπων [168]. Τα εργαλεία αυτά χρησιμοποιούνται προκειμένου να εξαχθούν «κρυφές» πληροφορίες, κρίσιμης σημασίας για την κλινική πρακτική, από μεγάλα σύνολα δεδομένων. Οι πολύπλοκοι αλγόριθμοι που χρησιμοποιούν τα CDSS έχουν τη δυνατότητα να συνδυάζουν μη γραμμικά συσχετιζόμενα χαρακτηριστικά, όπως είναι επιδημιολογικά δεδομένα, ιατρικό και οικογενειακό ιστορικό, αποτελέσματα εξετάσεων, βιοδείκτες και πληροφορίες που αφορούν στο εκάστοτε θεραπευτικό σχήμα όπως είναι το είδος και η δοσολογία των φαρμάκων. Με αυτόν τον τρόπο, τα CDSS μπορούν να παρέχουν στους κλινικούς γιατρούς ειδικές συστάσεις για τη διαχείριση των ασθενών βοηθώντας στη λήψη χρήσιμων αποφάσεων. Οι αποφάσεις αυτές αφορούν τόσο τη διάγνωση όσο και το σχεδιασμό της θεραπείας.

Όσον αφορά στον καρκίνο του μαστού, ένα CDSS θα μπορούσε να υποστηρίξει τους κλινικούς γιατρούς σε αποφάσεις αναφορικά με τη διάγνωση καθώς δύναται να παρέχει πληροφορίες σχετικά με τον κίνδυνο ανάπτυξης της νόσου σε συγκεκριμένες περιπτώσεις ασθενών. Επίσης, θα μπορούσε να προτείνει το καταλληλότερο θεραπευτικό σχήμα ανάλογα την περίπτωση [169], [170]. Πληθώρα δημοσιευμένων μελετών αναφέρονται σε CDSS για την υποστήριξη της διαχείρισης των ασθενών με καρκίνο του μαστού που βασίζονται στην ανάλυση εικόνας της Μαστογραφίας, του Υπερήχου μαστών και της Μαγνητικής

Τομογραφίας, εφαρμόζοντας πληθώρα αλγορίθμων μηχανικής μάθησης (π.χ. KNN, SVM, Binary – LG) προκειμένου να αυξήσουν τη διαγνωστική απόδοση και να εντοπίζουν με μεγάλη ακρίβεια, ευαισθησία και ειδικότητα τους κακοήθεις και καλοήθεις όγκους [171], [172]. Αρκετές ερευνητικές ομάδες αξιοποίησαν τα προαναφερθέντα μοντέλα εκτίμησης κινδύνου, όπως οι G.F Stark et al. [173], που εμπλούτισαν το μοντέλο Gail [124], με εύκολα προσβάσιμα προσωπικά δεδομένα υγείας προκειμένου να προβλέψουν τον πενταετή κίνδυνο καρκίνου του μαστού με τη βοήθεια των εργαλείων μηχανικής μάθησης. Πολλές μελέτες χρησιμοποιούν την, ανοιχτή στο επιστημονικό κοινό, βάση δεδομένων Wisconsin Breast Cancer Database (WBCD)[162] [174], [160]. Η εν λόγω βάση δεδομένων περιλαμβάνει 32 χαρακτηριστικά που έχουν εξαχθεί από την ψηφιοποιημένη εικόνα της εξέτασης FNA της μάζας του μαστού για 699 περιστατικά. Τα χαρακτηριστικά περιγράφουν τη μορφολογία των πυρήνων των κυττάρων που υπάρχουν στην εικόνα. Πολλές μελέτες συγκρίνουν διαφορετικούς αλγορίθμους μηχανικής μάθησης όπως είναι η γραμμική Παλινδρόμηση (Linear Regression - LR), ο Multilayer Perceptron (MLP), ο αλγόριθμος κοντινότερου γείτονα (k -Nearest Neighbor - kNN), η Softmax Regression, ο Support Vector Machine (SVM), ο Naive Bayes (NB) και άλλοι, πάνω στο σύνολο δεδομένων WBCD αξιολογώντας την ακρίβεια της εκάστοτε ταξινόμησης αλλά και τις τιμές ευαισθησίας και ειδικότητας αναφορικά με τη διάγνωση της νόσου [175]. Άλλες ερευνητικές ομάδες επέλεξαν τον συνδυασμό αλγορίθμων μηχανικής μάθησης (Ensemble Learning) προκειμένου να επιτύχουν καλύτερη προγνωστική και διαγνωστική απόδοση απ' ότι θα μπορούσε να αποκτηθεί από οποιονδήποτε μεμονωμένο αλγόριθμο μηχανικής μάθησης και πάλι αξιοποιώντας το σύνολο δεδομένων WBCD [176]. Ο Asti και οι συνεργάτες του [177], χρησιμοποίησαν τέσσερις ταξινομητές: SVM, NB, C4.5 (Δέντρο αποφάσεων) και k-NN οι οποίοι συγκαταλέγονται στους πιο σημαντικούς αλγόριθμους εξόρυξης δεδομένων για την ερευνητική κοινότητα και μεταξύ των κορυφαίων 10 αλγορίθμων εξόρυξης δεδομένων γενικά. Σκοπός τους ήταν να αξιολογήσουν την αποτελεσματικότητα των αλγορίθμων όσον αφορά στην ακρίβεια, στην εγκυρότητα, στην ευαισθησία και στην ειδικότητα έχοντας ως βάση δεδομένων την WBCD. Ο κύριος στόχος της ερευνητικής ομάδας ήταν να εκτιμηθεί η ορθότητα στην ταξινόμηση δεδομένων σε σχέση με την αποδοτικότητα και την αποτελεσματικότητα κάθε αλγορίθμου όσον αφορά την ακρίβεια, την ευαισθησία και την ειδικότητα. Τα πειραματικά αποτελέσματα έδειξαν ότι ο SVM ταξινομητής δίνει την υψηλότερη ακρίβεια (97,13%) με το χαμηλότερο ποσοστό σφάλματος. Επιπρόσθετα, η μελέτη των D.A Omondiaigbe et al. [178], πρότεινε μια υβριδική λύση για τη διάγνωση του καρκίνου του μαστού μειώνοντας σημαντικά τον αριθμό των χαρακτηριστικά της WBCD, με τη χρήση της ανάλυσης γραμμικής διάκρισης (LDA) και, στη συνέχεια, εφαρμόζοντας τον SVM ταξινομητή στο νέο μειωμένο σύνολο χαρακτηριστικών. Η προτεινόμενη προσέγγιση πέτυχε ακρίβεια 98,82%, ευαισθησία 98,41%, ειδικότητα 99,07%. Μια ακόμη ερευνητική ομάδα, των Alireza Osareh et al.[179], στη μελέτη τους χρησιμοποίησαν 2 δημόσια διαθέσιμα σύνολα δεδομένων. Το πρώτο σύνολο περιείχε δείγματα αναρροφήσεων με λεπτή βελόνα από ογκιδία του μαστού (FNA) και το δεύτερο σύνολο

δεδομένων αποτελούνταν από γονιδιακές μικροσυστοιχίες. Η καλύτερη συνολική ακρίβεια διάγνωσης που πέτυχαν μέσω SVM ταξινομητή ήταν 98,80% και 96,33% αντίστοιχα για τα δυο σύνολα δεδομένων.

Μια ακόμα ερευνητική ομάδα, των M.M.Rahman et al. [180], χρησιμοποίησε γνωστούς αλγορίθμους μηχανικής μάθησης προκειμένου αφενώς να βρει τους πιο σημαντικούς βιοδείκτες του καρκίνου του μαστού και αφετέρου να βελτιώσει τη διαγνωστική απόδοση και να συμβάλει στην έγκαιρη ανίχνευση του καρκίνου του μαστού. Στην μελέτη τους χρησιμοποίησαν ένα σύνολο δεδομένων που περιείχε ανθρωπομετρικά και κλινικά χαρακτηριστικά, και συγκεκριμένα το δείκτη μάζας σώματος, την ηλικία και τα επίπεδα γλυκόζης, MCP-1 χημειοκίνης, ρεζιστίνης και ινσουλίνης. Η ακρίβεια ταξινόμησης που επετεύχθη ήταν 93,9% (95% CI: 93,2-94,6%), η ευαισθησία 95,1% (95% CI: 94,4-95,8%) και η ειδικότητα 94,0% (95% CI: 93,3-94,7%). Αντίστοιχα, οι M.F Aslan et al. [181], αξιοποίησαν τα αποτελέσματα των αιματολογικών εξετάσεων των γυναικών χρησιμοποιώντας διάφορες μεθόδους μηχανικής μάθησης προκειμένου να επιτύχουν την έγκαιρη διάγνωση του καρκίνου του μαστού. Κάποιες από τις βασικές μεθόδους που χρησιμοποιήθηκαν ήταν τα ANN, SVM και ο αλγόριθμος K-Nearest Neighbor (k-NN). Το σύνολο δεδομένων που χρησιμοποιήθηκε ελήφθη από τη βιβλιοθήκη UCI [182]. Σε αυτό το σύνολο δεδομένων χρησιμοποιήθηκαν η ηλικία, ο BMI, η γλυκόζη, η ινσουλίνη, η η λεπτίνη, η αδιπονεκτίνη, η ρεζιστίνη, MCP1 χημειοκίνη κ.ά.

Με κατεύθυνση τη βελτίωση της διαχείρισης ασθενών με καρκίνο του μαστού, το ευρωπαϊκό έργο DESIREE (έργο του HORIZON 2020) στοχεύει στην ανάπτυξη διαδικτυακών υπηρεσιών για τη διαχείριση του πρωτοπαθούς καρκίνου του μαστού από κέντρα και μονάδες μαστού. Στα πλαίσια του έργου περιγράφεται ένα σύστημα υποστήριξης αποφάσεων με βάση τις υπάρχουσες κατευθυντήριες γραμμές (guidelines) για τη διαχείριση της νόσου (Guidelines Based Decision Support System - GL-DSS). Για τη δημιουργία του εν λόγω GL - DSS χρησιμοποιήθηκαν διάφορες κλινικές πρακτικές οδηγίες για τον καρκίνο του μαστού (clinical practice guidelines - CPGs) για την παροχή ειδικών εξατομικευμένων συστάσεων για τον ασθενή. Ο στόχος ήταν ο συνδυασμός και η διεύρυνση των υπάρχουσών κλινικών συστάσεων με τη βοήθεια του GL - DSS [183]. Για τη δημιουργία του GL – DSS χρησιμοποιήθηκαν τα εργαλεία του σημασιολογικού ιστού, οπότε τα δεδομένα και σημασιολογικά μοντέλα ενσωματώθηκαν σε ένα μοναδικό οντολογικό πλαίσιο [184].

Διάφορα συστήματα υποστήριξης κλινικής απόφασης θα ήταν χρήσιμα σε κρίσιμες αποφάσεις όπως είναι η επιλογή των εξετάσεων διαλογής στις οποίες πρέπει να υποβληθεί μια γυναίκα, η ερμηνεία των αποτελεσμάτων των εξετάσεων και η απόφαση για το αν μια γυναίκα πρέπει να παραπεμφθεί ή όχι σε κάποια επεμβατική ή δαπανηρή διαγνωστική εξέταση. Οι αποφάσεις αυτές καθοδηγούνται επί του παρόντος από κατευθυντήριες γραμμές κλινικής πρακτικής (Clinical Practice Guidelines - CPGs), οι οποίες

δεν είναι βέβαιο ότι λειτουργούν αποτελεσματικά στον γενικό πληθυσμό δεδομένης της ανομοιομορφία του. Η ερευνητική ομάδα των A.M. Alaa et al. [185], πρότεινε ένα εξατομικευμένο σύστημα υποστήριξης που προσαρμόζεται στα χαρακτηριστικά των ατόμων, προκειμένου να διασφαλιστεί ότι οι σωστές εξετάσεις συνιστώνται στη σωστή γυναίκα. Το σύστημά τους, ConfidentCare δημιουργεί ομάδες ασθενών με παρόμοια χαρακτηριστικά (π.χ. απεικονιστικά χαρακτηριστικά, ηλικία, πυκνότητα μαστού, οικογενειακό ιστορικό κ.λπ.), συνδυάζει τα στοιχεία του προσωπικού τους ιατρικού φακέλου με τα αποτελέσματα της Μαστογραφίας, του Υπερήχου μαστών και της MRI, και με τη βοήθεια των αλγορίθμων μηχανικής μάθησης προτείνει τις κατάλληλες πολιτικές πρακτικές ελέγχου που θα ήταν αποτελεσματικότερες για τη διαχείριση κάθε ομάδας ατόμων ξεχωριστά.

Με παρόμοια μέθοδο, αυτή της ομαδοποίησης των ασθενών βάσει των χαρακτηριστικών τους, λειτούργησε η ομάδα των M Alamelumangai [186], προτείνοντας ένα σύστημα υποστήριξης κλινικής απόφασης που συνδυάζει τα στοιχεία του αναλυτικού ατομικού ιστορικού (ηλικία, πυκνότητα μαστών, οικογενειακό ιστορικό, προηγούμενες βιοψίες) με τα αποτελέσματα της Μαστογραφίας, Υπερήχου μαστών και MRI (βαθμολογία κατά BI-RADS και για τις 3 εξετάσεις).

4.3 Συζήτηση – Συμπεράσματα

Μετά από τη εκτεταμένη μελέτη της διεθνούς βιβλιογραφίας, συγκεντρώσαμε τα συναφέστερα, σε σχέση με το προτεινόμενο, μοντέλα και παρουσιάσαμε τη διαγνωστική απόδοσή τους σε επίπεδα ακρίβειας της εκτίμησης του κινδύνου ανάπτυξης καρκίνου του μαστού. Στον Πίνακα 6, παραθέσαμε τα ονόματα των μοντέλων ή των ερευνητικών ομάδων που τα ανέπτυξαν, τις μεθόδους που χρησιμοποιήθηκαν για την ανάπτυξή τους και το είδος των χαρακτηριστικών που συμπεριλαμβάνονταν στο εκάστοτε μοντέλο.

Πίνακας 6: Διαγνωστική απόδοση των στατιστικών μοντέλων εκτίμησης κινδύνου του καρκίνου του μαστού και των ευφυών συστημάτων υποβοηθούμενης διάγνωσης

<i>Μέθοδοι</i>	<i>Χαρακτηριστικά/ Παράγοντες που συμπεριλαμβάνονται στο μοντέλο</i>	<i>Ακρίβεια εκτίμησης (%) κινδύνου για ανάπτυξη καρκίνου του μαστού</i>
<i>Στατιστικά μοντέλα</i>		
(1989) Gail 1 & 2 [124]	Ορμονικοί (ηλικία κατά την εμμηνарχή, ηλικία κατά τον πρώτο τοκετό), παθολογικοί (αριθμός προηγούμενων βιοψιών μαστού, προσωπικό ιστορικό	55% και 61% αντίστοιχα

		άτυπης υπερπλασίας), κληρομονικοί (αριθμός των θηλυκών συγγενών πρώτου βαθμού με καρκίνο του μαστού)			
(2006) Barlow/BCSC [127]	LR	<p><u>Για προεμμηνοπαυσιακές γυναίκες:</u> ηλικία, πυκνότητα μαστού, αριθμός των συγγενών πρώτου βαθμού με καρκίνο του μαστού και προηγούμενες επεμβάσεις στο μαστό</p> <p><u>Για μετεμμηνοπαυσιακές γυναίκες:</u> δημογραφικοί παράγοντες, BMI, ορμονικοί παράγοντες (ηλικία της γέννησης του πρώτου παιδιού, χρήση ορμονικής θεραπείας, τύπος εμμηνόπαυσης), κληρονομικοί παράγοντες (αριθμός πρώτου βαθμού συγγενών με καρκίνο του μαστού), μαστογραφικά ευρήματα (πυκνότητα του μαστού, προηγούμενα μαστογραφικά ευρήματα), και προηγούμενες επεμβάσεις στο μαστό</p>	63,1%	και	62,4% αντίστοιχα
(2021) [131]	<i>Tice</i> LR	Πυκνότητα μαστού, ιστορικό των καρκίνων του μαστού σε συγγενείς πρώτου βαθμού, προσωπικό ιστορικό βιοψίας μαστού, φυλή και εθνικότητα.	69 %		
(2015) BC [133]	<i>BBD</i> – LR	Ιστολογικά ευρήματα (συνολική ιστολογική εντύπωση, αριθμός εστιών άτυπης υπερπλασίας, έκταση της λοβιακής αναδίπλωσης, παρουσία ακτινωτής ουλής, σκληρυντική αδένωση),			

		ορμονικοί παράγοντες, κληρονομικοί παράγοντες	
(2013) Rosner Colditz [137]	LR	Αναπαραγωγικοί παράγοντες κινδύνου, προηγούμενη καλοήθης νόσος του μαστού, οικογενειακό ιστορικό καρκίνου του μαστού για συγγενείς πρώτου βαθμού, χρήση ορμονών μετά την εμμηνόπαυση, BMI, ύψος και κατανάλωση αλκοόλ	60 %
(2014) Tworoger [138]	LR	Χαρακτηριστικά μοντέλου Rosner – Colditz με προσθήκη επιπέδων οιστρώνης, προλακτίνης και τεστοστερόνης	62 %
(2011) Petracci [139]	LR	Ηλικία εμμηναρχής, ηλικία πρώτου τοκετού, αριθμός των προηγούμενων βιοψιών μαστού, αριθμός των συγγενών πρώτου βαθμού με καρκίνο του μαστού, ενδεχόμενη επαγγελματική σωματική δραστηριότητα, μορφωτικό επίπεδο, κατανάλωση αλκοόλ, σωματική δραστηριότητα στον ελεύθερο χρόνο και BMI	62 % για γυναίκες κάτω των 50 ετών και 57% για τις γυναίκες 50 ετών και άνω
(1993) Claus [141]	LR	Αριθμός συγγενών πρώτου (όχι δεύτερου) βαθμού με καρκίνο του μαστού ή καρκίνο των ωοθηκών (λαμβάνοντας υπόψη την ηλικία τους κατά τη διάγνωση)	71,6 %
<i>Ευφυή μοντέλα</i>			

(2019) <i>G. F Stark et al.</i> [173]	LR, LDA (Linear Discriminant Analysis), NB, DT, SVM, NN	1) Δεδομένα από το μοντέλο Gail [124] 2) Δεδομένα από το μοντέλο Gail [124] και 8 ακόμη παράγοντες κινδύνου (ηλικία εμμηνόπαυσης, δείκτης τρέχουσας χρήσης ορμονών, αριθμός ετών χρήσης ορμονών, BMI, έτη καπνίσματος πακέτων τσιγάρων, έτη χρήσης αντισύλληψης, αριθμός γεννήσεων και δείκτης προσωπικού προηγούμενου ιστορικού καρκίνου)	1) 56,7% (καλύτερη απόδοση από NN) 2) 61,3% (καλύτερη απόδοση από LR & LDA)
(2017) <i>M. Nilashi et. al</i> [174]	EM, PCA, CART και fuzzy rule-based	WDBC και σύνολα δεδομένων μαστογραφικής μάζας	93,2% και 94,1% αντίστοιχα
(2019) <i>E. A. Bayrak</i> [175]	Sequential Minimal Optimization (SMO), LibSVM, MLP, Voted Perceptron	WDBC	96,9%, 95,7, 95,4 και 90,9 % αντίστοιχα
(2018) <i>D. C Banerjee</i> [176]	Ensemble ML	WDBC	95,6 %
(2016) <i>H.Asri</i> [177]	SVM5, NB, C4.5 και k-NN	WDBC	97,1% (SVM)

(2019)	<i>D.A Omondiagbe et al.</i> [178]	LDA + SVM	WDBC	98,8%
(2010)	<i>A. Osareh et al.</i> [179]	SVM	Δείγματα από ογκιδία του μαστού (FNA) και δεδομένα από γονιδιακές μικροσυστοιχίες	98,8%
(2021)	<i>M.M.Rahman et al.</i> [180]	SVM	Ανθρωπομετρικά και κλινικά χαρακτηριστικά (BMI, ηλικία, τα επίπεδα γλυκόζης, MCP-1 χημειοκίνης, ρεζιστίνης και ινσουλίνης)	93,9%
(2018)	<i>M.F Aslan et al.</i> [181]	ANN, Extreme Learning Machine (ELM), Knn, SVM	Ηλικία, BMI, αιματολογικές εξετάσεις (γλυκόζη, η ινσουλίνη, λεπτίνη, αδιπνονεκτίνη, ρεζιστίνη, MCP1, χημειοκίνη)	79,4%3, 80%. 77,5% και 73,5 % αντίστοιχα
(2016)	<i>A.M. Alaa et al.</i> [185]	ML algorithms	Απεικονιστικά χαρακτηριστικά από μαστογραφία, υπέρηχο και MRI, ηλικία, πυκνότητα μαστού, οικογενειακό ιστορικό κ.λπ	94,4%

Όπως προκύπτει από τον παραπάνω πίνακα (Πίνακας 6), διαπιστώθηκε ότι οι μέθοδοι που έχουν αναπτυχθεί για την εκτίμηση του κινδύνου και την έγκαιρη διάγνωση του καρκίνου του μαστού, χρησιμοποιούν γενετικές πληροφορίες, δεδομένα από επεμβατικές εξετάσεις, όπως η CNB και η FNA, δεδομένα από την ανάλυση εικόνων μαστογραφίας, US και MRI, πληροφορίες από εξέταση αιματολογικού υλικού και χαρακτηριστικά που λαμβάνονται από τα προσωπικά ιατρικά αρχεία των γυναικών.

Οι μέθοδοι που χρησιμοποίησαν τα δεδομένα της βάσης WDBC ή άλλα δεδομένα από εξετάσεις FNA, είχαν σαν αποτέλεσμα πολύ υψηλές τιμές ακρίβειας της εκτίμησης του κινδύνου του καρκίνου του μαστού

παρόλα αυτά πρόκειται για αποτελέσματα ανάλυσης του βιολογικού υλικού της βιοψίας άρα προέρχονται από επεμβατική διαδικασία. Άλλες ερευνητικές ομάδες που συμπεριέλαβαν την εξέταση της MRI στα συστήματά τους επίσης παρουσίασαν αυξημένη ακρίβεια διάγνωσης, ωστόσο η λύση του προτείνουν δεν είναι οικονομικά αποδοτική.

Εν κατακλείδι, εξ όσων γνωρίζουμε, δεν έχει δημοσιευθεί μέχρι σήμερα στη βιβλιογραφία ένα ισοδύναμο ευφύς σύστημα με αυτό που προτείνεται στην παρούσα μελέτη, δηλαδή ένα σύστημα που συνδυάζει μη επεμβατικά και οικονομικά αποδοτικά διαγνωστικά εργαλεία μαζί με τα χαρακτηριστικά του ατομικού και οικογενειακού ιατρικού ιστορικού, όπως παρουσιάζεται παρακάτω, επιτυγχάνοντας τόσο υψηλές τιμές διαγνωστικής απόδοσης.

Συγκεκριμένα, στο πλαίσιο της παρούσας διατριβής διερευνάται ο ρόλος ενός νέου έξυπνου βιοπληροφορικού CDSS το οποίο συνδυάζει τα αποτελέσματα δύο διαφορετικών διαγνωστικών εξετάσεων, της μαστογραφίας και του υπέρηχου μαστού μαζί με τα χαρακτηριστικά του προσωπικού ιατρικού φακέλου ασθενούς όπως παρουσιάζονται αναλυτικά στο κεφάλαιο που ακολουθεί (Κεφάλαιο 5). Όπως αναλύθηκε στο υποκεφάλαιο 3.2, η μαστογραφία και ο US είναι οι συνήθεις και λιγότερο δαπανηρές διαγνωστικές εξετάσεις ωστόσο έχουν χαμηλή σχετικά ευαισθησία και ειδικότητα στη διάγνωση, όπως διαφαίνεται στην εκτεταμένη ανασκόπηση του 2022 [4], στην οποία διαπιστώθηκε ότι οι τιμές για τη μαστογραφία ήταν 54,5% (εύρος 27% - 86,8%) και 85,5% (εύρος 62,9% - 98,8%) αντίστοιχα (17 μελέτες) και για το υπερηχογράφημα 67,2% (εύρος 26,9% - 87,5%) και 76,8% (εύρος 18,8% - 96,9%) αντίστοιχα (14 μελέτες).

Σκοπός της παρούσας διατριβής είναι ο συνδυασμός των αποτελεσμάτων/χαρακτηριστικών των διαγνωστικών εξετάσεων της μαστογραφίας και του US μαστών μαζί με δεδομένα ατομικού ιστορικού προκειμένου να αυξηθεί η διαγνωστική απόδοση και να προσδιοριστεί ο πραγματικός κίνδυνος ανάπτυξης καρκίνου του μαστού με τα διαγνωστικά εργαλεία που διαθέτουν οι κλινικοί γιατροί και βρίσκονται στην καθημερινή κλινική πρακτική. Τα εν λόγω διαγνωστικά μέσα είναι οικονομικά για το σύστημα υγείας καθώς αποτελούν τις βασικές εξετάσεις που πραγματοποιούνται για την διάγνωση του καρκίνου του μαστού. Από το προτεινόμενο CDSS επιλέχθηκε να μη συμπεριληφθεί η MRI που αποτελεί μια δαπανηρή εξέταση και είναι χρήσιμη σε συγκεκριμένες περιπτώσεις μαστών [187], η γενετική εξέταση, η οποία είναι επίσης δαπανηρή και αφορά σε μικρό ποσοστό γυναικών και όλες οι μορφές βιοψίας καθώς στόχος του CDSS είναι να είναι μη επεμβατικό αποκλείοντας ακόμα και την FNA που είναι ελάχιστα επεμβατική αλλά όπως παρουσιάζεται στο υποκεφάλαιο 3.1.2 ενέχει κάποιους κινδύνους.

Κεφάλαιο 5^ο

Παρουσίαση του Ιατρικού Προβλήματος

5.1 Ιατρικά Δεδομένα: Περιγραφή και Οργάνωση

Ένα κρίσιμης σημασίας χαρακτηριστικό των ερευνητικών μελετών αποτελεί ο προσδιορισμός του πληθυσμού - στόχου. Προκειμένου να γίνει η κατάλληλη επιλογή του πληθυσμού - στόχου θα πρέπει να προσδιοριστούν τα χαρακτηριστικά της ομάδας ενδιαφέροντος με βάση τις μεταβλητές της μελέτης ή την ποιοτική εμπειρία, ώστε τα αποτελέσματα της μελέτης να μπορούν να εξαχθούν με ακρίβεια και να μεταφερθούν πίσω στον πληθυσμό ενδιαφέροντος. Πολύ συχνά τα αρχικά χαρακτηριστικά είναι σχετικά εύκολο να εφαρμοστούν, ωστόσο, απαιτούνται περισσότερες λεπτομέρειες για τον πλήρη ορισμό και τη λειτουργικότητα των ορίων του πληθυσμού-στόχου. Ένα κρίσιμο βήμα για τη οριστικοποίηση των κριτηρίων είναι ο σαφής ορισμός των βασικών όρων και η διασφάλιση ότι οι ορισμοί που χρησιμοποιούνται είναι σαφείς [188].

Όταν το πλαίσιο δειγματοληψίας έχει δύο ή περισσότερες ομάδες ενδιαφέροντος, ή στρώματα, τότε η στρωματοποιημένη δειγματοληψία (Stratified sampling) συμβάλλει στη διασφάλιση της αναλογικής εκπροσώπησης κάθε ομάδας ενδιαφέροντος. Αρχικά, πρέπει να καθορίζεται η αναλογία κάθε ομάδας ενδιαφέροντος στο πλαίσιο δειγματοληψίας. Στη στρωματοποιημένη δειγματοληψία καθορίζονται οι αναλογίες και στη συνέχεια γίνεται απλή τυχαία δειγματοληψία σε κάθε στρώμα μέχρι να επιτευχθεί το ανάλογο μέγεθος υποδείγματος. Η απλή τυχαία δειγματοληψία σε κάθε στρώμα επιτρέπει την πιθανολογική εκπροσώπηση άλλων (δευτερευόντων) χαρακτηριστικών που μπορεί να μην ενδιαφέρουν τη μελέτη, ενώ η στρωματοποιημένη δειγματοληψία εξασφαλίζει την πιθανολογική εκπροσώπηση των ομάδων ενδιαφέροντος [188].

Βασιζόμενοι στην επιστημονικά τεκμηριωμένη θεωρία δειγματοληψίας για ερευνητικές μελέτες [188], επιλέξαμε τη στρωματοποιημένη δειγματοληψία των γυναικών που απευθύνονται στο Κέντρο Μαστού του Γενικού Νοσοκομείου «Ελενα Βενιζέλου», προκειμένου να προσδιοριστεί κατάλληλα ο πληθυσμός - στόχος για τη διεξαγωγή της έρευνάς μας.

Προκειμένου να λάβουμε πρόσβαση στα ιατρικά δεδομένα των γυναικών κατατέθηκε σχετικό αίτημα στο Διοικητικό και Επιστημονικό συμβούλιο του νοσοκομείου. Στο αίτημα παρουσιάστηκε αναλυτικά το θέμα

της διδακτορικής διατριβής. Το Διοικητικό και Επιστημονικό συμβούλιο αποδέχθηκε το αίτημα και επετράπη η πρόσβαση, ανάλυση και επεξεργασία των δεδομένων για ερευνητικούς σκοπούς, μετά από τη σχετική δέσμευση για την κρυπτογράφηση και προστασία των ευαίσθητων ιατρικών δεδομένων των ασθενών. Όλα τα ιατρικά δεδομένα (αποτελέσματα διαγνωστικών εξετάσεων, παθολογοανατομικές εκθέσεις και αναλυτικό προσωπικό και οικογενειακό ιστορικό) καταχωρήθηκαν και αποθηκεύτηκαν ανώνυμα προκειμένου να υποβληθούν σε επεξεργασία.

Η μελέτη βασίστηκε σε έναν πληθυσμό γυναικών οι οποίες απευθύνθηκαν στο Κέντρο Μαστού προκειμένου να υποβληθούν σε κλινικές και διαγνωστικές εξετάσεις για καρκίνο του μαστού (πληθυσμός – στόχος). Τα δεδομένα των γυναικών ήταν σε ξεχωριστούς φακέλους με αναγνωριστικό Αριθμό Ιστορικού για κάθε περίπτωση. Ο φάκελος αυτός περιελάμβανε μια φόρμα με πληροφορίες που αφορούν στα προσωπικά στοιχεία της ασθενούς αλλά και στο οικογενειακό ιστορικό της. Συγκεκριμένα η φόρμα του ατομικού ιατρικού φακέλου έχει τα εξής πεδία:

- | | |
|---------------------------------------------------------------------------------|-------------------------------------------------|
| ✚ Ημερομηνία καταγραφής της φόρμας | ✚ Ρυθμός έμμηνου ρύσης |
| ✚ Αριθμός Ιστορικού | ✚ Εγκυμοσύνες |
| ✚ Ασφαλιστικό Ταμείο ασθενούς | ✚ Τέκνα |
| ✚ Ονοματεπώνυμο ασθενούς | ✚ Αποβολές |
| ✚ Ημερομηνία Γέννησης | ✚ Εκτρώσεις |
| ✚ Τόπος γέννησης | ✚ Ηλικία 1 ^{ου} και τελευταίου τοκετού |
| ✚ Τόπος διαμονής | ✚ Θηλασμός - Διάρκεια Θηλασμού |
| ✚ Βάρος και Ύψος | ✚ Ορμόνες - είδος και διάρκεια |
| ✚ Γενική κατάσταση ασθενούς (αφορά σε παρούσα κλινική εικόνα και λήψη φαρμάκων) | ✚ Αντισυλληπτικά - είδος και διάρκεια |
| ✚ Προηγηθείσες παθήσεις μαστού | ✚ Ομάδα αίματος |
| ✚ Ατομικό αναμνηστικό (Παθήσεις - Εγχειρήσεις) | ✚ Επάγγελμα - μορφωτικό επίπεδο |
| ✚ Οικογενειακό αναμνηστικό | ✚ Κάπνισμα - ποσότητα και διάρκεια |
| ✚ Έναρξη έμμηνου ρύσης | ✚ Αλκοόλ - διάρκεια |
| ✚ Τελευταία ημερομηνία έμμηνου ρύσης ή έναρξη Εμμηνόπαυσης | ✚ Συμπτώματα αναφερόμενα από τον ασθενή |
| | ✚ Μέγεθος στήθους |

Στον φάκελο κάθε ασθενούς υπήρχαν επίσης τα αποτελέσματα από τις διαγνωστικές εξετάσεις στις οποίες είχε υποβληθεί η ασθενής ανάλογα με την περίπτωση. Στις περισσότερες περιπτώσεις υπήρχε το αποτέλεσμα της μαστογραφίας (MAMMO) και του υπερήχου μαστών (US) ενώ σε κάποιες περιπτώσεις συμπεριλαμβανόταν το αποτέλεσμα της μαγνητικής τομογραφίας μαστών (MRI). Υπήρχαν φάκελοι με συγκεντρωμένες όλες τις πληροφορίες και τις περαιτέρω εξετάσεις στις οποίες υποβάλλεται η ασθενής πριν το χειρουργείο μαστού (αιματολογικές εξετάσεις, ακτινογραφία θώρακος, αξονική κοιλίας, αξονική οστών κ.τ.λ.). Υπήρχαν άλλοι φάκελοι με μοναδικό στοιχείο τη προαναφερθείσα φόρμα ατομικών και οικογενειακών στοιχείων (ατομικός ιατρικός φάκελος) και μια μαστογραφία αναφοράς καθώς επρόκειτο για την πρώτη «αναγνωριστική» επίσκεψη των γυναικών αυτών στο Κέντρο Μαστού. Στην παρούσα διδακτορική διατριβή επιλέξαμε να χρησιμοποιήσουμε εκείνους τους φακέλους των ασθενών που συμπεριελάμβαναν τις πληροφορίες που επρόκειτο να χρησιμοποιηθούν για τη διεξαγωγή της μελέτης, και συγκεκριμένα όλοι οι φάκελοι που αξιοποιήθηκαν περιλάμβαναν τον ατομικό ιατρικό φάκελο, τη μαστογραφία και τον υπέρηχο της εκάστοτε περίπτωσης.

Ποιες περιπτώσεις γυναικών συμπεριλήφθηκαν στη Διδακτορική Διατριβή

Στη διδακτορική διατριβή συμπεριελήφθησαν 489 γυναίκες εκ των οποίων οι 287 είχαν παθολογοανατομική έκθεση υπέρ κακοήθειας και οι 202 υπέρ καλοήθειας. Σε όλες οι περιπτώσεις οι γυναίκες υποβλήθηκαν σε χειρουργική επέμβαση μαστού, είτε σε συντηρητική επέμβαση μαστού (Ουγκεκτομή) είτε σε Μαστεκτομή. Όπως αναφέρθηκε και προηγουμένως, ο υπό μελέτη πληθυσμός αφορά πληθυσμό παραπομπής σε χειρουργική επέμβαση μαστού (πληθυσμός διαλογής - στόχος). Από τα 489 περιστατικά του υπό μελέτη πληθυσμού, κάποια παραπέμφθηκαν για χειρουργική επέμβαση λόγω ύποπτων ευρημάτων στην μαστογραφία, ενώ τα υπόλοιπα που είτε παρουσίασαν χαμηλότερου κινδύνου μαστογραφία είτε βρίσκονταν διαγνωστικά στη λεγόμενη «γκρίζα ζώνη», παραπέμφθηκαν διότι οι χειρουργοί μαστού έκριναν ότι έχρηζαν περαιτέρω διερεύνησης λόγω κλινικών ευρημάτων που διαπίστωσαν οι ίδιοι.

Προκειμένου να καταλήξουμε στο μέγεθος του υπό μελέτη πληθυσμού αλλά και στην αναλογία «αρνητικών» και «θετικών» περιστατικών, ανέτρεξαμε στο αρχείο των χειρουργικών επεμβάσεων του νοσοκομείου, όπου καταγράφονται όλες οι επεμβάσεις και τα αποτελέσματα των παθολογοανατομικών εκθέσεών τους και πραγματοποιήσαμε σωματοποιημένη δειγματοληψία. Σύμφωνα με το αρχείο των χειρουργείων, η αναλογία θετικών και αρνητικών αποτελεσμάτων ήταν περίπου 60% και 40% αντίστοιχα, ποσοστά που διαπιστώθηκαν σχετικά «σταθερά» κάθε χρόνο. Επομένως, ο υπό μελέτη πληθυσμός αποτελεί αντιπροσωπευτικό δείγμα ενός πληθυσμού παραπομπής σε χειρουργική επέμβαση (πληθυσμός διαλογής) εξασφαλίζοντας την πιθανολογική εκπροσώπηση των ομάδων ενδιαφέροντος. Η βάση αυτή

χρησιμοποιήθηκε τόσο για την ανάλυση των δεδομένων όσο και για την κατασκευή των ευφών αλγορίθμων και μοντέλων που παρουσιάζονται στην παρούσα διατριβή. Ως εκ τούτου, στα επόμενα κεφάλαια, όπου αναφερόμαστε σε δεδομένα, αναφερόμαστε στα περιστατικά αυτά.

Έχοντας τις 489 παθολογοανατομικές εκθέσεις που προέκυψαν μετά τη χειρουργική επέμβαση, η ιστολογική διάγνωση κάθε περιστατικού χρησιμοποιήθηκε ως το «Gold Standard» της παρούσας μελέτης. Ανάλογα με το αποτέλεσμα της ιστολογικής διάγνωσης, τα περιστατικά ταξινομήθηκαν στις 2 ακόλουθες κατηγορίες:

- I. Παθολογοανατομική έκθεση υπέρ Καλοήθειας (202/489)
- II. Παθολογοανατομική έκθεση υπέρ Κακοήθειας (287/489)

Το εύλογο ερώτημα που προκύπτει είναι ο λόγος που οι γυναίκες, οι οποίες αποδείχθηκε (μέσω της παθολογοανατομικής τους έκθεσης) ότι είχαν καλοήθη ευρήματα υπεβλήθησαν σε χειρουργική επέμβαση. Η απάντηση είναι σύνθετη και πολυπαραγοντική. Όπως έχει γίνει σαφές από τα προηγούμενα κεφάλαια, η μαστογραφία θεωρείται η εξέταση αναφοράς από τους χειρουργούς μαστού [3]. Στις περιπτώσεις που η μαστογραφία έχει σαφή εικόνα υπέρ κακοήθειας (π.χ. BI-RADS 4+) ο χειρουργός μαστού προβαίνει σε χειρουργική επέμβαση όπως προτρέπουν οι αντίστοιχες κατευθυντήριες γραμμές. Το πρόβλημα προκύπτει, όπως εξηγήσαμε και στο Κεφάλαιο 3.1.1, όταν το αποτέλεσμα της μαστογραφίας ανήκει στη λεγόμενη «γκρίζα ζώνη» (π.χ. BI-RADS 0 & 3), μην επιτρέποντας στον ειδικό να επιτελέσει τη διάγνωση με σιγουριά ώστε να αποφύγει όλους τους πιθανούς κινδύνους για την γυναίκα. Σε αυτές τις περιπτώσεις ο/η χειρουργός μαστού καλείται να διερευνήσει, μέσω της επέμβασης, την πιθανότητα ύπαρξης υποβόσκουσας κακοήθειας. Όπως διαπιστώνεται από τα αποτελέσματα τις παρούσας έρευνας, δεν είναι λίγες οι περιπτώσεις που τα αποτελέσματα της μαστογραφίας συνηγορούν ψευδώς υπέρ καλοήθειας. Συγκεκριμένα, παρατηρήθηκαν ακόμα και περιπτώσεις με BI-RADS 1 & 2 στη μαστογραφία για τις οποίες, λόγω άλλων παραγόντων (προσωπικό/οικογενειακό ιστορικό, αποτελέσματα εξετάσεων), οι χειρουργοί οδηγήθηκαν σε χειρουργική επέμβαση διερεύνησης εξαιτίας της υποψίας καρκίνου. Αυτές οι διαγνωστικές ελλείψεις της μαστογραφίας δημιούργησαν την ανάγκη για υποβοήθηση και βελτίωση της διαχείρισης των γυναικών αυτών.

Για την ανάλυση των δεδομένων και την ανάπτυξη των αλγορίθμων και των ευφών μοντέλων που παρουσιάζονται στην παρούσα διατριβή, από το συνολικό πλήθος των διαθέσιμων περιστατικών που υπήρχαν στη βάση δεδομένων, εξαιρέθηκαν περιστατικά που είχαν σημαντικό αριθμό ελλιπών δεδομένων σε βασικές εξετάσεις, όπως παραδείγματος χάριν περιπτώσεις όπου η μαστογραφία ή ο υπέρηχος δε συμπεριελάμβαναν βασικές πληροφορίες για τη μορφολογία της σκίασης ή την κατάσταση των μαστών.

Είναι σημαντικό να τονιστεί το γεγονός ότι στην παρούσα διδακτορική διατριβή αξιοποιήσαμε τα δεδομένα που συλλέγονται κατά την καθημερινή κλινική πράξη μέσω του ατομικού ιατρικού φακέλου και των βασικών διαγνωστικών εξετάσεων που συλλέγονται σε ένα Κέντρο Μαστού. Τα χαρακτηριστικά που συλλέχθηκαν, χρησιμοποιήθηκαν και αξιολογήθηκαν, στο προτεινόμενο CDSS, αντιπροσωπεύουν το μεγαλύτερο μέρος των παραγόντων κινδύνου που αναλύθηκαν διεξοδικά στο υποκεφάλαιο 2.6 και σύμφωνα με τις γνώσεις μας δεν έχει δημοσιευθεί μέχρι σήμερα κάποιο αντίστοιχο CDSS που να συνδυάζει τόσους πολλούς παράγοντες κινδύνου ανάπτυξης και εντοπισμού της νόσου αξιοποιώντας δεδομένα που συλλέγονται στην καθημερινή κλινική πράξη. Στη συνέχεια παρουσιάζονται αναλυτικά οι λόγοι επιλογής των συγκεκριμένων χαρακτηριστικών με βάση τη σημασία τους στην ανάπτυξη του καρκίνου του μαστού αλλά και τη συμβολή τους για την έγκαιρη διάγνωση καρκινικών αλλοιώσεων (π.χ. μορφολογικά χαρακτηριστικά αλλοιώσεων και μαστών), όπως αυτά προέκυψαν με τα την εκτενή μελέτη της διεθνούς βιβλιογραφίας.

5.2 Προεπεξεργασία των Ιατρικών Δεδομένων

Το πρώτο βήμα για την επεξεργασία των ιατρικών δεδομένων ήταν η ψηφιοποίηση των εγγράφων που συμπεριλαμβάνονταν στους φακέλους των ασθενών σε έντυπη μορφή. Τα αρχεία που ψηφιοποιήθηκαν ήταν η φόρμα με τις πληροφορίες για το ατομικό και οικογενειακό ιστορικό των ασθενών και το έντυπο αποτέλεσμα της μαστογραφίας και του υπερήχου μαστών.

Προκειμένου να είναι δυνατή η ανάλυση και επεξεργασία των επιδημιολογικών αλλά και ιατρικών δεδομένων πραγματοποιήθηκε πολύ σχολαστική εξαγωγή ποιοτικής πληροφορίας από τις εξετάσεις (μαστογραφία και υπέρηχος μαστών) καθώς οι εξετάσεις περιείχαν την πληροφορία σε μορφή κειμένου. Το επόμενο στάδιο ήταν ο μετασχηματισμός όλης της πληροφορίας σε μαθηματική πληροφορία, ήτοι η κωδικοποίηση της πληροφορίας σε αριθμούς.

Για κάθε ένα περιστατικό του υπό μελέτη πληθυσμού δημιουργήθηκε ένα διάνυσμα χαρακτηριστικών το οποίο αποτελείται από 35 μεταβλητές. Οι μεταβλητές αυτές προκύπτουν από τα αποτελέσματα της προαναφερθείσας φόρμας και των δύο διαγνωστικών δοκιμασιών (μαστογραφία και υπέρηχος μαστών). Οι 35 μεταβλητές παρουσιάζονται αναλυτικά στον Πίνακα 7. Οι μεταβλητές αυτές χαρακτηρίζουν κάθε περιστατικό (εκφράζουν τα αποτελέσματα των εξετάσεων και των προσωπικών πληροφοριών κάθε ασθενούς) και επομένως αποτελούν τις μεταβλητές εισόδου (μεταβλητές διανύσματος χαρακτηριστικών) των μοντέλων που πρόκειται να παρουσιαστούν στις παρακάτω υποενότητες. Είναι σημαντικό να τονιστεί ότι οι μεταβλητές είναι αριθμητικές, δυαδικές και κατηγορικές, επομένως απαιτείται η ανάλυση και επεξεργασία των δεδομένων χρησιμοποιώντας κατάλληλους αλγορίθμους που να μπορούν να χειριστούν επιτυχώς τους εν λόγω τύπους δεδομένων ταυτόχρονα όπως αυτοί παρουσιάζονται παρακάτω.

Η επιλογή των χαρακτηριστικών του διανύσματος εισόδου των μοντέλων βασίστηκε σε ανασκόπηση της διεθνούς βιβλιογραφίας για τους τους σημαντικότερους και καθοριστικότερους παράγοντες κινδύνου για τον καρκίνο του μαστού. Αρχικά, η ηλικία της γυναίκας φαίνεται να επηρεάζει σημαντικά το επίπεδο κινδύνου εμφάνισης καρκίνου του μαστού σύμφωνα με τη διεθνή επιστημονική και ιατρική κοινότητα [34], [35], [37]. Επιπλέον, οι αναπαραγωγικοί παράγοντες, συμπεριλαμβανομένης της ηλικίας έναρξης της εμμηνου ρύσεως, του αριθμού των τοκετών, του αριθμού των αμβλώσεων ή των αποβολών [43], της ηλικίας του πρώτου τοκετού [41], της ηλικίας του τελευταίου τοκετού [40], του θηλασμού, της χρήσης αντισυλληπτικών χαπιών από το στόμα, της ηλικίας έναρξης της εμμηνόπαυσης και της πιθανής εμμηνόπαυσιακής ορμονοθεραπείας, έχει αποδειχθεί, σε μεγάλο αριθμό μελετών, ότι διαδραματίζουν σημαντικό ρόλο στην εκτίμηση του επιπέδου κινδύνου εμφάνισης καρκίνου του μαστού [35],[38], [39]. Ένας άλλος παράγοντας που φαίνεται να επηρεάζει τον κίνδυνο εμφάνισης καρκίνου του μαστού φαίνεται να είναι ο BMI και η παχυσαρκία, καθώς έχουν δημοσιευτεί αρκετές μελέτες για το θέμα, αλλά με αντικρουόμενα αποτελέσματα [45], [53], [54], [57]. Στη συνέχεια, συγκεκριμένα χαρακτηριστικά του μαστού, όπως η πυκνότητά του [59], [60], [61] και το μέγεθός του [62], έχουν μελετηθεί εκτενώς για την επίδρασή τους στον κίνδυνο εμφάνισης καρκίνου του μαστού. Σύμφωνα με την ιατρική κοινότητα και τις σχετικές μελέτες, ορισμένοι από τους σημαντικότερους παράγοντες κινδύνου είναι οι διαταραχές του προσωπικού ιατρικού ιστορικού της γυναίκας (προηγούμενες γυναικολογικές διαταραχές, προηγούμενες παθήσεις του μαστού, διαταραχές του θυρεοειδούς [189], [190], υπέρταση [86], υψηλά επίπεδα χοληστερίνης [87]) και το οικογενειακό ιστορικό καρκίνου του μαστού και άλλων καρκίνων [66], [68], [69]. Ένας άλλος παράγοντας κινδύνου σχετίζεται με τον τρόπο ζωής της γυναίκας, συγκεκριμένα με την περίοδο καπνίσματος και την ποσότητα των τσιγάρων που κάπνιζε κατά τη διάρκεια αυτής της περιόδου [85].

Όσον αφορά στα χαρακτηριστικά των απεικονιστικών εξετάσεων που χρησιμοποιήθηκαν στο διάνυσμα εισόδου, βασιστήκαμε στις σημαντικότερες, με βάση τη διεθνή βιβλιογραφία, πληροφορίες που μπορούν να ληφθούν μέσω της εξέτασης MAMMO και της US μαστού. Συγκεκριμένα, ένα από τα σημαντικότερα σημεία που προσέχουν οι χειρουργοί μαστού κατά τη διαγνωστική διαδικασία με βάση τη MAMMO είναι εάν ο όγκος ή η βλάβη που ανιχνεύεται αποτελεί νέο εύρημα σε σχέση με προηγούμενη σχετική εξέταση. Ένα άλλο χαρακτηριστικό που τονίζεται κατά την εξήγηση των ευρημάτων MAMMO είναι εάν οι μαστοί έχουν μικροοζώδη μορφολογία και αποτιτανώσεις [104]. Η αρχιτεκτονική διαταραχή του μαστού, που ανιχνεύεται κατά την εξέταση της MAMMO, φαίνεται να αποτελεί πολύ σημαντικό παράγοντα κινδύνου για τα μαστογραφικά ευρήματα, όπως τονίζεται σε πολυάριθμες δημοσιευμένες μελέτες [104], [191]. Συμπληρωματικά με τα ευρήματα της Μαστογραφίας, ο US μπορεί να παράσχει κρίσιμες πληροφορίες σχετικά με το σχήμα του όγκου που ανιχνεύεται κατά την εξέταση, για παράδειγμα ωοειδές σχήμα, ακτινωτό σχήμα κ.λ.π. Το σχήμα του όγκου που ανιχνεύεται αποτελεί ισχυρή ένδειξη του ευρήματος του US, καθώς

όσο πιο «ομαλό» είναι το σχήμα του όγκου (π.χ. ωοειδές ή στρογγυλό) τόσο πιθανότερο είναι να πρόκειται για καλοήγη βλάβη, σε αντίθεση με τις ακτινωτές ή ατρακτοειδείς βλάβες που «ενοχοποιούνται» ως πιο πιθανές για κακοήγη βλάβη. Ένα άλλο σημαντικό χαρακτηριστικό που περιέχει η έκθεση του US, είναι ο χαρακτηρισμός των ορίων της εντοπισμένης βλάβης ως προς την ομαλότητά τους. Επίσης, η αγγείωση του όγκου καθώς και η κατάσταση των λεμφαδένων, περιγράφονται συχνά στην έκθεση του US, πληροφορίες που ενδιαφέρουν ιδιαίτερα τους κλινικούς ιατρούς, καθώς μπορεί να αποτελέσουν τα αποκαλυπτικά σημάδια μιας κακοήθους βλάβης [104].

Στον παρακάτω πίνακα παρουσιάζονται αναλυτικά οι μεταβλητές που προκύπτουν από τη φόρμα χαρακτηριστικών και τις διαγνωστικές εξετάσεις (MAMMO και US μαστού), καθώς και οι τιμές που λαμβάνουν αυτές οι μεταβλητές (Μεταβλητές Διανύσματος Χαρακτηριστικών).

Πίνακας 7: Μεταβλητές που προκύπτουν από τη φόρμα χαρακτηριστικών και τις διαγνωστικές εξετάσεις και χαρακτηρίζουν κάθε ασθενή (Μεταβλητές Διανύσματος Χαρακτηριστικών)

	Ονομασία Μεταβλητής	Περιγραφή	Εύρος τιμών
1	Acr	Βαθμολόγηση της πυκνότητας των μαστών με βάση την κλίμακα ACR	1= Οι μαστοί είναι σχεδόν εντελώς λιπώδεις 2= Υπάρχουν διασκορπισμένες περιοχές ινώδους και αδενικού ιστού 3= Οι μαστοί είναι ετερογενώς πυκνοί 4= Οι μαστοί είναι εξαιρετικά πυκνοί, γεγονός που μειώνει την ευαισθησία της μαστογραφίας
2	Age	Ηλικία ασθενούς	Αριθμητική μεταβλητή
3	Bmi	Δείκτης Μάζας Σώματος (Body Mass Index)= Βάρος/υψος ²	1= (< 18.5 kg/m ²) Επικίνδυνα χαμηλό σωματικό βάρος 2= (18.5-24.9 kg/m ²) Φυσιολογικό βάρος 3= (25-29.9 kg/m ²) Βάρος μεγαλύτερο του φυσιολογικού

		4= (>30 kg/m ²) Παχυσαρκία
4	Breast Size	Μέγεθος στήθους 0= Μικρό στήθος 1= Μεσαίο στήθος 2= Μεγάλο στήθος 3= Πολύ μεγάλο στήθος
5	Y of Smoking	Διάρκεια καπνίσματος σε έτη Αριθμητική μεταβλητή
6	Cigarettes_daily	Αριθμός τσιγάρων ανά ημέρα Αριθμητική μεταβλητή
7	Fem Diseases	Γυναικολογικές Παθήσεις 0= Δεν υπάρχουν γυναικολογικές παθήσεις στο ιστορικό της ασθενούς, 1= Πολύποδας μήτρας/ Κύστη ωοθηκών/ Πολυκυστικές ωοθήκες/ Εξωμήτριος κύηση, 2= Ινομώματα, 3= Καρκίνος τραχήλου/ Καρκίνος ενδομήτριου
8	Breast Dis. (Prev)	Προηγηθείσες Παθήσεις Μαστού 0= Δεν υπάρχουν προηγηθείσες παθήσεις μαστού στο ιστορικό της ασθενούς, 1= Κύστες, 2= Ινοαδένωμα 3= Μαστίτιδα 4= Αποτιτανώσεις 5= Καρκίνος μαστού
9	Hypertension (bn)	Η ασθενής πάσχει από Υπέρταση 0= Όχι 1= Ναι

10	Thyroid Problems (bn)	Η ασθενής αντιμετωπίζει προβλήματα με το Θυρεοειδή αδένα	0= Όχι 1= Ναι
11	High Cholesterol (bn)	Η ασθενής έχει υψηλή χοληστερόλη	0= Όχι 1= Ναι
12	Cancer hist {1d_S}	Μητέρα, Πατέρας, Τέκνα ή Αδελφός/ή με Καρκίνο του Μαστού	Αριθμητική μεταβλητή
13	Cancer hist {2d_S}	Θεός/α, Ξάδερφος/η, Γιαγιά/Παππούς, Ανιψιός/ά με Καρκίνο του Μαστού	Αριθμητική μεταβλητή
14	Cancer hist {1d_O}	Μητέρα, Πατέρας, Τέκνα ή Αδελφός/ή με άλλο τύπο Καρκίνου (εκτός του καρκίνου του μαστού)	Αριθμητική μεταβλητή
15	Cancer History {2d_O}	Θεός/α, Ξάδερφος/η, Γιαγιά/Παππούς, Ανιψιός/ά με άλλο τύπο Καρκίνου (εκτός του καρκίνου του μαστού)	Αριθμητική μεταβλητή
16	Onset of menstruation	Ηλικία έναρξης Έμμηνου ρύσης	Αριθμητική μεταβλητή
17	Menopause (bn)	Η ασθενής βρίσκεται σε Εμμηνόπαυση	0= Όχι 1= Ναι
18	Menopause	Ηλικία έναρξης Εμμηνόπαυσης	Αριθμητική μεταβλητή
19	Pregnancies	Εγκυμοσύνες (Αριθμός)	Αριθμητική μεταβλητή
20	Miscarriages	Εκτρώσεις (Αριθμός)	Αριθμητική μεταβλητή

21	Abortions	Αποβολές (Αριθμός)	Αριθμητική μεταβλητή
22	1st Preg_age	Ηλικία 1 ^{ου} τοκετού	Αριθμητική μεταβλητή
23	Last Preg_age	Ηλικία τελευταίου οκετού	Αριθμητική μεταβλητή
24	Breastfeeding	Συνολική διάρκεια θηλασμού τέκνων σε μήνες	0= Καθόλου 1= Έως και 12 μήνες 2= Πάνω από 12 μήνες
25	Contracting Pills	Διάρκεια λήψης αντισυλληπτικών χαπιών σε μήνες	Αριθμητική μεταβλητή
26	Other hormones [Y]	Διάρκεια λήψης ορμονών (εκτός θυροξίνης) σε έτη	Αριθμητική μεταβλητή
27	Thyroid hormones [Y]	Διάρκεια λήψης θυροξίνης σε έτη	Αριθμητική μεταβλητή
28	New Finding (bn)	Νέο εύρημα στη μαστογραφία σε σχέση με την προηγούμενη	0= Όχι 1= Ναι
29	Micro_Morphology (bn)	Μαστοί με μικροζώδη μορφολογία (κατά τη Μαστογραφία)	0= Όχι 1= Ναι
30	Arc_Disturbance (bn)	Διαταραχή της αρχιτεκτονικής των μαστών (από τη Μαστογραφία)	0= Όχι 1= Ναι
31	Tumor Shape	Σχήμα οζώδους σκίασης (από τον Υπέρηχο μαστών)	0= Δεν έχει σκίαση κάποιος μαστός 1= Στρογγυλό 2= Ωοειδές

		<p>3= Μακρολοβωτό</p> <p>4= Λοβωτό</p> <p>5= Ατρακτοειδές</p> <p>6= Μικρολοβωτό</p> <p>7= Ακανόνιστο</p> <p>8= Ακτινωτό</p>
32	Tumor Margins (bn)	<p>Όρια οζώδους σκίασης (από τον Υπέρηχο μαστών)</p> <p>0= Ανώμαλα όρια</p> <p>1= Ομαλά όρια</p>
33	Vascularity (bn)	<p>Αιμάτωση – Αγγείωση της οζώδους σκίασης (από τον Υπέρηχο μαστών)</p> <p>0= Φτωχή/ Όχι αυξημένη/ Φυσιολογική</p> <p>1= Αυξημένη</p>
34	Lymph Nodes	<p>Μορφολογία λεμφαδένων (από τον Υπέρηχο μαστών)</p> <p>0= Όχι παθολογική διόγκωση/ όχι παρουσία</p> <p>1= Καλοήθους μορφολογίας</p> <p>2= Ενδομαστικοί</p> <p>3= Μασχαλιαίοι λεμφαδένες</p> <p>4= Ύποπτη μορφολογία</p> <p>5= Παθολογικά διογκωμένοι</p>
35	Calcifications	<p>Μορφολογία αποτιτανώσεων (από τη Μαστογραφία)</p> <p>0= Όχι ύπαρξη</p> <p>1= Καλοήθειες</p> <p>2= Αδρές</p> <p>3= Όχι συρρέουσες</p> <p>4= Διάσπαρτες</p> <p>5= Συρρέουσες</p> <p>6= Ύποπτες</p>

Για κάθε ένα περιστατικό χρησιμοποιήθηκε ως μεταβλητή εξόδου (ή μεταβλητή στόχος) το αποτέλεσμα της ιστολογικής εξέτασης όπως αυτό προέκυψε από την παθολογοανατομική έκθεση, καθώς αυτό εκφράζει την πραγματική κατάσταση των μαστών κάθε γυναίκας (Gold Standard). Στην παρούσα διατριβή το αποτέλεσμα της ιστολογικής εξέτασης κωδικοποιήθηκε με τον εξής τρόπο:

Κλάση 1: Αρνητικό για Ca μαστού

Κλάση 2: Θετικό για Ca μαστού

5.3 Παρουσίαση του Ιατρικού Προβλήματος

Ο πληθυσμιακός έλεγχος με χρήση της μαστογραφίας έχει μειώσει δραματικά τα ποσοστά του καρκίνου του μαστού παγκοσμίως [192]. Προς το παρόν, σε πολλές χώρες η πρόληψη του καρκίνου του μαστού είναι βασισμένη σε συχνές και επαναλαμβανόμενες Μαστογραφίες, ακολουθούμενες από Υπέρηχο μαστών και εάν χρειαστεί από ιστολογική εξέταση στο βιολογικό υλικό της βιοψίας [3]. Ωστόσο, η αξιολόγηση των ευρημάτων της Μαστογραφίας αποτελεί μια δύσκολη διεργασία και μπορεί να επιτευχθεί μόνο από πολύ καλά εκπαιδευμένο ιατρικό προσωπικό. Ως εκ τούτου, η ερμηνεία τους επηρεάζεται από υποκειμενικούς παράγοντες και είναι επιρρεπής σε διαγνωστικά σφάλματα [193].

Η ένταξη της MRI αποτελεί μια διαγνωστική πρακτική που ενδείκνυται σε περιπτώσεις πολύ πυκνών μαστών (ACR 3-4) και αποτελεί μια υψηλού κόστους εξέταση για το σύστημα υγείας [187]. Στην παρούσα διδακτορική διατριβή παρουσιάζεται ένα μοντέλο πρόβλεψης κινδύνου εμφάνισης της νόσου του καρκίνου του μαστού που δεν συμπεριλαμβάνει τη διαγνωστική εξέταση της MRI. Ο λόγος που συμβαίνει αυτό είναι γιατί ο στόχος της δημιουργίας του εν λόγω μοντέλου είναι να αποτελέσει μια οικονομικά βιώσιμη και αποδοτική λύση για τα συστήματα υγείας όπου προτείνεται να εφαρμοστεί. Όσον αφορά στην FNA, η εν λόγω εξέταση είναι οικονομικά αποδοτική και ελάχιστα επεμβατική όμως παρόλα αυτά ενέχει ορισμένους κινδύνους όπως αυτοί καταγράφονται από τους J. T. Casaubon et al. [6] και αναλύονται στο υποκεφάλαιο 3.1.2. Επιπλέον, η εν λόγω ερευνητική ομάδα, έπειτα από ενδελεχή μελέτη, τονίζει την ανάγκη αξιολόγησης του αποτελέσματος από έμπειρο κυτταροπαθολόγο γεγονός που επηρεάζει την ακρίβεια της διάγνωσης. Ως εκ τούτου, παρά τα πλεονεκτήματα που εμφανίζει η μέθοδος, επιλέξαμε να μην την συμπεριλάβουμε στην εν λόγω μελέτη λόγω των ανωτέρω μειονεκτημάτων της ως προς τη διαγνωστική διαδικασία [6]. Σκοπός μας λοιπόν είναι η ανάπτυξη ενός συστήματος το οποίο να παρουσιάζει υψηλή διαγνωστική απόδοση, χωρίς όμως τη χρήση της MRI και της FNA, προκειμένου να είναι οικονομικά αποδοτικό και ταυτόχρονα μη επεμβατικό.

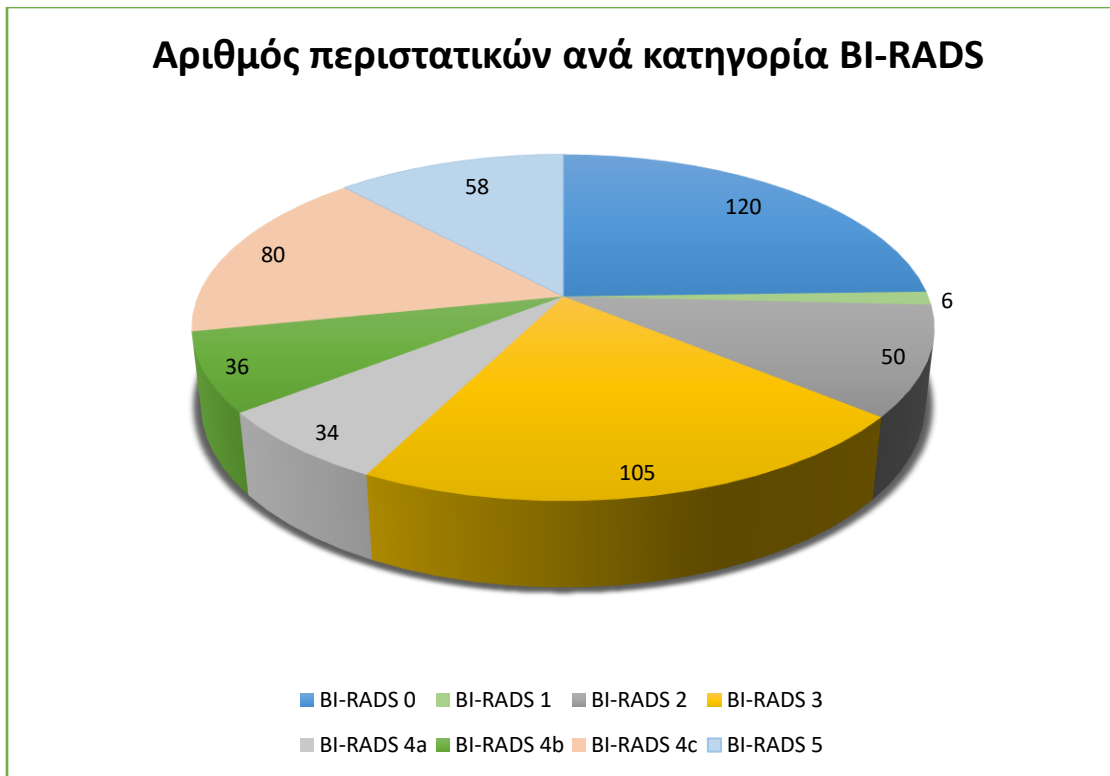
Σύμφωνα με τις ευρωπαϊκές οδηγίες και τα πρωτόκολλα για τη διαχείριση και διαλογή γυναικών με πιθανή ανάπτυξη καρκίνου του μαστού ως εξέταση αναφοράς θεωρείται η μαστογραφία. Στη συνέχεια ως

συμπληρωματικό διαγνωστικό εργαλείο χρησιμοποιείται ο υπέρηχος μαστών και σε πιο σπάνιες ειδικές περιπτώσεις εξαιρετικά πυκνών μαστών πραγματοποιείται και MRI [3]. Οι συστάσεις αυτές [3] αφορούν κυρίως τις γυναίκες με μέσο κίνδυνο για καρκίνο του μαστού και όχι αυξημένο κίνδυνο λόγω γενετικής προδιάθεσης (μεταλλάξεις στα BRCA1 και BRCA2 γονίδια ή λόγω φυλής / εθνικότητας [194]). Οι συστάσεις υποδηλώνουν επίσης ότι η χάραξη αποτελεσματικής πολιτικής για την αντιμετώπισή της ανάπτυξης καρκίνου του μαστού απαιτεί ουσιαστική συζήτηση και συμμετοχή από διάφορα ενδιαφερόμενα μέρη.

Για τον υπολογισμό των μέτρων απόδοσης της Μαστογραφίας αξιοποιήθηκε η βαθμονόμηση BI-RADS, που αναλύεται στο Κεφάλαιο 3.1.1 και αποτυπώνει το βαθμό επικινδυνότητας των μαστογραφικών ευρημάτων. Η κατάταξη με βάση την κλίμακα BI-RADS χρησιμοποιήθηκε ως συγκριτικό μέσο για να αξιολογηθεί η μαστογραφία ως διαγνωστικό εργαλείο εκτίμησης κινδύνου εμφάνισης καρκίνου του μαστού με βάση τα ευρήματα. Οι μαστογραφίες που συμπεριελήφθησαν στην παρούσα διδακτορική διατριβή ήταν αξιολογημένες με βάση την BI-RADS κλίμακα και ανήκαν σε κάποια από τις προαναφερθείσες κατηγορίες και υποκατηγορίες. Συγκεκριμένα, στον παρακάτω πίνακα (Πίνακας 8) παρουσιάζονται τα 489 περιστατικά σύμφωνα με την κατηγορία BI-RADS στην οποία ανήκουν:

Πίνακας 8: Αριθμός περιστατικών ανά κατηγορία σύμφωνα με τη βαθμονόμηση BI-RADS

	BI- RADS 0	BI- RADS 1	BI- RADS 2	BI- RADS 3	BI- RADS 4a	BI- RADS 4b	BI- RADS 4c	BI- RADS 5
Αριθμός Περιστατικών ανά κατηγορία	120	6	50	105	34	36	80	58



Εικόνα 16: Γραφική απεικόνιση περιστατικών ανά κατηγορία BI-RADS

Στη συνέχεια χρησιμοποιήθηκαν διαφορετικά «κατώφλια» θετικότητας για τον υπολογισμό της διαγνωστικής απόδοσης της μαστογραφίας.

Η χρήση διαφορετικών κατωφλίων θετικότητας για τη συγκεκριμένη εξέταση είναι μια απαραίτητη πρακτική καθώς η εξέταση αυτή λαμβάνει πολλαπλά επίπεδα αποτελεσμάτων. Είναι σημαντικό να τονιστεί ότι από την ομαδοποίηση εξαιρέθηκαν οι μαστογραφίες που ανήκαν στην κατηγορία BI-RADS 0, αφού λόγω του χαρακτηρισμού τους ως ανεπαρκούς εξέτασης (βλ. Κεφάλαιο 3.1.1.1) δε δύναται να καταταχθούν σε κάποιο από τα δυο προαναφερθέντα γκρουπ. Η ομαδοποίηση των BI-RADS κατηγοριών έγινε με χρήση διαφορετικών κατωφλίων θετικότητας, όπως προαναφέρθηκε, καθώς βασίστηκε στα θεωρητικά ποσοστά εκτίμησης κινδύνου που αναφέρονται στη βιβλιογραφία, στοχεύοντας στη διερεύνηση της αντιστοίχισης των εν λόγω ποσοστών στα πραγματικά δεδομένα της έρευνας.

Η κατηγοριοποίηση κατά BI-RADS είναι το μόνο μέτρο που έχουμε στη διάθεσή μας για την αξιολόγηση της διαγνωστικής απόδοσης της Μαστογραφίας, καθώς οι κατηγορίες της κλίμακας BI-RADS (εκτός της κατηγορίας 0 και 6) χαρακτηρίζονται από την πιθανότητα ύπαρξης καρκίνου του μαστού με βάση τα μαστογραφικά ευρήματα. Τα διαφορετικά κατώφλια θετικότητας χρησιμοποιήθηκαν προκειμένου να αποδείξουμε ότι η Μαστογραφία αποδίδει μεγάλο αριθμό κυρίως ψευδώς αρνητικών και λιγότερα ψευδώς θετικών αποτελεσμάτων σε όλες τις περιπτώσεις, όπως και να μετακινηθεί το κατώφλι θετικότητας.

Αναλυτικά όλες οι διαφορετικές δοκιμές και τα αποτελέσματα των ιστολογικών εξετάσεων σε σχέση με την πρόβλεψη (μέσω BI-RADS) της Μαστογραφίας, φαίνονται στους παρακάτω Πίνακες Σύγκρισης (βλ. Παράρτημα):

Πίνακας 9: Πίνακας Σύγκρισης 1 - Αποτελέσματα της Μαστογραφίας με κατώφλι θετικότητας (BI-RADS 4a+)

	Θετικό ιστολογικό αποτέλεσμα	Αρνητικό ιστολογικό αποτέλεσμα
BI-RADS 4a, 4b, 4c, 5	177	31
BI-RADS 1,2,3	50	111

Πίνακας 10: Πίνακας Σύγκρισης 2 - Αποτελέσματα της Μαστογραφίας με κατώφλι θετικότητας τη βαθμολογία (BI-RADS 4b+)

	Θετικό ιστολογικό αποτέλεσμα	Αρνητικό ιστολογικό αποτέλεσμα
BI-RADS 4b, 4c, 5	166	8
BI-RADS 1,2,3, 4a	61	134

Πίνακας 11: Πίνακας Σύγκρισης 3 - Αποτελέσματα της Μαστογραφίας με κατώφλι θετικότητας (BI-RADS 4c+)

	Θετικό ιστολογικό αποτέλεσμα	Αρνητικό ιστολογικό αποτέλεσμα
BI-RADS 4c, 5	136	2
BI-RADS 1,2,3, 4a, 4b	91	140

Παρά το γεγονός ότι η κάθε κατηγορία της κλίμακας BI-RADS περιλαμβάνει μια εκτίμηση κινδύνου, βλέπουμε με βάση τα δοθέντα δεδομένα ότι σε πολλές περιπτώσεις το αποτέλεσμα της ιστολογικής εξέτασης διαφέρει από το αναμενόμενο με βάση την κατηγορία στην οποία ανήκει η αντίστοιχη μαστογραφία. Συγκεκριμένα, όπως φαίνεται από τα αποτελέσματα των ιστολογικών εξετάσεων που έχουμε στη διάθεσή μας, και διαφαίνεται από τους πίνακες σύγκρισης που παρατίθενται (Πίνακας 9, 10 και 11), οι κατηγορίες BI-RADS 3 και BI-RADS 4, ενέχουν σε πολλές περιπτώσεις κίνδυνο, αφού δεν είναι λίγες οι φορές που ενώ αναμένεται αρνητικό αποτέλεσμα ιστολογικής από τους ειδικούς, η πραγματικότητα

τους φέρνει αντιμέτωπους με τη νόσο του καρκίνου. Μετά από την εκτεταμένη συγκεντρωτική μελέτη των N. Aristokli et al. [4], που παρουσιάστηκε στο υποκεφάλαιο 3.2.6, διαφαίνεται ότι η διαγνωστική ακρίβεια της μαστογραφίας με όρους ευαισθησίας και ειδικότητας κυμαίνεται σε ποσοστά κοντά στο 60% και 90% αντίστοιχα παρακινώντας την επιστημονική κοινότητα να αναζητήσει μεθόδους «ενίσχυσης» του διαγνωστικού εργαλείου της μαστογραφίας. Τα ποσοστά αυτά επισημαίνονται επίσης σε πληθώρα άλλων μελετών [195], [196], [197].

Προκειμένου να αξιολογηθεί η διαγνωστική απόδοση της Μαστογραφίας με βάση τα δεδομένα της διδακτορικής διατριβής υπολογίστηκε η ευαισθησία, η ειδικότητα, η θετική προγνωστική αξία (PPV), η αρνητική προγνωστική αξία (NPV) και ο Youden's index των μεθόδων (βλ. Παράρτημα Α), όσον αφορά στον εντοπισμό του καρκίνου του μαστού.

Πίνακας 12: Διαγνωστική απόδοση της μαστογραφίας για την ανίχνευση του καρκίνου του μαστού στον υπό μελέτη πληθυσμό

	Μαστογραφία (κατώφλι BI-RADS 4a)	Μαστογραφία (κατώφλι BI-RADS 4b)	Μαστογραφία (κατώφλι BI-RADS 4c)
<i>Accuracy (%)</i>	78,0	81,0	74,8
<i>Sensitivity (%)</i>	77,9	73,1	59,9
<i>Specificity (%)</i>	78,1	94,4	98,6
<i>PPV (%)</i>	88,5	95,4	98,5
<i>NPV (%)</i>	68,9	68,7	59
<i>Youden's index</i>	0,56	0,58	0,59
<i>F1 score</i>	0,82	0,83	0,74

Όπως βλέπουμε στον Πίνακα 12, στην παρούσα έρευνα διαπιστώνεται ότι όσο μετακινείται το κατώφλι θετικότητας πιο αυστηρά ως προς τις πιθανότητες που δίνονται στη βιβλιογραφία για τις κατηγορίες BI-RADS 0, τόσο μειώνεται η ευαισθησία και αυξάνεται η ειδικότητα της εξέτασης. Συγκεκριμένα η

ευαισθησία κυμαίνεται από 77,9 έως 59,9% και αντίστοιχα η ειδικότητα από 78,1 - 98,6 %, γεγονός που αποδεικνύει ότι στις περιπτώσεις όπου εντοπίζονται με μεγάλη ακρίβεια οι αρνητικές περιπτώσεις (κατώφλι BI-RADS 4c), χάνονται πολλά περιστατικά με κακοήθεια υποδεικνύοντας την επιτακτική ανάγκη υποβοήθησης της λήψης αποφάσεων εκ μέρους των κλινικών.

5.4 Διαχείριση Περιστατικών με Ελλιπή Δεδομένα

Σε πρακτικές εφαρμογές, για κάποια περιστατικά ενδεχομένως να λείπουν ορισμένες τιμές από συγκεκριμένα χαρακτηριστικά (δεδομένα με ελλιπείς τιμές). Συγκεκριμένα στη δική μας περίπτωση, εφόσον συμπεριλαμβάνουμε πραγματικά δεδομένα γυναικών που απευθύνθηκαν στο Κέντρο Μαστού έχοντας πραγματοποιήσει τις εξετάσεις τους σε διαφορετικά απεικονιστικά κέντρα, είναι ευνόητο ότι οι αναφορές της μαστογραφίας και του υπερήχου τους δεν περιείχαν ακριβώς τις ίδιες πληροφορίες. Επίσης λόγω της καταγραφής των δεδομένων του προσωπικού ιατρικού φακέλου από το προσωπικό του Κέντρου μαστού, αντιλαμβανόμαστε ότι παρουσιάστηκαν ορισμένες ελλείψεις σε πληροφορίες αφού ο ανθρώπινος παράγοντας είναι κάτι που επηρεάζει την εν λόγω διαδικασία. Οι περιπτώσεις αυτές είναι αρκετά συνηθισμένες σε ένα Κέντρο μαστού ή Νοσοκομείο και οδηγούν σε περιστατικά με ελλιπή αποτελέσματα εξετάσεων (ελλιπείς τιμές).

Το πρόβλημα των ελλιπών τιμών είναι ένα πρόβλημα το οποίο απασχολεί τις περισσότερες εφαρμογές μηχανικής μάθησης. Ως εκ τούτου, έχουν γίνει πάρα πολλές προσπάθειες για την αποτελεσματική αντιμετώπισή του [198]. [Το πρόβλημα των ελλιπών τιμών στη μηχανική μάθηση είναι διττό' αφορά αφενός την εκμάθηση μοντέλων με σύνολα δεδομένων που περιέχουν ελλιπείς τιμές, αφετέρου την ταξινόμηση προτύπων με ελλιπή δεδομένα. Έχουν προταθεί διάφορες τεχνικές για την αντιμετώπιση των δύο αυτών προβλημάτων. Οι περισσότερες από αυτές μπορούν να ομαδοποιηθούν στις παρακάτω κατηγορίες [198]:

- Τεχνικές εκτίμησης και συμπλήρωσης των ελλιπών τιμών χρησιμοποιώντας στατιστικές μεθόδους
- Τεχνικές εκτίμησης και συμπλήρωσης των ελλιπών τιμών χρησιμοποιώντας αλγόριθμους μηχανικής μάθησης
- Αλγόριθμοι μηχανικής μάθησης που διαχειρίζονται εγγενώς δεδομένα με ελλιπείς τιμές (χωρίς να πραγματοποιείται κάποια εκτίμηση και συμπλήρωση των ελλιπών τιμών).

Πέρα από τις ανωτέρω τεχνικές υπάρχει και η μέθοδος διαγραφής, σύμφωνα με την οποία τα δείγματα με ελλιπή δεδομένα διαγράφονται εντελώς και δεν χρησιμοποιούνται σε κανένα στάδιο. Στην περίπτωση που το πρόβλημα είναι η εκπαίδευση ενός ταξινομητή, είναι στην κρίση του σχεδιαστή εάν θα διαγράψει τα δείγματα με ελλιπείς τιμές ή όχι' εάν τα δείγματα αυτά είναι λίγα ίσως είναι προτιμητέο να διαγραφούν

παρά να εφαρμοστεί κάποια μέθοδος διαχείρισης ελλিপών τιμών, εάν όμως τα δείγματα αυτά αποτελούν το μεγαλύτερο μέρος του συνόλου δεδομένων τότε η μέθοδος διαγραφής δεν αποτελεί λύση. Γίνεται εύκολα αντιληπτό ότι η μέθοδος αυτή δεν αποτελεί λύση στην περίπτωση που επιθυμείται η ταξινόμηση ενός προτύπου με ελλιπείς τιμές [199].

Οι τεχνικές εκτίμησης και συμπλήρωσης δεδομένων (τεχνικές Καταλογισμού Δεδομένων – Data Imputation) αφορούν στην εκτίμηση και συμπλήρωση των ελλিপών τιμών σε αρχικό στάδιο (στάδιο προεπεξεργασίας δεδομένων), πριν πραγματοποιηθεί οποιαδήποτε άλλη διαδικασία, προκειμένου να δημιουργηθεί ένα πλήρες σετ δεδομένων. Με τον τρόπο αυτό, ο ταξινομητής (είτε κατά την εκπαίδευσή του, είτε κατά την ταξινόμηση) επεξεργάζεται πλήρη δεδομένα και δεν γνωρίζει ότι υπάρχουν ελλιπείς τιμές, καθώς αυτές έχουν αντικατασταθεί με τις εκτιμήσεις τους.

Όσον αφορά στις τεχνικές που χρησιμοποιούν στατιστικές μεθόδους για την εκτίμηση των ελλিপών τιμών, αυτές περιλαμβάνουν σχήματα όπως:

- Συμπλήρωση με μηδενικά
- Συμπλήρωση με την αδέσμευτη μέση τιμή η οποία έχει υπολογιστεί από τις υπόλοιπες τιμές που υπάρχουν για το χαρακτηριστικό
- Συμπλήρωση με τη δεσμευμένη μέση τιμή, αν υπάρχει η εκτίμηση της συνάρτησης πυκνότητας πιθανότητας για τις ελλιπείς τιμές, με δεδομένες τις υπόλοιπες τιμές που έχουν παρατηρηθεί

Στην παρούσα μελέτη επιλέξαμε ως καταλληλότερες τις τεχνικές εκτίμησης και συμπλήρωσης δεδομένων. Αρχικά ορίσαμε ένα κατώτατο κατώφλι προκειμένου να διαχωρίσουμε τα δεδομένα με υψηλό ποσοστό ελλিপών τιμών και να αποκλείσουμε κάποια από το να συμπεριληφθούν στη μελέτη, έτσι όταν πάνω από το 60% των τιμών μιας μεταβλητής έλειπαν, η μεταβλητή αυτή αποκλειόταν από τη μελέτη. Ακολουθώντας τον προαναφερθέντα κανόνα, αποκλείσαμε την ομάδα αίματος (74% ελλιπείς τιμές), την κατανάλωση αλκοόλ (70% ελλιπείς τιμές) και τη συμπάγεια του όγκου (64% ελλιπείς τιμές) που αφορούσε σε ένα μορφολογικό χαρακτηριστικό που παρείχε η έκθεση του υπερήχου. Για τα υπόλοιπα χαρακτηριστικά, αντικαταστήσαμε τις ελλείπουσες τιμές εφαρμόζοντας τη στρατηγική επαναληπτικού υπολογισμού (IterativeImputer) καθώς είναι μια εξελιγμένη προσέγγιση που περιλαμβάνει τον ορισμό ενός μοντέλου για την πρόβλεψη κάθε ελλείποντος χαρακτηριστικού ως συνάρτηση όλων των άλλων χαρακτηριστικών [199].

Τέλος η Python (έκδοση 3.6.13) χρησιμοποιήθηκε για την προεπεξεργασία των κλινικών μας δεδομένων όπως και για τη δημιουργία των αλγορίθμων Τεχνητής Νοημοσύνης και την ανάλυση των δεδομένων που θα παρουσιαστεί στη συνέχεια. Η διαθεσιμότητα πολυάριθμων βιβλιοθηκών και εργαλείων ανοικτού κώδικα καθιστούν την Python ιδανική επιλογή για την ανάπτυξη των μοντέλων.

Κεφάλαιο 6^ο

Αξιολόγηση και Κατάταξη των Χαρακτηριστικών του Διανύσματος Εισόδου με χρήση Μεθόδων Επιλογής Χαρακτηριστικών και Ανάπτυξη των Μοντέλων Πρόβλεψης/Ταξινόμησης

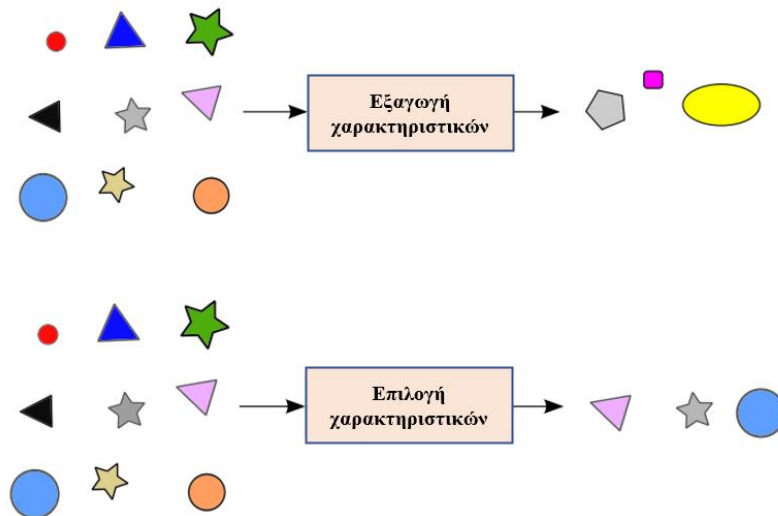
6.1 Επιλογή Χαρακτηριστικών στη Μηχανική Μάθηση

Στην εποχή των μεγάλων δεδομένων, τα πολλά δείγματα και χαρακτηριστικά έχουν αυξήσει σημαντικά τον υπολογιστικό φόρτο και έχουν προκαλέσει τη λεγόμενη κατάρα της «διαστασιμότητας», η οποία αναφέρεται σε διάφορα φαινόμενα που προκύπτουν κατά την ανάλυση δεδομένων σε χώρο υψηλών διαστάσεων. Η επιλογή χαρακτηριστικών είναι ένα από τα πιο αποτελεσματικά εργαλεία για τη μείωση της διάστασης και την αντιμετώπιση του παραπάνω προβλήματος και έχει εφαρμοστεί ευρέως σε πολλούς τομείς, π.χ. αναγνώριση προτύπων, εξόρυξη δεδομένων και μηχανική μάθηση, ως αποτελεσματική τεχνολογία προεπεξεργασίας δεδομένων [200].

Σε γενικές γραμμές, η επιλογή χαρακτηριστικών παίζει ζωτικής σημασίας ρόλο στην αποτροπή του φαινομένου της υπερπροσαρμογής (overfitting), στη διευκόλυνση της οπτικοποίησης δεδομένων, στη μείωση των απαιτήσεων αποθήκευσης και του υπολογιστικού κόστους, και στη βελτίωση της ακρίβειας των αλγορίθμων. Επιπλέον, η επιλογή χαρακτηριστικών διατηρεί το φυσικό νόημα των αρχικών χαρακτηριστικών και παρέχει καλύτερη αναγνωσιμότητα και ερμηνευσιμότητα για το μοντέλο δεδομένων [200].

Πέρα από τα ανωτέρω πλεονεκτήματα, η επιλογή των σημαντικότερων χαρακτηριστικών από το σύνολο όλων των χαρακτηριστικών που περιγράφουν ένα πρόβλημα στοχεύει στο να οδηγήσει στην καλύτερη διαίσθηση για την αποτελεσματικότερη αντιμετώπισή του από τους ειδικούς επιστήμονες. Το πλεονέκτημα αυτό είναι ιδιαίτερα χρήσιμο στο ιατρικό πρόβλημα που περιγράφεται στην παρούσα διατριβή καθώς τα σημαντικότερα χαρακτηριστικά που θα εξαχθούν από τις μεθόδους που παρουσιάζονται στο υποκεφάλαιο 6.1.3, θα βοηθήσουν στην ενίσχυση της διαγνωστικής διαδικασίας μέσω της καλύτερης διαίσθησης που θα δοθεί στους χειρουργούς μαστού ως προς τον κίνδυνο ανάπτυξης καρκίνου του μαστού.

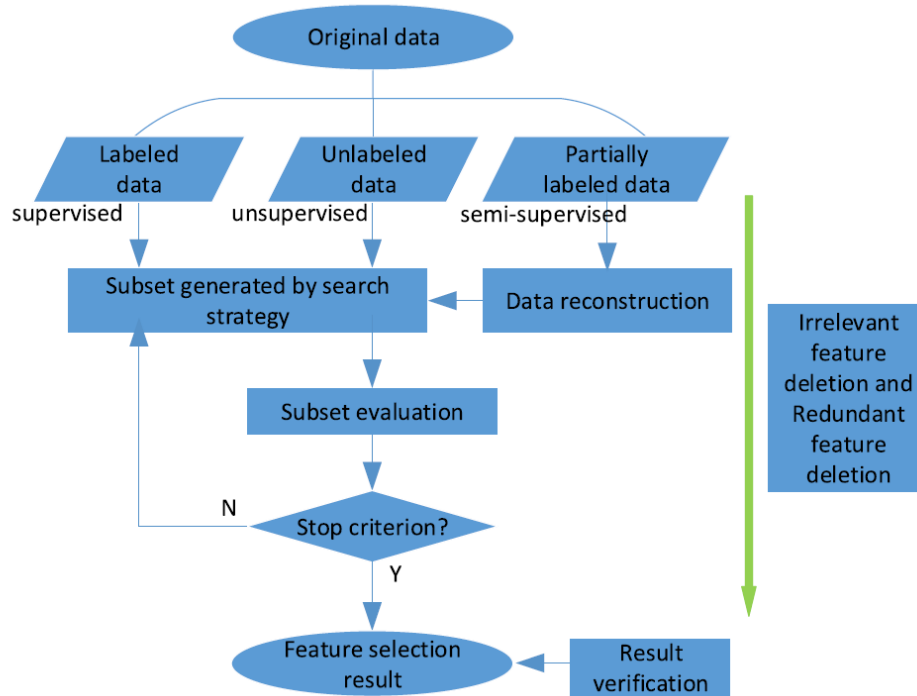
Αναλυτικότερα, η επιλογή χαρακτηριστικών αναφέρεται στη διαδικασία διάκρισης ενός υποσυνόλου από ένα αρχικό σύνολο χαρακτηριστικών με στόχο την απόκτηση του πιο κατατοπιστικού υποσυνόλου επιλέγοντας σημαντικά χαρακτηριστικά ή διαγράφοντας ασήμαντα χαρακτηριστικά από το αρχικό σύνολο χαρακτηριστικών. Παίζει ρόλο στη συμπίεση της κλίμακας επεξεργασίας δεδομένων, όπου αφαιρούνται τα περιττά και άσχετα χαρακτηριστικά. Η τεχνική επιλογής χαρακτηριστικών μπορεί να θεωρηθεί κάποιες φορές ως στάδιο προεπεξεργασίας των αλγορίθμων μηχανικής μάθησης και τα καλά αποτελέσματα της επιλογής χαρακτηριστικών μπορούν να βελτιώσουν την ακρίβεια μάθησης, να μειώσουν το χρόνο μάθησης, και να απλοποιήσουν τα αποτελέσματά της. Ειδικότερα, η επιλογή και η εξαγωγή χαρακτηριστικών είναι δύο τρόποι μείωσης της διάστασης του διανύσματος των χαρακτηριστικών (βλ. Εικόνα 17). Σε αντίθεση με την επιλογή χαρακτηριστικών, η εξαγωγή χαρακτηριστικών πρέπει συνήθως να μετασχηματίσει τα αρχικά δεδομένα σε χαρακτηριστικά με ικανότητα ισχυρής αναγνώρισης προτύπων, όπου τα αρχικά δεδομένα μπορούν να θεωρηθούν ως χαρακτηριστικά με αδύναμη ικανότητα αναγνώρισης [201]. Η επιλογή χαρακτηριστικών, η οποία αποτελεί μείζον ερευνητικό θέμα για την επιστημονική κοινότητα εδώ και δεκαετίες, χρησιμοποιείται σε πολλούς τομείς, όπως αναγνώριση εικόνων [202], [203], ανάκτηση εικόνων [204], εξόρυξη κειμένου [204], ανάλυση δεδομένων βιοπληροφορικής [205], [206], διάγνωση σφαλμάτων [207], και ούτω καθεξής.



Εικόνα 17: Απεικόνιση της διαφοράς των διαδικασιών Εξαγωγής χαρακτηριστικών (Feature extraction) και Επιλογής χαρακτηριστικών (Feature selection) [208].

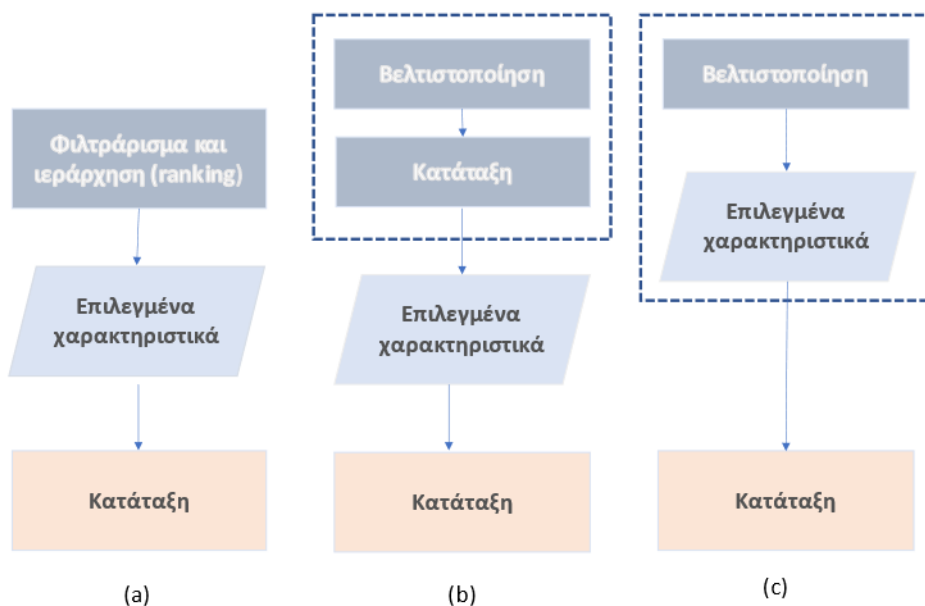
Σύμφωνα με τη θεωρία, πολλές μέθοδοι επιλογής χαρακτηριστικών βασίζονται στη στατιστική [209], [210] ή τη θεωρία της πληροφορίας [211] και μπορούν να κατηγοριοποιηθούν σύμφωνα με διάφορα πρότυπα [201]:

- ✚ Σύμφωνα με τα χρησιμοποιούμενα δεδομένα εκπαίδευσης (επισημασμένα, μη επισημασμένα, ή μερικώς επισημασμένα), οι μέθοδοι επιλογής χαρακτηριστικών μπορούν να χωριστούν σε επιβλεπόμενες (supervised), μη επιβλεπόμενες (unsupervised) και ημιεπιβλεπόμενες (semi-supervised) όπως φαίνονται στην Εικόνα 18.



Εικόνα 18: Πλαίσιο για την επιλογή χαρακτηριστικών [201].

- ✚ Σύμφωνα με τη σχέση τους με τις μεθόδους μάθησης, οι μέθοδοι επιλογής χαρακτηριστικών μπορούν να ταξινομηθούν σε μοντέλα φίλτρου (filter), περιτυλίγματος (wrapper), και ενσωματωμένα μοντέλα (embedded models) (βλ. Εικόνα 19).



Εικόνα 19: Λογικό διάγραμμα για τη σχέση μεταξύ των μοντέλων φίλτρου, ενσωματωμένων μοντέλων και περιτυλίγματος [212]

- ✚ Σύμφωνα με το κριτήριο αξιολόγησης, οι μέθοδοι επιλογής χαρακτηριστικών μπορούν να προκύψουν από τη συσχέτιση (correlation), την ευκλείδεια απόσταση (Euclidean distance), τη συνέπεια (consistency), την εξάρτηση (dependence) και το μέτρο πληροφορίας (information measure).
- ✚ Σύμφωνα με τις στρατηγικές αναζήτησης, οι μέθοδοι επιλογής χαρακτηριστικών μπορούν να χωριστούν σε εμπρόσθιες (forward selection), οπίσθιες (backward selection), τυχαία και υβριδικά μοντέλα.
- ✚ Σύμφωνα με τον τύπο της εξόδου, οι μέθοδοι επιλογής χαρακτηριστικών μπορούν να χωριστούν σε μοντέλα κατάταξης χαρακτηριστικών (weighting) και μοντέλα επιλογής υποσυνόλων (subset selection).

Τα μοντέλα φίλτρου εξετάζουν μόνο τη συσχέτιση μεταξύ του χαρακτηριστικού και της ετικέτας της κλάσης. Σε σύγκριση με τα μοντέλα περιτυλίγματος, έχουν μικρότερο υπολογιστικό κόστος. Το κριτήριο αξιολόγησης είναι κρίσιμο για τα μοντέλα φίλτρου. Εν τω μεταξύ, τα ενσωματωμένα μοντέλα επιλέγουν το χαρακτηριστικό στη διαδικασία εκπαίδευσης του μοντέλου μάθησης και το αποτέλεσμα της επιλογής του χαρακτηριστικού εξέρχεται αυτόματα, ενώ η διαδικασία εκπαίδευσης έχει ολοκληρωθεί [201].

Το βασικό πλεονέκτημα των μεθόδων φιλτραρίσματος έναντι των μεθόδων περιτυλίγματος είναι ότι διακρίνονται από απλότητα και έχουν χαμηλότερο υπολογιστικό κόστος και πολυπλοκότητα. Ιδιαίτερα στην περίπτωση δεδομένων υψηλών διαστάσεων, οι μέθοδοι περιτυλίγματος έχουν πολύ πιο υψηλό υπολογιστικό κόστος καθώς κάθε συνδυασμός χαρακτηριστικών πρέπει να αξιολογηθεί από τον ταξινομητή. Αντίθετα το φιλτράρισμα εκτελείται μόνο μία φορά και με το αποτέλεσμα του μπορούν να δοκιμαστούν διάφοροι ταξινομητές. Οι μέθοδοι φιλτραρίσματος είναι εμπειρικά επιτυχείς αλγόριθμοι αφού ικανοποιούν το βασικό τους στόχο, την απομάκρυνση των λιγότερο σχετικών χαρακτηριστικών. Επιπλέον, οι μέθοδοι φιλτραρίσματος έχουν καλύτερες ιδιότητες γενίκευσης επειδή δεν εξαρτώνται από οποιοδήποτε αλγόριθμο εκμάθησης σε αντίθεση με τις μεθόδους περιτυλίγματος που έχουν μεγαλύτερη πιθανότητα υπερπροσαρμογής (overfitting). Ωστόσο, για τον ίδιο ακριβώς λόγο το κύριο μειονέκτημα των μεθόδων φιλτραρίσματος είναι ότι αγνοούν την αλληλεπίδραση με τον ταξινομητή: ο καλύτερος συνδυασμός χαρακτηριστικών μπορεί να μην είναι ανεξάρτητος από τις επαγωγικές και αναπαραστατικές προτιμήσεις του αλγορίθμου μάθησης που θα χρησιμοποιηθεί για την κατασκευή του ταξινομητή. Αντιθέτως, οι μέθοδοι περιτυλίγματος έχουν καλύτερη απόδοση επειδή η FS βελτιστοποιείται ειδικά για τον επιλεγμένο ταξινομητή και κατά συνέπεια μπορούν να βρουν ένα μικρό υποσύνολο χαρακτηριστικών με μεγάλη ακρίβεια [213].

Η απόδοση της μεθόδου επιλογής χαρακτηριστικών συνήθως αξιολογείται από κάποιο μοντέλο μηχανικής μάθησης. Τα συνήθως χρησιμοποιούμενα μοντέλα μηχανικής μάθησης περιλαμβάνουν τα NaiveBayes, KNN, SVM, BPNN, K-means, την ιεραρχική ομαδοποίηση, την ομαδοποίηση με βάση την πυκνότητα και ούτω καθεξής [214]–[216]. Μια καλή μέθοδος επιλογής χαρακτηριστικών θα πρέπει να έχει υψηλή ακρίβεια εκμάθησης αλλά με τη μικρότερη υπολογιστική επιβάρυνση (χρονική και χωρική πολυπλοκότητα) [201].

Το μέτρο χαρακτηριστικών (feature measure) ή το κριτήριο αξιολόγησης αποτελεί τη βάση της επιλογής χαρακτηριστικών και παίζει σημαντικό ρόλο στη διαδικασία. Στο πλαίσιο των προβλημάτων ταξινόμησης, βέλτιστο κριτήριο επιλογής θα ήταν το ποσοστό σφάλματος $E(S)$ που βασίζεται στον κανόνα Bayes και φαίνεται στον τύπο (1) υπό τη συνεχή ή διακριτή συνθήκη, όπου $c_i \in C$ είναι μια κλάση από όλες τις πιθανές κλάσεις C που υπάρχουν στα δεδομένα [201].

$$(1) E(S) = \sum_s p(S) \left(1 - \max_i (p(c_i|S)) \right)$$

Όπως φαίνεται από τον τύπο (1), το $E(S)$ έχει τη μορφή αθροίσματος και το $p(S) \left(1 - \max_i (p(c_i|S)) \right)$ είναι μη γραμμικό και μη αρνητικό. Το $E(S)$ έχει ένα ανώτερο όριο που φαίνεται στον τύπο (2), όπου $H(C|S)$ είναι η υπό συνθήκη εντροπία του C δεδομένου του S [201].

$$(2) E(S) \leq H(C|S) / 2$$

Καθότι είναι δύσκολο να υπολογιστεί άμεσα το $E(S)$, επειδή το S είναι ο συνδυασμός των χαρακτηριστικών, οι ερευνητές προτιμούν να χρησιμοποιούν άλλα μέτρα, όπως η συσχέτιση, η εξάρτηση και η απόσταση [201].

Τα μέτρα αυτά χρησιμοποιούνται συχνά σε μετασχηματισμούς για ειδικές εφαρμογές. Επιπλέον, υπάρχουν και άλλα μέτρα, όπως η βαθμολογία Laplacian, η βαθμολογία Fisher και ούτω καθεξής. Τα μέτρα πληροφορίας απαιτούν το χαρακτηριστικό να είναι σε διακριτό τύπο, επομένως η διακριτοποίηση του χαρακτηριστικού πρέπει να υλοποιηθεί πριν από την επιλογή του. Οι μέθοδοι διακριτοποίησης περιλαμβάνουν τις μεθόδους Equal-Depth, Equal-Width, τη χειροκίνητη ρύθμιση κ.λπ. [217]. Μια άλλη συχνά χρησιμοποιούμενη μέθοδος είναι το ελάχιστο μήκος περιγραφής (Minimum Description Length - MDL) [201].

6.1.1 Επιβλεπόμενη Επιλογή Χαρακτηριστικών (Supervised Feature Selection)

Η επιβλεπόμενη επιλογή χαρακτηριστικών (supervised feature selection) είναι προσανατολισμένη στο πρόβλημα της ταξινόμησης και χρησιμοποιεί, ως θεμελιώδη αρχή, τη συνάφεια ή τη συσχέτιση μεταξύ του χαρακτηριστικού και της ετικέτας της κλάσης. Οι τεχνικές της επιβλεπόμενης επιλογής χαρακτηριστικών μπορούν να χρησιμοποιηθούν για επισημασμένα δεδομένα (με γνωστές ετικέτες κλάσεων) και χρησιμοποιούνται για τον εντοπισμό των σημαντικών χαρακτηριστικών με σκοπό την αύξηση της αποτελεσματικότητας των επιβλεπόμενων μοντέλων που αφορούν σε ταξινόμηση και παλινδρόμηση. Στο ιατρικό πρόβλημα που μελετάει η παρούσα διατριβή η μεταβλητή εξόδου αποτελείται από δυο κλάσεις (Κλάση 1: Αρνητικό για Ca μαστού, Κλάση 2: Θετικό για Ca μαστού) επομένως εφόσον είναι γνωστές οι ετικέτες των κλάσεων εξόδου αλλά και του διανύσματος των χαρακτηριστικών εισόδου πρόκειται για πρόβλημα επιβλεπόμενης επιλογής χαρακτηριστικών [201].

Η σημασία των χαρακτηριστικών μπορεί να αξιολογηθεί με τα μέτρα συνάφειας. Για ένα δεδομένο σύνολο δεδομένων $D = (X, C)$ με ένα σύνολο χαρακτηριστικών $X = \{x_1, x_2, \dots, x_n\}$ και μια κλάση C , το επιβλεπόμενο μοντέλο στοχεύει στην εύρεση ενός βέλτιστου υποσυνόλου χαρακτηριστικών που μεγιστοποιεί την ακρίβεια ταξινόμησης. Έχουν προταθεί πολλά επιβλεπόμενα μοντέλα, όπως το CFS [218] και το ReliefF [219], [220]. Η CFS είναι μια μέθοδος επιλογής υποσυνόλου και χρησιμοποιεί κυρίως ευρυστικές προσεγγίσεις για την αξιολόγηση της επίδρασης του μεμονωμένου χαρακτηριστικού που αντιστοιχεί σε κάθε κατηγορία, προκειμένου να ληφθεί ένα βέλτιστο υποσύνολο χαρακτηριστικών. Η μέθοδος ReliefF [219] επεκτείνεται από την Relief [220] για να υποστηρίξει το πρόβλημα των πολλαπλών κατηγοριών. Η κύρια ιδέα του μοντέλου είναι ότι λαμβάνει την ευκλείδεια απόσταση ως δείκτη συσχέτισης

και στη συνέχεια προσπαθεί να σταθμίσει τα χαρακτηριστικά ανάλογα με το πόσο καλά διαφοροποιούν τα περιστατικά διαφορετικών κατηγοριών.

Το μέτρο εξάρτησης Hilbert-Schmidt [221] που φαίνεται στον τύπο (3) είναι ένα γενικό πλαίσιο για την επιλογή χαρακτηριστικών, όπου το $J(\cdot)$ μετρά την εξάρτηση ενός υποσυνόλου χαρακτηριστικών από το C . Η ιδέα αυτού του πλαισίου είναι ότι το καλό υποσύνολο χαρακτηριστικών πρέπει να μεγιστοποιεί το $J(\cdot)$, γεγονός που μετατρέπει την επιλογή χαρακτηριστικών σε ένα πρόβλημα βελτιστοποίησης.

$$(3) D_{FS} = \arg \max_{S \subseteq F} [J(S)]$$

Στην πραγματικότητα, το μέτρο εξάρτησης και το μέτρο συσχέτισης είναι ουσιαστικά το ίδιο. Οι μέθοδοι επιλογής χαρακτηριστικών φιλτραρίσματος χρησιμοποιούν συνήθως κριτήρια αξιολόγησης για να ενισχύσουν τη συσχέτιση μεταξύ του χαρακτηριστικού και της ετικέτας της κλάσης και να μειώσουν τη συσχέτιση μεταξύ των χαρακτηριστικών. Επιπλέον, η συσχέτιση μεταξύ των χαρακτηριστικών συχνά αντικαθίσταται από τον πλεονασμό (redundancy) ή την ποικιλομορφία (diversity) (ή απόσταση - distance). Τα μέτρα της συσχέτισης (relevance), του πλεονασμού και της ποικιλομορφίας μπορεί να είναι ίδια ή διαφορετικά.

6.1.1.1 Επιβλεπόμενο μοντέλο φίλτρου βασισμένο στη συσχέτιση και τον πλεονασμό

Η ανάλυση συσχέτισης (Relevance) (κλάση - χαρακτηριστικό) και πλεονασμού (Redundancy) (χαρακτηριστικό - χαρακτηριστικό) είναι η βάση αυτού του τύπου μεθόδων. Αυτά τα μοντέλα χρησιμοποιούν την ευκλείδεια απόσταση, τη συσχέτιση Pearson και τα μέτρα πληροφορίας για την ανάλυση συσχέτισης και πλεονασμού [222]. Τα χαρακτηριστικά σε ένα αρχικό σύνολο δεδομένων μπορούν να χωριστούν σε τέσσερις ομάδες [223]: (α) εντελώς άσχετα και θορυβώδη χαρακτηριστικά, (β) ασθενώς συναφή και πλεονάζοντα χαρακτηριστικά, (γ) ασθενώς συναφή και μη πλεονάζοντα χαρακτηριστικά και (δ) πολύ συναφή χαρακτηριστικά. Τα αποτελέσματα της επιβλεπόμενης επιλογής χαρακτηριστικών (Supervised feature selection) θα πρέπει να περιλαμβάνουν τις ομάδες (γ) και (δ). Η ανάλυση συσχέτισης και πλεονασμού μπορεί να μετατραπεί σε δύο προβλήματα βελτιστοποίησης: $\max(\text{Relevance}(S; C))$ και $\min(\text{Redundancy}(S))$.

Ένα κλασικό κριτήριο για την επιλογή χαρακτηριστικών με βάση την ανάλυση συσχέτισης και πλεονασμού είναι το mRMR (Max - Relevance and Min - Redundancy), το οποίο χρησιμοποιεί την αμοιβαία πληροφορία ως μέτρο αξιολόγησης. Στον τύπο (4), το πρώτο στοιχείο είναι η πρώτη τάξης έκφραση της ανάλυσης συσχέτισης και το δεύτερο στοιχείο ανήκει στην ανάλυση πλεονασμού. Αυτή η έκφραση μπορεί να επεκταθεί περαιτέρω σε μορφή υπό συνθήκη αμοιβαίας πληροφορίας όπως παρουσιάζεται στον τύπο (5) [224].

$$(4) \text{ mRMR} : \max \left[\frac{1}{|S|} \sum_{x_i \in S} I(x_i; C) - \frac{1}{|S|^2} \sum_{x_i \in S} \sum_{x_j \in S} I(x_i; x_j) \right]$$

$$(5) \text{ CmRMR} : \max \left[\frac{1}{|S|} \sum_{x_i \in S} I(x_i; C) - \frac{1}{|S|^2} \sum_{x_i \in S} \sum_{x_j \in S} (I(x_i; x_j) - I(x_i; x_j|C)) \right]$$

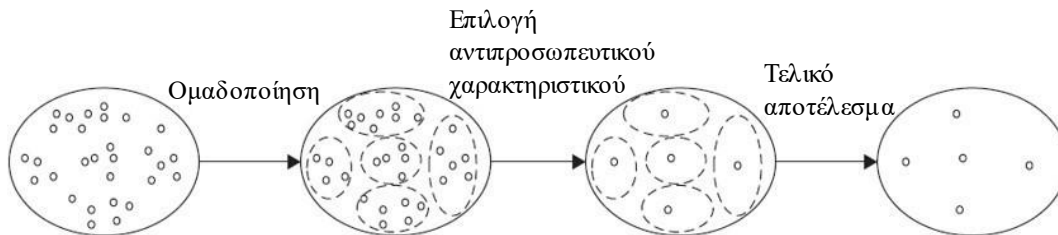
Στο παρελθόν, οι μέθοδοι επιλογής χαρακτηριστικών που βασίζονται στην αμοιβαία πληροφορία έχουν σημειώσει μεγάλη επιτυχία. Αυτές οι μέθοδοι [225]–[227] επικεντρώνονται κυρίως στην ανάλυση της συσχέτισης με χρήση αμοιβαίων πληροφοριών. Οι Peng et al. [224] πρότειναν το mRMR που εφαρμόζει αυτό το κριτήριο με μια επαυξητική αναζήτηση πρώτης τάξης. Ωστόσο, το mRMR, που βασίζεται στην αμοιβαία πληροφορία, ελαχιστοποιεί μόνο την αμοιβαία πληροφορία χαρακτηριστικών και αγνοεί την απόδοση ταξινόμησης των υποψήφιων χαρακτηριστικών, η οποία μπορεί να επηρεάζεται από τα επιλεγμένα χαρακτηριστικά. Στη συνέχεια, παρατίθεται η υπό συνθήκη ανάλυση αμοιβαίας πληροφορίας προκειμένου να ξεπεραστεί αυτό το πρόβλημα. Οι μέθοδοι επιλογής χαρακτηριστικών που βασίζονται στην υπό συνθήκη αμοιβαία πληροφορία έχουν προσελκύσει σημαντική προσοχή [228]–[230]. Οι μέθοδοι επιλογής χαρακτηριστικών που βασίζονται στη θεωρία πληροφοριών επιλέγουν υποσύνολα χαρακτηριστικών που μεγιστοποιούν την πληροφορία όσον αφορά στην ετικέτα της κλάσης. Η προσεγγιστική έκφραση και η αυξητική ευριστική αναζήτηση είναι προσεγγίσεις που χρησιμοποιούνται συνήθως από αυτές τις μεθόδους, επειδή ο άμεσος υπολογισμός της αμοιβαίας πληροφορίας μεταξύ του υποσυνόλου χαρακτηριστικών και της ετικέτας κλάσης είναι δύσκολος.

6.1.1.2 Επιβλεπόμενο μοντέλο φίλτρου βασισμένο στη συσχέτιση και την ποικιλομορφία

Πολλοί ερευνητές μελετούν μεθόδους επιλογής χαρακτηριστικών με έμφαση στην ποικιλομορφία (Diversity) μεταξύ των χαρακτηριστικών [231], [232]. Στη βιβλιογραφία [231], η απόκλιση Kullback-Leibler χρησιμοποιείται για τη δημιουργία ενός πλαισίου επιλογής χαρακτηριστικών, το οποίο εστιάζει στην ξεχωριστή τάξη. Αυτός ο τύπος μοντέλου θα μπορούσε επίσης να εκφραστεί με δύο προβλήματα βελτιστοποίησης: $\max(\text{Relevance}(\hat{S}; C))$ και $\max(\text{Diversity}(\hat{S}))$. Οι Dhillon et al. [233] πρότειναν ένα συνολικό κριτήριο για την ομαδοποίηση χαρακτηριστικών/λέξεων και παρουσίασαν μια γρήγορη, διαιρετική μέθοδο για την ταξινόμηση κειμένου. Η μέθοδος αυτή μεγιστοποιεί την απόκλιση Jensen-Shannon μεταξύ των συστάδων χαρακτηριστικών.

Οι επιλογές χαρακτηριστικών με βάση την ομαδοποίηση [234]–[236] συνήθως πραγματοποιούνται με όρους μεγιστοποίησης της ποικιλομορφίας. Η βασική διαδικασία (όπως φαίνεται στην Εικόνα 20) σε αυτές τις μεθόδους αποτελείται από τρία βήματα: πρώτα, επιλέγεται το κατάλληλο μέτρο απόστασης για να

σχηματιστεί ο χώρος των χαρακτηριστικών, στη συνέχεια, τα χαρακτηριστικά ομαδοποιούνται με τη μέθοδο ομαδοποίησης (clustering) και τέλος, επιλέγεται το αντιπροσωπευτικό χαρακτηριστικό κάθε ομάδας για να παραχθεί το αποτέλεσμα της επιλογής.



Εικόνα 20: Η διαδικασία επιλογής χαρακτηριστικών με βάση την ομαδοποίηση (clustering) [201].

Το αντιπροσωπευτικό χαρακτηριστικό είναι συχνά το χαρακτηριστικό με τη μεγαλύτερη σχέση με την ετικέτα της κλάσης. Οι Ienco και Meo [236] ομαδοποίησαν ιεραρχικά τα σύνολα χαρακτηριστικών σύμφωνα με τις συσχετίσεις τους και επέλεξαν το καλύτερο χαρακτηριστικό από κάθε ομάδα για να σχηματίσουν το τελικό υποσύνολο χαρακτηριστικών χρησιμοποιώντας μια μέθοδο συσκευασίας (packing). Παρά το γεγονός ότι αυτή η μέθοδος δεν χρειάζεται να ρυθμίσει καμία παράμετρο, αυξάνει το χρονικό κόστος και προσθέτει προκατάληψη (bias) στη μέθοδο μάθησης. Οι Witten και Tibshirani [235] εισήγαγαν την αραιή K-means (sparse K-means) και την ιεραρχική ομαδοποίηση σε ένα πλαίσιο ομαδοποίησης χαρακτηριστικών. Αρχικά, το σύνολο χαρακτηριστικών ομαδοποιείται και χωρίζεται σε πολλαπλές συστάδες χαρακτηριστικών. Στη συνέχεια, χρησιμοποιούνται παράγοντες ποινής (penalty factors) τύπου Lasso για την επιλογή των αντιπροσωπευτικών χαρακτηριστικών από κάθε συστάδα, τα οποία αποτελούν το τελικό υποσύνολο χαρακτηριστικών. Στο περιβάλλον μάθησης με επίβλεψη, ο Liu [234] χρησιμοποίησε τη μέθοδο συσσωρευτικής ιεραρχικής ομαδοποίησης για να διαιρέσει το σύνολο χαρακτηριστικών και να δημιουργήσει το τελικό υποσύνολο χαρακτηριστικών αφαιρώντας τα χαρακτηριστικά που απέχουν πολύ μεταξύ τους. Ο Zhao [237] χρησιμοποίησε τον συντελεστή μέγιστης πληροφορίας ως μέτρο συσχέτισης των χαρακτηριστικών, πραγματοποίησε την ομαδοποίηση διάδοσης της συγγένειας (affinity propagation clustering) του υποσυνόλου χαρακτηριστικών και στη συνέχεια επέλεξε το κεντροειδές (centroid) από κάθε ομάδα ως το αντιπροσωπευτικό χαρακτηριστικό της ομάδας.

Επιπλέον, έχουν παρουσιαστεί πολλές μέθοδοι επιλογής χαρακτηριστικών που βασίζονται στην ομαδοποίηση και χρησιμοποιούν τη θεωρία της πληροφορίας. Οι Au et al. [238] πρότειναν μια μέθοδο η οποία επιλέγει ένα μέτρο πληροφορίας για τη μέτρηση της συσχέτισης μεταξύ των χαρακτηριστικών, ομαδοποιεί τα χαρακτηριστικά με μια μέθοδο ομαδοποίησης τύπου K-means και στη συνέχεια επιλέγει το

αντιπροσωπευτικό χαρακτηριστικό. Αυτή η μέθοδος είναι πολύ αποτελεσματική για την ταξινόμηση γονιδιακών δεδομένων. Η διαδικασία ομαδοποίησης μπορεί να οδηγήσει σε ατέρμονες επαναλήψεις, επομένως ο αριθμός τους πρέπει να προκαθοριστεί προκειμένου να σταματήσει η διαδικασία ομαδοποίησης. Η μέθοδος FAST [206], που προτάθηκε από τον Song, χρησιμοποιεί τη μέθοδο της ιεραρχικής ομαδοποίησης. Αυτή η μέθοδος θεωρεί κάθε ελάχιστο δένδρο ως συστάδα και χρησιμοποιεί τη συμμετρική αβεβαιότητα ως μέτρο. Η FAST είναι κατάλληλη για δεδομένα υψηλής διάστασης, αλλά ο αριθμός των επιλεγμένων χαρακτηριστικών δεν μπορεί να καθοριστεί χειροκίνητα. Ο Liu πρότεινε επίσης μια μέθοδο επιλογής χαρακτηριστικών ομαδοποίησης MFC [239] που βασίζεται στο ελάχιστο δένδρο διάσπασης (minimum spanning tree). Σε αντίθεση με τη FAST, το μέτρο που χρησιμοποιεί η MFC, είναι το μέτρο της απόστασης της πληροφορίας, δηλαδή η διακύμανση της πληροφορίας. Ο Sotoca [240] παρουσίασε μια μέθοδο επιλογής χαρακτηριστικών βασισμένη στην ιεραρχική ομαδοποίηση mRR, όπου η υπό συνθήκη αμοιβαία πληροφορία λαμβάνεται ως μέτρο ποικιλομορφίας. Γενικά, όσο μικρότερος είναι ο πλεονασμός μεταξύ των χαρακτηριστικών, τόσο μεγαλύτερη είναι η ποικιλομορφία μεταξύ τους. Ωστόσο, το μέτρο του πλεονασμού δεν μπορεί να χρησιμοποιηθεί άμεσα ως ποικιλομορφία. Οι παραπάνω μέθοδοι χρησιμοποιούν το SU [206], την παραλλαγή (variant) της πληροφορίας [239] και την υπό συνθήκη αμοιβαία πληροφορία [240] ως μέτρα ποικιλομορφίας, αντίστοιχα, αλλά όχι την αμοιβαία πληροφορία όπως στο κριτήριο mRMR.

Η μέθοδος που βασίζεται στην ομαδοποίηση μπορεί να επιλέξει ένα άσχετο χαρακτηριστικό, καθώς τα άσχετα χαρακτηριστικά συχνά ομαδοποιούνται μαζί. Επομένως, είναι καλύτερο να διαγράφονται τα άσχετα χαρακτηριστικά πριν από την ομαδοποίησή τους. Πρόσφατα, ο Zou [241] πρότεινε τη μέθοδο κατάταξης MRMD (Max - Relevance Max - Distance) χρησιμοποιώντας ένα υβριδικό μέτρο για την κατάταξη των χαρακτηριστικών. Αυτή η μέθοδος παρουσιάζει καλύτερη σταθερότητα για την επιλογή χαρακτηριστικών, αλλά και μικρότερο χρόνο εκτέλεσης σε σύγκριση με τη μέθοδο mRMR [241].

6.1.1.3 Επιβλεπόμενο μοντέλο περιτυλίγματος

Τα μοντέλα περιτυλίγματος (wrapper) λαμβάνουν το σφάλμα ταξινόμησης ή το ποσοστό ακρίβειας ως πρότυπο αξιολόγησης χαρακτηριστικών. Το αποτέλεσμα της επιλογής χαρακτηριστικών συχνά παράγεται ταυτόχρονα με εκείνο του μοντέλου μάθησης, επειδή η μέθοδος μάθησης περιλαμβάνεται στην επιλογή χαρακτηριστικών. Σε σύγκριση με το μοντέλο φίλτρου, το μοντέλο περιτυλίγματος μπορεί να επιτύχει υψηλότερη ακρίβεια ταξινόμησης και τείνει να έχει μικρότερο μέγεθος υποσυνόλου, ωστόσο έχει φτωχή ικανότητα γενίκευσης (generalization capability) και υψηλή χρονική πολυπλοκότητα. Η μέθοδος SVM-RFE, η οποία προτάθηκε από τους Guyon et al. [242], είναι μια ευρέως μελετημένη και εφαρμοσμένη μέθοδος wrapper που χρησιμοποιεί SVM για τη μέτρηση της απόδοσης των χαρακτηριστικών και κατασκευάζει έναν ταξινομητή με υψηλές βαθμολογίες απόδοσης [243]. Οι Michalak και Kwasnicka [244]

πρότειναν μια δυαδική μέθοδο επιλογής χαρακτηριστικών wrapper. Στη μέθοδο αυτή παρουσιάζεται ένα σύστημα κατάταξης με βάση τις πυκνότητες των κλάσεων για δυαδικά δεδομένα. Στη βιβλιογραφία [245], χρησιμοποιείται επίσης μια μέθοδος δύο σταδίων που χρησιμοποιεί μια μέθοδο φίλτρου χαμηλού κόστους για την κατάταξη των χαρακτηριστικών και μια «δαπανηρή» μέθοδο wrapper για την περαιτέρω εξάλειψη των άσχετων μεταβλητών.

Το μοντέλο φίλτρου συνδυάζεται συχνά με το μοντέλο περιτύγματος για να σχηματίσει την υβριδική μέθοδο επιλογής χαρακτηριστικών. Γενικά, η υβριδική επιλογή χαρακτηριστικών αποτελείται από δύο στάδια: το πρώτο στάδιο είναι η χρήση του φίλτρου για τη μείωση του μεγέθους του χώρου των χαρακτηριστικών με την αφαίρεση των άσχετων και των θορυβωδών χαρακτηριστικών - το δεύτερο στάδιο είναι η χρήση του περιτυλίγματος για την εύρεση του βέλτιστου υποσυνόλου χαρακτηριστικών από τα διατηρούμενα χαρακτηριστικά. Οι Akadi et al. [246] συνδύασαν τη μέθοδο mRMR και τον γενετικό αλγόριθμο για να αποκτήσουν μια υβριδική μέθοδο φίλτρου - περιτυλίγματος. Οι Cadenas et al. [247] πρότειναν μια υβριδική μέθοδο Filter-Wrapper με χρήση Fuzzy Random Forest, η οποία ονομάζεται FRFfs.

6.2 Τεχνικές Επιλογής Χαρακτηριστικών από το σύνολο των 35 Χαρακτηριστικών του Διανύσματος Εισόδου

Στο παρόν υποκεφάλαιο χρησιμοποιούνται οι τεχνικές κατάταξης χαρακτηριστικών (μοντέλα φίλτρου) προκειμένου να αξιολογηθούν και καταταχθούν τα χαρακτηριστικά του διανύσματος εισόδου ως προς τη σημασία τους για την κατάταξη ενός περιστατικού σε μία εκ των δύο κλάσεων της μεταβλητής εξόδου. Τα διαθέσιμα περιστατικά όπως εξηγούμε στο υποκεφάλαιο 5.2 είναι χωρισμένα σε 2 κλάσεις (Κλάση 1: Αρνητικό για Ca μαστού και Κλάση 2: Θετικό για Ca μαστού). Εφαρμόζοντας τα ακόλουθα μέτρα και τους αλγορίθμους κατάταξης αξιολογήθηκε ξεχωριστά κάθε μια από τις 35 μεταβλητές (χαρακτηριστικά) που χαρακτηρίζει κάθε περιστατικό (βλ. Πίνακας 7):

6.2.1 Information Gain

Κέρδος πληροφορίας (Information Gain - IG) είναι η αλλαγή στην εντροπία της κλάσης από μια προηγούμενη κατάσταση σε μια κατάσταση όπου είναι γνωστή η τιμή ενός χαρακτηριστικού. Σε αυτή την περίπτωση χρησιμοποιείται για να δείξει τη συνάφεια των χαρακτηριστικών. Υπολογίζεται χρησιμοποιώντας τα ακόλουθα βήματα [248]:

Βήμα πρώτο: Δεδομένου ενός χαρακτηριστικού A και μιας κλάσης C, το πρώτο βήμα είναι ο υπολογισμός της Εντροπίας (H) πριν από την παρατήρηση του χαρακτηριστικού A, η οποία δίνεται από τον τύπο [248]:

$$(6) H(C) = -\sum p(c) \log_2 p(c)$$

Όπου $H(C)$ είναι η εντροπία πριν την παρατήρηση μιας μεταβλητής a στην κλάση C , $p(c)$ είναι η κατανομή πιθανότητας των παρατηρήσεων c και η μεταβλητή c είναι μέλος της κλάσης C .

Βήμα δεύτερο: Είναι ο υπολογισμός της Εντροπίας μετά την παρατήρηση ενός χαρακτηριστικού A , η οποία δίνεται από τον ακόλουθο τύπο [248]:

$$(7) H(C/A) = \sum p(a) \sum P(c|a) \log_2 p(c)$$

Όπου a είναι ένα μέλος του συνόλου A ($a \in A$) και c είναι μέλος του C ($c \in C$).

Αυτό προσδιορίζει την εντροπία μετά την παρατήρηση χαρακτηριστικών που ανήκουν σε υποσύνολα του κύριου συνόλου δεδομένων. Όσο μικρότερη είναι η τιμή της εντροπίας η οποία επιτυγχάνεται από τον διαχωρισμό των δεδομένων με βάση τη μεταβλητή A , τόσο αποτελεσματικότερα η μεταβλητή αυτή διαχωρίζει τα δεδομένα στις σωστές κλάσεις. Αυτά τα δύο πρώτα βήματα είναι σημαντικά επειδή δίνουν εντροπίες που χρησιμοποιούνται για να προκύψει το κέρδος πληροφορίας στο επόμενο βήμα.

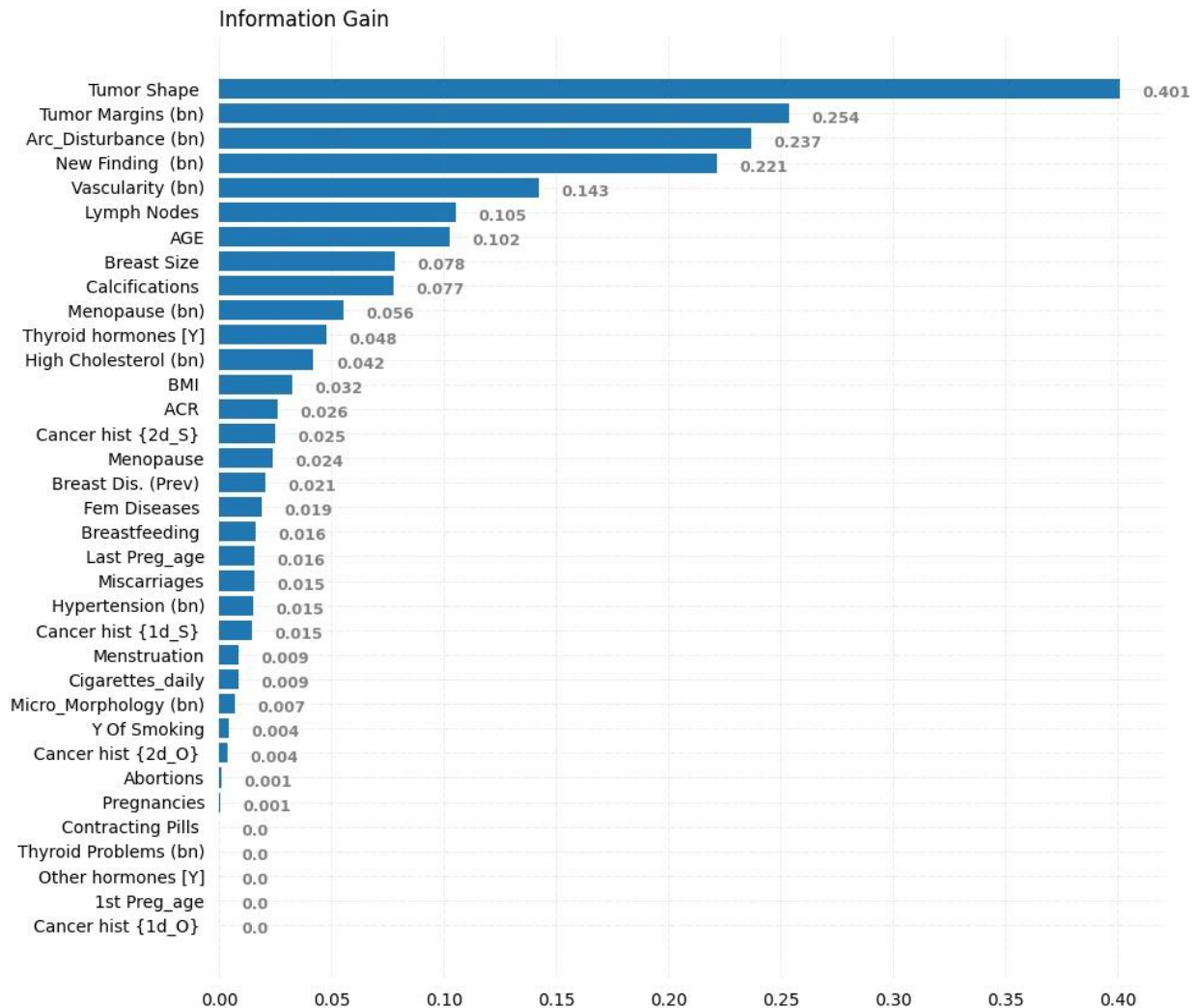
Το τελευταίο βήμα είναι ο υπολογισμός του κέρδους πληροφορίας. Το κέρδος πληροφορίας του χαρακτηριστικού A είναι η διαφορά μεταξύ της εντροπίας πριν από την παρατήρηση του χαρακτηριστικού a στο A και της εντροπίας μετά την παρατήρησή του. Αυτό δίνεται από τον παρακάτω τύπο [248]:

$$(8) IG(A/C) = H(C) - H(C/A)$$

Όπου $IG(A/C)$ είναι το κέρδος πληροφορίας του χαρακτηριστικού A δεδομένης της κλάσης C , $H(C)$ είναι η εντροπία πριν την παρατήρηση του A και $H(C/A)$ είναι η εντροπία μετά την παρατήρηση.

Στην παρούσα έρευνα, υπολογίζεται το κέρδος πληροφορίας κάθε χαρακτηριστικού και ορίζεται ένα κατώφλι για τον καθορισμό του τελικού συνόλου χαρακτηριστικών που χρησιμοποιείται αργότερα για την ταξινόμηση. Οι τιμές του κέρδους πληροφορίας για κάθε χαρακτηριστικό παρουσιάζονται στην Εικόνα 21. Όπως βλέπουμε το χαρακτηριστικό με την υψηλότερη τιμή κέρδους πληροφορίας είναι το σχήμα της οζώδους σκίασης (Tumor Shape), πληροφορία η οποία παρέχεται από την έκθεση του US και όπως επιβεβαιώνει και η βιβλιογραφία [104] αποτελεί καθοριστικής σημασίας παράγοντα κινδύνου για τον καρκίνο του μαστού. Στη συνέχεια, δεύτερο στη σειρά κατάταξης ακολουθεί το χαρακτηριστικό που αφορά στα περιθώρια του εντοπισμένου ευρήματος (Tumor Margins) που επίσης εντοπίζεται στον US και υπάρχει ισχυρή βιβλιογραφία που αποδεικνύει τη σημασία του [104]. Ακολουθούν δυο εξαιρετικά σημαντικά χαρακτηριστικά που αντλούμε από την έκθεση της MAMMO, αυτό της διαταραχής της αρχιτεκτονικής των μαστών (Arc_Disturbance) και το χαρακτηριστικό που αφορά στην ύπαρξη νέου ευρήματος σε σύγκριση με παλαιότερη μαστογραφία (New finding). Στη συνέχεια, η αγγείωση (Vascularity) της οζώδους αλλοίωσης, η κατάσταση των λεμφαδένων (Lymph Nodes), η ηλικία της γυναίκας (age), το μέγεθος του

στήθους (breast size), το είδος των αποτιτανώσεων (Calcifications) και η εμμηνόπαυση (Menopause (bn)), συμπληρώνουν τη δεκάδα των χαρακτηριστικών με τη μεγαλύτερη τιμή κέρδους πληροφορίας. Στη λίστα κατάταξης ακολουθούν τα χαρακτηριστικά του ατομικού ιατρικού φακέλου της γυναίκας, με εκείνα που αφορούν σε αναπαραγωγικούς παράγοντες να προηγούνται εκείνων που αφορούν στο οικογενειακό ιστορικό.



Εικόνα 21: Οι τιμές του κέρδους πληροφορίας (IG) για τα 35 χαρακτηριστικά του διανύσματος εισόδου

6.2.2. mRMR

Η μέθοδος ελάχιστου πλεονασμού και μέγιστης σχετικότητας mRMR (minimum redundancy maximum relevance) είναι ένα φίλτρο με επιλογή προς τα εμπρός (forward selection). Σε κάθε βήμα, η μέθοδος επεκτείνει το σύνολο των επιλεγμένων χαρακτηριστικών με το επόμενο, ώστε να μεγιστοποιείται η συνάφεια προς τη μεταβλητή - στόχο και ταυτόχρονα να ελαχιστοποιείται ο πλεονασμός μεταξύ των ήδη επιλεγμένων μεταβλητών. Στόχος της μεθόδου είναι να καταλήξει στα χαρακτηριστικά τα οποία

παρουσιάζουν ταυτόχρονα τη μικρότερη συσχέτιση μεταξύ τους (minimum redundancy) και την υψηλότερη συνάφεια (μέγιστη σχετικότητα) με την κλάση (maximum relevance) [249].

Έστω ότι το σύνολο δεδομένων αναπαρίσταται από έναν δισδιάστατο πίνακα, του οποίου οι γραμμές είναι οι παρατηρήσεις και οι στήλες οι επεξηγηματικές μεταβλητές. Συμβολίζουμε $\{\chi_1, \chi_2 \dots \chi_n\}$ ως το σύνολο των n παρατηρήσεων και $\{X_1, X_2 \dots X_k\}$ ως το σύνολο των k επεξηγηματικών μεταβλητών ή χαρακτηριστικών. Έστω ότι η μεταβλητή στόχος Y είναι ένα διάνυσμα $(y_1, y_2 \dots y_n)$. Στη γενική περίπτωση όταν οι τιμές των επεξηγηματικών μεταβλητών είναι πραγματικοί αριθμοί, τότε οι μεταβλητές αποτελούν τον k - διάστατο χώρο R^k . Οι παρατηρήσεις είναι σημεία δεδομένων σε αυτόν τον υψηλοδιάστατο χώρο. Η εν λόγω μέθοδος αποσκοπεί στην εύρεση του υποσυνόλου S των επεξηγηματικών μεταβλητών που χαρακτηρίζουν βέλτιστα τη μεταβλητή - στόχο Y . Κάτω από τον βέλτιστο χαρακτηρισμό, θα εννοήσουμε την υψηλότερη στατιστική εξάρτηση της μεταβλητής - στόχου Y από το επιλεγμένο υποσύνολο S . Το κριτήριο αυτό είναι γνωστό ως μέγιστη εξάρτηση (maximal dependency). Στην mRMR, η αμοιβαία πληροφορία χρησιμοποιείται ως μέτρο εξάρτησης δύο τυχαίων μεταβλητών. Η αμοιβαία πληροφορία μεταξύ του συνόλου των επεξηγηματικών μεταβλητών S και της μεταβλητής - στόχου Y ορίζεται ως [249]:

$$(9) I(S; Y) = \iint p(S, Y) \log \frac{p(S, Y)}{p(S)p(Y)} dSdY,$$

Όπου $p(Y)$ είναι η συνάρτηση πυκνότητας πιθανότητας της συνεχούς τυχαίας μεταβλητής Y . Οι όροι $p(S)$ και $p(S, Y)$ συμβολίζουν τις κοινές συναρτήσεις πυκνότητας πιθανότητας των συνόλων συνεχών τυχαίων μεταβλητών S και $S \cup \{Y\}$, αντίστοιχα.

Στην περίπτωση των συνεχών τυχαίων μεταβλητών, ο υπολογισμός της αμοιβαίας πληροφορίας είναι δύσκολος, επειδή οι πυκνότητες είναι όλες άγνωστες και είναι δύσκολο να εκτιμηθούν από τα δεδομένα. Επομένως, προτιμάται η διακριτοποίηση των τιμών ή η προσέγγιση της πυκνότητας. Στην περίπτωση των διακριτών μεταβλητών Z και Y , η αμοιβαία πληροφορία ορίζεται ως εξής [249]:

$$(10) I(Z; Y) = \sum_{z_i} \sum_{y_i} P(Z = z_i, Y = y_i) \log \frac{P(Z = z_i, Y = y_i)}{P(Z = z_i) P(Y = y_i)}$$

Οι πιθανότητες P προσδιορίζονται από τους πίνακες ενδεχομένων που καλύπτουν την καταμέτρηση των τιμών εμφάνισης των μεταβλητών Z και Y . Ο στόχος του εν λόγω αλγορίθμου είναι να βρεθεί το υποσύνολο S των επεξηγηματικών μεταβλητών στο οποίο η Y εμφανίζει την υψηλότερη εξάρτηση που εκφράζεται από τον τύπο 11 χρησιμοποιώντας την αμοιβαία πληροφορία ως [249]:

$$(11) \max D(S, Y), \quad D = I(S; Y)$$

Επειδή το κριτήριο της μέγιστης εξάρτησης είναι δύσκολο να εφαρμοστεί, η προτιμώμενη εναλλακτική λύση είναι η χρήση της μέγιστης συνάφειας (Max - Relevance) ως κριτήριο. Ο όρος Max-Relevance αναζητά την υψηλότερη εξάρτηση της μεταβλητής - στόχου από τις επιμέρους επιλεγμένες μεταβλητές, δηλαδή αναζητά τις μεταβλητές που ικανοποιούν το εξής κριτήριο [249]:

$$(12) \max D(S, Y), \quad D = \frac{1}{|S|} \sum_{X_i \in S} I(X_i; Y).$$

Η παραπάνω εξίσωση προσεγγίζει την Max - Dependency με την υψηλότερη μέση εξάρτηση της μεταβλητής - στόχου Y από τις επιμέρους μεταβλητές από το σύνολο S . Ωστόσο, οι επεξηγηματικές μεταβλητές που επιλέγονται με αυτή την προσέγγιση μπορεί να είναι περιττές. Εάν μία από τις περιττές μεταβλητές παραλειφθεί, η απόδοση πρόβλεψης δεν μειώνεται. Επομένως, η συνθήκη για τον ελάχιστο πλεονασμό (Min - Redundancy) για την απόρριψη περιττών μεταβλητών προστίθεται [249]:

$$(13) \min R(S), \quad R = \frac{1}{|S|^2} \sum_{X_i, X_j \in S} I(X_i, X_j; Y).$$

Το κριτήριο που συνδυάζει και τις δύο συνθήκες είναι το mRMR. Η ακόλουθη συνάρτηση αντιπροσωπεύει το απλούστερο τρόπο για την ταυτόχρονη βελτιστοποίηση της συνάφειας D και του πλεονασμού R [249]:

$$(14) \max \Phi(D, R), \quad \Phi = D - R, \quad \text{αντίστοιχα } \Phi = \frac{D}{R}$$

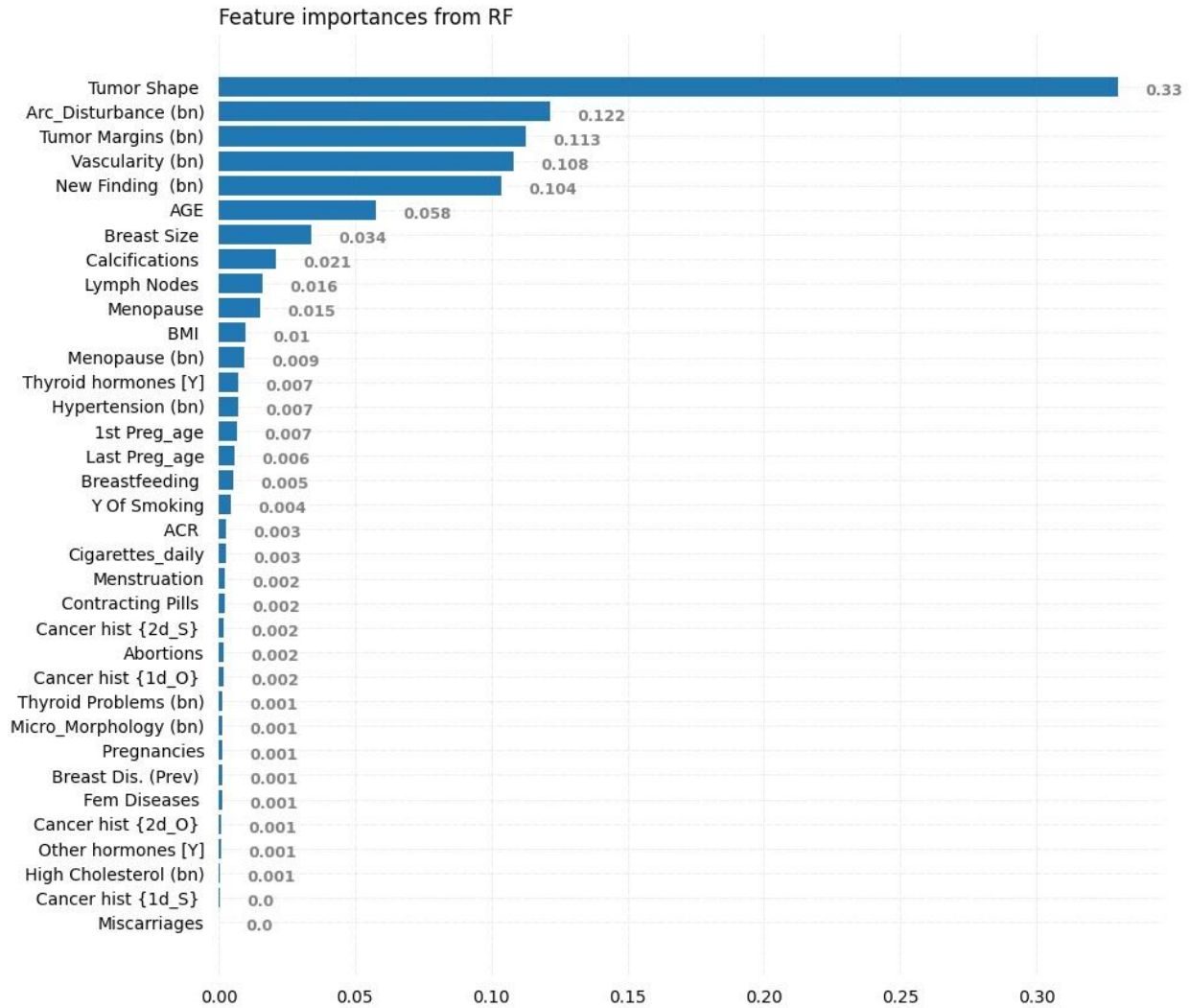
Όπως προαναφέρθηκε, στόχος της συνάρτησης είναι να βρεθεί το χαρακτηριστικό εκείνο το οποίο μεγιστοποιεί τη συνάφεια μεταξύ αυτού και της κλάσης - στόχου και ταυτόχρονα ελαχιστοποιεί τον πλεονασμό μεταξύ αυτού και των υπολοίπων που έχουν ήδη επιλεγεί. Αρχικά, ως πρώτο κατατάσσεται το πιο συναφές χαρακτηριστικό, αυτό δηλαδή με τη μεγαλύτερη αμοιβαία πληροφορία με την κλάση, και απομακρύνεται από το σύνολο των χαρακτηριστικών [249]. Ο mRMR επαναλαμβάνει την ίδια διαδικασία για τα υπόλοιπα χαρακτηριστικά μέχρι να προκύψει η τελική λίστα κατάταξης που παρουσιάζεται στον Πίνακα 13.

6.2.3 Random Forest

Ο αλγόριθμος Random Forest έχει ενσωματωμένη τη σημαντικότητα των χαρακτηριστικών που μπορεί να υπολογιστεί μέσω του δείκτη gini importance. Όπως αναλύθηκε ανωτέρω ο RF αποτελείται από ένα σύνολο δέντρων αποφάσεων, κάθε δέντρο είναι ένα σύνολο εσωτερικών κόμβων και φύλλων. Στον εσωτερικό κόμβο, το επιλεγμένο χαρακτηριστικό χρησιμοποιείται για τη λήψη απόφασης σχετικά με τον τρόπο

διαχωρισμού του συνόλου δεδομένων σε δύο ξεχωριστά σύνολα με παρόμοιες αποκρίσεις στο εσωτερικό τους. Τα χαρακτηριστικά για τους εσωτερικούς κόμβους επιλέγονται με κάποιο κριτήριο, το οποίο για τις διαδικασίες ταξινόμησης μπορεί να είναι η gini impurity («ακαθαρότητα») ή το IG, ενώ για την παλινδρόμηση είναι η μείωση της διακύμανσης. Μπορούμε να μετρήσουμε τον τρόπο με τον οποίο κάθε χαρακτηριστικό μειώνει την «ακαθαρότητα» του διαχωρισμού (το χαρακτηριστικό με τη μεγαλύτερη μείωση επιλέγεται για τον εσωτερικό κόμβο). Για κάθε χαρακτηριστικό μπορούμε να συλλέξουμε το κατά πόσον (μέσο όρο) μειώνει την «ακαθαρότητα». Ο μέσος όρος σε όλα τα δέντρα του δάσους είναι το μέτρο της σημαντικότητας του χαρακτηριστικού. Το μεγαλύτερο πλεονέκτημα αυτής της μεθόδου είναι η ταχύτητα υπολογισμού - καθώς όλες οι απαραίτητες τιμές υπολογίζονται κατά τη διάρκεια της εκπαίδευσης του Random Forest. Τα μειονεκτήματα της μεθόδου είναι η τάση να προτιμώνται (να επιλέγονται ως σημαντικά) αριθμητικά και κατηγορικά χαρακτηριστικά με πολλές διαφορετικές τιμές. Επιπλέον, στην περίπτωση συσχετιζόμενων χαρακτηριστικών μπορεί να επιλέξει ένα από τα χαρακτηριστικά και να παραμελήσει τη σημασία του δεύτερου (γεγονός που μπορεί να οδηγήσει σε λανθασμένα συμπεράσματα).

Στην παρούσα μελέτη χρησιμοποιήσαμε το μέτρο Gini importance προκειμένου να κατατάξουμε σε σειρά σημαντικότητας τα χαρακτηριστικά εισόδου με χρήση του αλγορίθμου Random Forest. Στην Εικόνα 22 παρουσιάζεται η σειρά κατάταξης που προέκυψε με υψηλότερης σημασίας χαρακτηριστικό να αναδεικνύεται το σχήμα της οζώδους σκίασης (Tumor Shape) όπως και στην περίπτωση του IG.



Εικόνα 22: Κατάταξη των 35 χαρακτηριστικών εισόδου με χρήση του αλγορίθμου Random Forest και με βάση το μέτρο Gini importance.

6.2.4 Λίστες Κατάταξης των Χαρακτηριστικών Εισόδου

Στον παρακάτω πίνακα (Πίνακας 13) παρουσιάζονται οι επιμέρους λίστες κατάταξης των χαρακτηριστικών εισόδου ως προς την αποτελεσματικότητα/σημαντικότητα τους για την κατάταξη ενός περιστατικού σε μία εκ των δύο κλάσεων της μεταβλητής εξόδου. Παρατηρώντας τον Πίνακα 13 διαπιστώνουμε ότι και οι τρεις τεχνικές κατάταξης συμφωνούν στα 5 πρώτα χαρακτηριστικά των λιστών και ακολουθούν παρόμοια κατάταξη μέχρι περίπου τα 10 χαρακτηριστικά με μικρές διαφοροποιήσεις. Είναι εμφανές ότι τα χαρακτηριστικά που εξάγονται από τις απεικονιστικές εξετάσεις (μαστογραφία & υπέρηχο) βρίσκονται στις υψηλότερες θέσεις αναδεικνύοντας έτσι τη σημαντικότητά τους ανεξάρτητα από τη μέθοδο επιλογής χαρακτηριστικών. Όπως διαπιστώνουμε από τις λίστες κατάταξης, τα πρώτα 5 κοινά χαρακτηριστικά

αφορούν σε πληροφορίες που εξάγονται από την αναφορά της μαστογραφίας (Arc_Disturbance (bn), New Finding (bn)) και του υπερήχου (Tumor Shape, Tumor Margins (bn), Vascularity (bn)).

Συγκεκριμένα, όπως βλέπουμε το χαρακτηριστικό με την υψηλότερη τιμή κέρδους πληροφορίας (Information Gain) είναι το σχήμα της οζώδους σκίασης (Tumor Shape), πληροφορία η οποία παρέχεται από την έκθεση του US και όπως επιβεβαιώνει και η βιβλιογραφία αποτελεί καθοριστικής σημασίας παράγοντα κινδύνου για τον καρκίνο του μαστού [104]. Συγκεκριμένα, όσο πιο ακανόνιστο είναι το σχήμα της οζώδους αλλοίωσης τόσο περισσότερο ενοχοποιείται για ύπαρξη καρκινικού όγκου. Στη συνέχεια, δεύτερο στη σειρά κατάταξης με βάση το κέρδος πληροφορίας ακολουθεί το χαρακτηριστικό που αφορά στα περιθώρια του εντοπισμένου ευρήματος (Tumor Margins) που επίσης εντοπίζεται στον US και υπάρχει ισχυρή βιβλιογραφία που αποδεικνύει τη σημασία του, καθώς η μη ομαλότητα των περιθωρίων της εντοπισμένης αλλοίωσης υποδεικνύει αυξημένη πιθανότητα ύπαρξης κακοήθειας [104]. Αντίστοιχα στη λίστα κατάταξης που προέκυψε με τη μέθοδο mRMR, τα δύο αυτά χαρακτηριστικά είναι επίσης τα δύο πρώτα με αντίστροφη σειρά, ενώ για τη λίστα κατάταξης σύμφωνα με το δείκτη Gini Importance πρώτο χαρακτηριστικό είναι και πάλι το σχήμα της εντοπισμένης αλλοίωσης κατά τον υπέρηχο ενώ δεύτερο ακολουθεί το χαρακτηριστικό της διαταραχής των μαστών (Arc_Disturbance) που αναφέρεται στην έκθεση της μαστογραφίας. Αναλυτικότερα, όπως έχει προαναφερθεί, η διαταραχή της αρχιτεκτονικής των μαστών θεωρείται ένας κρίσιμης σημασίας παράγοντας για την ενίσχυση της υποψίας υφιστάμενης κακοήθειας [104], [191]

Στη συνέχεια, όσον αφορά στη λίστα κατάταξης του Information Gain, ακολουθούν δυο εξαιρετικά σημαντικά χαρακτηριστικά που αντλούνται από την έκθεση της μαστογραφίας, αυτό της διαταραχής της αρχιτεκτονικής των μαστών (Arc_Disturbance) και το χαρακτηριστικό που αφορά στην ύπαρξη νέου ευρήματος σε σύγκριση με παλαιότερη μαστογραφία (New finding). Η λίστα με μια μικρή διαφοροποίηση συμφωνεί με αυτή του mRMR με τη διαφορά ότι στην τρίτη θέση είναι η ύπαρξη αγγείωσης (Vascularity) στο υπερηχογραφικό εύρημα, που αποτελεί επιβαρυντικό στοιχείο για την αξιολόγηση του περιστατικού [104] και έπειτα ακολουθεί η αρχιτεκτονική των μαστών. Στις αντίστοιχες θέσεις της λίστας κατά Gini importance απαντάμε το χαρακτηριστικό που αφορά στα περιθώρια της αλλοίωσης αλλά και την ύπαρξη αγγείωσης όπως και στις άλλες δύο λίστες.

Τα επόμενα χαρακτηριστικά σύμφωνα με τη λίστα κατάταξης του Information Gain είναι η ύπαρξη αγγείωσης (Vascularity) στο υπερηχογραφικό εύρημα και η κατάσταση των λεμφαδένων (Lymph Nodes) όπως αυτή αποτυπώνεται στην αναφορά του υπερήχου [104]. Για τη λίστα του mRMR και για αυτή του Gini importance, στην 5^η θέση βρίσκεται το χαρακτηριστικό που αφορά στην ύπαρξη νέου ευρήματος, και ακολουθούν το μέγεθος μαστών της γυναίκας (Breast size), στοιχείο που λαμβάνεται από τον ατομικό

ιατρικό φάκελο και όπως φαίνεται έχει σημαντική συμβολή στην εμφάνιση της νόσου [62], [63], και η ηλικία της γυναίκας αντίστοιχα [35], [36], [37]. Το ίδιο χαρακτηριστικό, η ηλικία της εξεταζόμενης γυναίκας, βρίσκεται στην 7^η θέση των λιστών Information Gain και mRMR και το μέγεθος μαστών στην αντίστοιχη θέση της λίστας Gini Importance.

Στη συνέχεια, στις λίστες κατάταξης παρατηρούνται κάποιες διαφοροποιήσεις, ωστόσο και στις τρεις λίστες εντοπίζονται σε σχετικά υψηλές θέσεις τα χαρακτηριστικά του μεγέθους του στήθους (Breast Size), της ύπαρξης αποτιτανώσεων (Calcifications) [19], της ηλικίας της εμμηνόπαυσης (Menopause) [46], [47], της κατάστασης των λεμφαδένων (Lymph Nodes), της διάρκειας λήψης ορμονών για την αντιμετώπιση προβλημάτων στο θυρεοειδή αδένα (Thyroid Hormones [Y]) [88], [89], της υψηλής χοληστερόλης (High Cholesterol) [87] και του δείκτη μάζας σώματος (BMI) [51], [52].

Έπειτα, παρατηρείται μεγαλύτερη διαφοροποίηση στις λίστες κατάταξης και ακολουθούν χαρακτηριστικά όπως είναι η πυκνότητα των μαστών (ACR) [60], η υπέρταση (Hypertension) [86] και το κάπνισμα (Cigarettes_daily) [85]. Αν εξαιρέσουμε το γεγονός ότι στη λίστα κατάταξης με την τεχνική mRMR, το χαρακτηριστικό που αφορά στους συγγενείς πρώτου βαθμού με καρκίνο του μαστού (Cancer hist {1d_S}) βρίσκεται στην 14^η θέση, και αντίστοιχα το που αφορά στους συγγενείς δεύτερου βαθμού με καρκίνο του μαστού (Cancer hist {2d_S}) βρίσκεται στην 15^η θέση [70], φαίνεται τα χαρακτηριστικά τα οποία αφορούν στο οικογενειακό ιστορικό δεν διαδραματίζουν σημαντικό ρόλο για την ύπαρξη καρκίνου του μαστού στη γυναίκα καθώς βρίσκονται σε αρκετά χαμηλές θέσεις στις λίστες κατάταξης με βάση τη σημαντικότητα των χαρακτηριστικών. Αυτή η διαπίστωση θα μπορούσε να εξηγηθεί με βάση τα στατιστικά στοιχεία που θέλουν μόνο ένα μικρό ποσοστό της τάξης του να οφείλεται σε κληρονομικούς παράγοντες (5 - 10% του συνόλου των καρκίνων μαστού) [79].

Πίνακας 13: Επιμέρους λίστες κατάταξης των χαρακτηριστικών εισόδου ως προς την αποτελεσματικότητα/σημαντικότητά τους για κάθε μια από τις τεχνικές κατάταξης

Σειρά κατάταξης (σημαντι- κότητα)	Information Gain	mRMR	Random Forest (Gini Importance)
1	Tumor Shape	Tumor Margins (bn)	Tumor Shape
2	Tumor Margins (bn)	Tumor Shape	Arc_Disturbance (bn)

3	Arc_Disturbance (bn)	Vascularity (bn)	Tumor Margins (bn)
4	New Finding (bn)	Arc_Disturbance (bn)	Vascularity (bn)
5	Vascularity (bn)	New Finding (bn)	New Finding (bn)
6	Lymph Nodes	Breast Size	Age
7	Age	Age	Breast Size
8	Breast Size	Menopause	Calcifications
9	Calcifications	Lymph Nodes	Lymph Nodes
10	Menopause (bn)	Thyroid Hormones [Y]	Menopause
11	Thyroid hormones [Y]	Menopause (bn)	BMI
12	High Cholesterol (bn)	Calcifications	Menopause (bn)
13	BMI	BMI	Thyroid hormones [Y]
14	ACR	Cancer hist {1d_S}	Hypertension (bn)
15	Cancer hist {2d_S}	Cigarettes_daily	1 st Preg_age
16	Menopause	Micro_Morphology (bn)	Last Preg_age
17	Breast Dis. (Prev)	Hypertension (bn)	Breastfeeding
18	Fem Diseases	1 st Preg_age	Y of Smoking
19	Breastfeeding	Thyroid Problems (bn)	ACR
20	Last Preg_age	High Cholesterol (bn)	Cigarettes_daily

21	Miscarriages	Cancer hist {2d_S}	Menstruation
22	Hypertension (bn)	Y Of Smoking	Contracting Pills
23	Cancer hist {1d_S}	Abortions	Cancer hist {2d_S}
24	Menstruation	Age of Last Pregnancy	Abortions
25	Cigarettes_daily	ACR	Cancer hist {1d_O}
26	Micro_Morphology (bn)	Other Hormones [Y]	Thyroid Problems (bn)
27	Y of Smoking	Contracting Pills	Micro_Morphology (bn)
28	Cancer hist {2d_O}	Total breastfeeding Months	Pregnancies
29	Abortions	Pregnancies	Breast Dis. (Prev)
30	Pregnancies	Fem Diseases	Fem Diseases
31	Contracting Pills	Menstruation	Cancer hist {2d_O}
32	Thyroid Problems (bn)	Can Cancer hist {1d_O}	Other hormones [Y]
33	Other hormones [Y]	Breast Dis. (Prev)	High Cholesterol (bn)
34	1 st Preg_age	Miscarriages	Cancer hist {1d_S}
35	Cancer hist {1d_O}	Cancer hist {2d_S}	Miscarriages

6.3 Μοντέλα Πρόβλεψης/Ταξινόμησης

Για τους σκοπούς της παρούσας διδακτορικής διατριβής, ερευνήθηκε η διαγνωστική απόδοση των αλγορίθμων Naive Bayes και Random Forest. Στόχος της κατασκευής των προαναφερθέντων μοντέλων

ταξινόμησης είναι η αξιολόγηση της ικανότητας τους να διακρίνουν την πραγματική κατάσταση των περιστατικών με τα οποία δοκιμάζονται και η επιλογή των συγκεκριμένων αλγορίθμων έγινε για λόγους που αναλύονται στη συνέχεια του κεφαλαίου.

6.3.1 Μπεϋζιανός Ταξινομητής Naive Bayes

Η πιθανοτική προσέγγιση του προβλήματος της ταξινόμησης βασίζεται στην πιθανοτική μοντελοποίηση των σχέσεων μεταξύ των χαρακτηριστικών και της μεταβλητής στόχου που αντιπροσωπεύει τις κλάσεις του προβλήματος. Για την ταξινόμηση ενός άγνωστου δείγματος, ένας Μπεϋζιανός ταξινομητής προβαίνει στην εκτίμηση των εκ των υστέρων πιθανοτήτων των κλάσεων του προβλήματος δοθέντων των τιμών των χαρακτηριστικών του προς ταξινόμηση δείγματος. Εν συνεχεία, ο αλγόριθμος ταξινομεί το δείγμα στην κλάση με τη μεγαλύτερη εκ των υστέρων πιθανότητα, δηλαδή στην κλάση που εμφανίζει τη μεγαλύτερη πιθανότητα να ανήκει το δείγμα [199].

Για ένα πρόβλημα ταξινόμησης M κλάσεων, $\omega_1, \omega_2, \dots, \omega_M$, και ένα άγνωστο δείγμα με διάνυσμα χαρακτηριστικών \mathbf{x} , ορίζονται M δεσμευμένες πιθανότητες ως προς τις κλάσεις $P(\omega_i|\mathbf{x}), i = 1, 2, \dots, M$. Η πιθανότητα $P(\omega_i|\mathbf{x})$ είναι η δεσμευμένη πιθανότητα της κλάσης ω_i δεδομένου του διανύσματος \mathbf{x} , αναπαριστά δηλαδή την πιθανότητα το άγνωστο δείγμα να ανήκει στην κλάση ω_i . Οι δεσμευμένες πιθανότητες αυτές είναι γνωστές ως εκ των υστέρων πιθανότητες των κλάσεων ω_i και μπορούν να υπολογιστούν με βάση το θεώρημα του Bayes:

$$(15) P(\omega_i|\mathbf{x}) = \frac{p(\mathbf{x}|\omega_i)P(\omega_i)}{p(\mathbf{x})}$$

όπου $p(\mathbf{x}|\omega_i)$ είναι η δεσμευμένη συνάρτηση πυκνότητας πιθανότητας του διανύσματος \mathbf{x} δεδομένου της κλάσης ω_i , $P(\omega_i)$ είναι η εκ των προτέρων πιθανότητα της κλάσης ω_i , και $p(\mathbf{x})$ είναι η περιθώρια συνάρτηση πυκνότητας πιθανότητας του διανύσματος \mathbf{x} . Η συνάρτηση πυκνότητας πιθανότητας $p(\mathbf{x}|\omega_i)$ είναι γνωστή και ως συνάρτηση πιθανοφάνειας (likelihood function) των ω_i ως προς \mathbf{x} . Στα ανωτέρω έχει γίνει η υπόθεση ότι τα χαρακτηριστικά μπορούν να πάρουν οποιαδήποτε τιμή στον l -διάστατο διανυσματικό χώρο χαρακτηριστικών. Αν τα χαρακτηριστικά παίρνουν διακριτές τιμές, τότε οι συναρτήσεις $p(\mathbf{x}|\omega_i)$ και $p(\mathbf{x})$ αντιπροσωπεύουν πιθανότητες και σημειώνονται ως $P(\mathbf{x}|\omega_i)$ και $P(\mathbf{x})$ αντίστοιχα [199].

Για την ταξινόμηση ενός άγνωστου δείγματος με διάνυσμα χαρακτηριστικών \mathbf{x} ο Μπεϋζιανός ταξινομητής προβαίνει στην εκτίμηση των εκ των υστέρων πιθανοτήτων $P(\omega_i|\mathbf{x})$ και κατατάσσει το δείγμα στην κλάση που εμφανίζει τη μεγαλύτερη πιθανότητα να ανήκει το δείγμα σε αυτήν. Επομένως, για να πραγματοποιηθεί

η ταξινόμηση αρκεί να υπολογισθούν οι απαιτούμενες πιθανότητες για την εκτίμηση των εκ των υστέρων πιθανοτήτων των κλάσεων.

Οι εκ των προτέρων πιθανότητες των κλάσεων $P(\omega_i)$, $i = 1, 2, \dots, M$, μπορούν να εκτιμηθούν από τα διαθέσιμα δεδομένα εκπαίδευσης. Θεωρώντας ένα πρόβλημα ταξινόμησης με δυο κλάσεις, έστω ω_1 και ω_2 , αν N είναι ο συνολικός αριθμός των δειγμάτων εκπαίδευσης και N_1, N_2 από αυτά ανήκουν στις κλάσεις ω_1, ω_2 αντίστοιχα, τότε $P(\omega_1) \approx N_1/N$ και $P(\omega_2) \approx N_2/N$. Οι δεσμευμένες συναρτήσεις πυκνότητας πιθανότητας $p(\mathbf{x}|\omega_i)$, $i = 1, 2, \dots, M$, περιγράφουν την κατανομή των χαρακτηριστικών σε κάθε μια από τις κλάσεις και μπορούν και αυτές να υπολογιστούν από τα διαθέσιμα δεδομένα. Τέλος, η συνάρτηση πυκνότητας πιθανότητας $p(\mathbf{x})$ υπολογίζεται από τις δύο προηγούμενες ποσότητες σύμφωνα με τον τύπο:

$$(16) p(\mathbf{x}) = \sum_{i=1}^M p(\mathbf{x}|\omega_i)P(\omega_i)$$

Με βάση τα ανωτέρω, στην περίπτωση ενός προβλήματος ταξινόμησης δύο κλάσεων, έστω ω_1 και ω_2 , ο κανόνας ταξινόμησης του Bayes μπορεί να εκφραστεί ως εξής:

$$\text{Αν } P(\omega_1|\mathbf{x}) > P(\omega_2|\mathbf{x}), \text{ το } x \text{ ταξινομείται στην } \omega_1$$

$$\text{Αν } P(\omega_1|\mathbf{x}) < P(\omega_2|\mathbf{x}), \text{ το } x \text{ ταξινομείται στην } \omega_2$$

Με βάση τον ανωτέρω κανόνα ταξινόμησης και το θεώρημα του Bayes, η απόφαση ταξινόμησης μπορεί να εκφραστεί σύμφωνα με τις ανισότητες:

$$(17) p(\mathbf{x}|\omega_1)P(\omega_1) \geq p(\mathbf{x}|\omega_2)P(\omega_2)$$

Η ποσότητα $p(\mathbf{x})$ δεν παίζει κανένα ρόλο στην απόφαση ταξινόμησης διότι είναι η ίδια για όλες τις κλάσεις. Επομένως, το πρόβλημα ανάγεται στην εκτίμηση των $P(\omega_i)$ και $p(\mathbf{x}|\omega_i)$. Αν και όπως αναφέρθηκε ανωτέρω οι εκ των προτέρων πιθανότητες των κλάσεων μπορούν εύκολα να υπολογισθούν από τα δεδομένα εκπαίδευσης, οι συναρτήσεις πυκνότητας πιθανότητας $p(\mathbf{x}|\omega_i)$, οι οποίες εκφράζουν την κατανομή των χαρακτηριστικών στις κλάσεις, δεν είναι εύκολο να εκτιμηθούν. Πρέπει είτε να γνωρίζουμε την κατανομή των χαρακτηριστικών, είτε να θεωρήσουμε ότι τα χαρακτηριστικά ακολουθούν μια γνωστή κατανομή, είτε να εκτιμήσουμε την κατανομή τους από τα δεδομένα. Όσον αφορά στην τελευταία περίπτωση, υπάρχουν διαθέσιμες διάφορες τεχνικές για την εκτίμηση των συναρτήσεων πυκνότητας πιθανότητας από τα δεδομένα εκπαίδευσης [250]. Ωστόσο, προκειμένου να προκύψουν ακριβείς εκτιμήσεις των συναρτήσεων πυκνότητας πιθανότητας πρέπει το πλήθος των δειγμάτων εκπαίδευσης να είναι αρκετά μεγάλο. Η ανάγκη για επιπλέον δεδομένα εκπαίδευσης αυξάνει εκθετικά σε σχέση με τη

διάσταση l του χώρου των χαρακτηριστικών. Αν από N δείγματα εκπαίδευσης μπορεί να προκύψει ακριβής εκτίμηση της κατανομής ενός χαρακτηριστικού (για διανυσματικό χώρο μιας διάστασης), θα χρειαστούν N^l δείγματα εκπαίδευσης για ένα διανυσματικό χώρο l διαστάσεων. Επομένως, στην περίπτωση μεγάλων διαστάσεων του χώρου χαρακτηριστικών (μεγάλων τιμών του l) δεν μπορούμε να προβούμε στην εκτίμηση της συνάρτησης πυκνότητας πιθανότητας με μεγάλη ακρίβεια αφού είναι πολύ δύσκολο να συγκεντρωθεί ένας πολύ μεγάλος αριθμός δεδομένων εκπαίδευσης.

Προκειμένου να ξεπεραστεί το ανωτέρω πρόβλημα, μπορούμε να προβούμε στην εκτίμηση των συναρτήσεων πυκνότητας πιθανότητας θεωρώντας ότι τα χαρακτηριστικά είναι μεταξύ τους ανεξάρτητα. Με δεδομένη την υπόθεση αυτή, η εκτίμηση της συνάρτησης πυκνότητας πιθανότητας είναι πιο απλή και ανάγεται στην παραγοντοποίηση των συναρτήσεων πυκνότητας πιθανότητας κάθε χαρακτηριστικού ξεχωριστά:

$$(18) p(x|\omega_i) = \prod_{j=1}^l p(x_j|\omega_i), \quad i = 1, 2, \dots, M$$

Επομένως το πρόβλημα ανάγεται στον υπολογισμό l συναρτήσεων πυκνότητας πιθανότητας μίας διάστασης. Στην περίπτωση αυτή χρειάζονται lN δείγματα εκπαίδευσης και όχι N^l .

Ο Μπεϋζιανός ταξινομητής ο οποίος λειτουργεί βάσει του ανωτέρω πλαισίου, δηλαδή λειτουργεί θεωρώντας ότι όλα τα χαρακτηριστικά είναι υπό συνθήκη ανεξάρτητα μεταξύ τους, ονομάζεται αφελής (naïve). Ο αφελής Μπεϋζιανός ταξινομητής (Naive Bayes - NB) αποτελεί τον πιο απλό πιθανοτικό ταξινομητή. Παρά την αφελή σχεδίαση και τις υπεραπλουστευμένες υποθέσεις ανεξαρτησίας των χαρακτηριστικών, έχει αποδειχθεί ότι αποδίδει ιδιαίτερα καλά σε πολλά πραγματικά περίπλοκα προβλήματα ταξινόμησης, ενώ ταυτόχρονα είναι εύρωστος σε παραβιάσεις της υπόθεσης της ανεξαρτησίας των χαρακτηριστικών [250]. Ο H. Zhang [251] έδειξε πως υπάρχουν βάσιμοι θεωρητικοί λόγοι για την απίθανη αποτελεσματικότητα που επιδεικνύει.

Ο αφελής Μπεϋζιανός ταξινομητής παρουσιάζει πολλά πλεονεκτήματα έναντι άλλων μοντέλων ταξινόμησης, με το πιο βασικό την απλότητά του στην κατασκευή και τη χρήση του. Καταρχήν, αποτελεί έναν από τους λίγους μη-παραμετρικούς ταξινομητές με την έννοια ότι δεν διαθέτει παραμέτρους που πρέπει να επιλεγούν από τον χρήστη κατά τη κατασκευή του. Ως εκ τούτου δεν απαιτεί κάποιο ξεχωριστό σύνολο δεδομένων (σύνολο επικύρωσης) για την εύρεση των βέλτιστων τιμών των παραμέτρων του, όπως απαιτούν τα περισσότερα μοντέλα ταξινόμησης. Η ιδιότητα του αυτή είναι πολύ σημαντική σε προβλήματα όπου η επιλογή των παραμέτρων των ταξινομητών είναι δύσκολη ή περίπλοκη, όπως παραδείγματος χάριν σε προβλήματα με λίγα δεδομένα εκπαίδευσης όπου δεν υπάρχει η δυνατότητα δημιουργίας ξεχωριστού

συνόλου επικύρωσης. Πρακτικώς, ο αφελής Μπεϋζιανός ταξινομητής λειτουργεί ευθέως και δεν απαιτεί ειδική σχεδίαση για να κατασκευαστεί αρκεί μόνο η εκπαίδευσή του με τα δεδομένα εκπαίδευσης, διαδικασία κατά την οποία εκτιμώνται οι κατανομές των χαρακτηριστικών και οι εκ των προτέρων πιθανότητες των κλάσεων. Από τη στιγμή που υπολογισθούν οι τιμές αυτές, ο ταξινομητής μπορεί να προβεί στην ταξινόμηση οποιουδήποτε άγνωστου δείγματος εφαρμόζοντας απλά το θεώρημα του Bayes. Επιπλέον, εκπαιδεύεται πιο γρήγορα από ότι οι περισσότεροι αλγόριθμοι ταξινόμησης: για την εκπαίδευση του είναι αρκετό ένα μόνο πέρασμα των δεδομένων εκπαίδευσης (γραμμική πολυπλοκότητα $O(N)$, όπου N το πλήθος των δειγμάτων εκπαίδευσης). Άλλα σημαντικά πλεονεκτήματα είναι η ανθεκτικότητα που παρουσιάζει απέναντι στο θόρυβο (επειδή ακολουθεί πιθανοτική προσέγγιση) και η δυνατότητά του να χρησιμοποιηθεί σε προβλήματα με ελλιπή δεδομένα (όπου απλά παραλείπονται οι ελλειπείς τιμές στον υπολογισμό των κατανομών).

Η απλότητά του σε συνδυασμό με την αποτελεσματικότητά του έχουν καταστήσει τον αφελή Μπεϋζιανό ταξινομητή ένα από τα βασικά μοντέλα ταξινόμησης. Ειδικά σε προβλήματα με αποκλειστικά κατηγορικά δεδομένα στα οποία αποδίδει ιδιαίτερα καλά, μπορεί να χρησιμοποιηθεί ως ένα βασικό μοντέλο για τη σύγκριση διαφόρων ταξινομητών μεταξύ τους.

Ο NB είναι ένας μη παραμετρικός ταξινομητής, λειτουργεί άμεσα και δεν απαιτεί ιδιαίτερο σχεδιασμό για την κατασκευή του αρκεί μόνο η εκπαίδευσή του με το σύνολο δεδομένων εκπαίδευσης, μια διαδικασία κατά την οποία εκτιμώνται οι κατανομές των χαρακτηριστικών και οι εκ των προτέρων πιθανότητες των κλάσεων [252]. Για την κατασκευή του μοντέλου μας, χρησιμοποιήσαμε τη μέθοδο GaussianNB που υλοποιεί τον αλγόριθμο Gaussian NB για την ταξινόμηση, δηλαδή η πιθανότητα των συνεχών χαρακτηριστικών θεωρείται ότι ακολουθεί την κατανομή Gaussian.

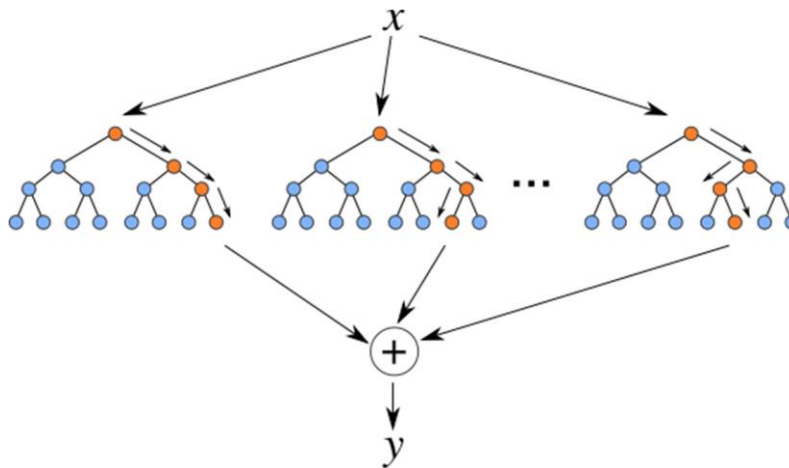
6.3.2 Τυχαία Δάση

Τα Τυχαία Δάση (Random Forests, RFs) αποτελούν έναν από τους καλύτερους και αποδοτικότερους ταξινομητές [253]. Ένα RF είναι ένας συλλογικός ταξινομητής δέντρων. Η βασική ιδέα της μεθόδου είναι η κατασκευή πολλών δέντρων ταξινόμησης (δάσος) και ο εν συνεχεία συνδυασμός των αποτελεσμάτων τους μέσω ενός συστήματος ψηφοφορίας. Το Random Forest είναι ένα σύνολο μη κλαδεμένων δέντρων ταξινόμησης, που δημιουργούνται από δείγματα bootstrap των δεδομένων εκπαίδευσης (δειγματοληψία με αντικατάσταση), χρησιμοποιώντας τυχαία επιλογή χαρακτηριστικών στη διαδικασία επαγωγής των δέντρων. Τα δέντρα εκπαιδεύονται ανεξάρτητα και οι προβλέψεις των δέντρων συνδυάζονται μέσω συνάθροισης (πλειοψηφική ψήφος για ταξινόμηση ή μέσος όρος για παλινδρόμηση) προκειμένου να παραχθεί η πρόβλεψη του συνόλου [254].

Σε ένα Τυχαίο Δάσος, κάθε δέντρο του κατασκευάζεται ως εξής [255]:

1. Ένα διαφορετικό τυχαίο υποσύνολο (random bootstrap) του συνόλου δεδομένων χρησιμοποιείται για την εκπαίδευση του δέντρου. Το υποσύνολο αυτό δημιουργείται από τυχαία επιλογή των $2/3$ των δειγμάτων του συνόλου δεδομένων. Το υπόλοιπο $1/3$ δεν χρησιμοποιείται για την εκπαίδευση του δέντρου, ενώ κρατείται εκτός για σκοπούς αξιολόγησης. Το τελευταίο ονομάζεται σύνολο out-of-bag (oob).
2. Εάν υπάρχουν διαθέσιμα M χαρακτηριστικά, τότε επιλέγονται τυχαία m ($m < M$) από αυτά για την κατασκευή του δέντρου.
3. Το δέντρο κατασκευάζεται σε βάθος, δηλαδή δεν λαμβάνει χώρα κλάδεμα.

Για την ταξινόμηση ενός δείγματος (x), αρχικά το δείγμα προωθείται σε όλα τα δέντρα και κάθε ένα από αυτά εξάγει το αποτέλεσμα ταξινόμησής του. Τα αποτελέσματα αυτά συνδυάζονται μέσω ενός συστήματος ψηφοφορίας προκειμένου να προκύψει το τελικό αποτέλεσμα ταξινόμησης του δάσους (y)



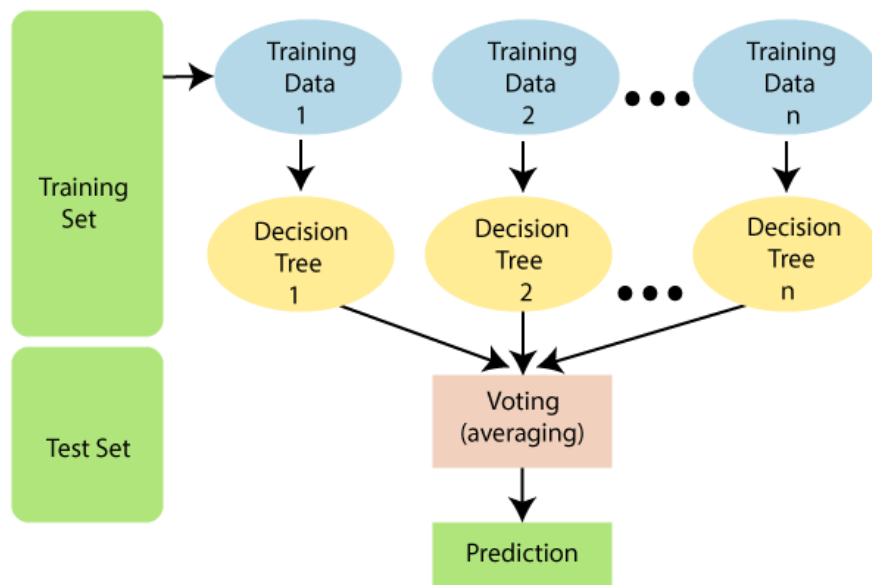
Εικόνα 23: Ταξινόμηση με τον αλγόριθμο Τυχαίων Δασών [199]

Τα out-of-bag δεδομένα μπορούν να χρησιμοποιηθούν για την εκτίμηση του σφάλματος γενίκευσης χωρίς την ανάγκη χρήσης ενός ξεχωριστού συνόλου ελέγχου (όπως για παράδειγμα απαιτείται στην περίπτωση των τεχνητών νευρωνικών δικτύων). Το out-of-bag error (oob error) είναι το σφάλμα ταξινόμησης των δειγμάτων τα οποία μένουν εκτός εκπαίδευσης των δέντρων κατά τη διαδικασία της δειγματοληψίας με αντικατάσταση. Εναλλακτικά, χρησιμοποιούνται για την εύρεση των παραμέτρων του ταξινομητή. Η μέθοδος διαθέτει δύο παραμέτρους: το πλήθος N των δέντρων που απαρτίζουν το δάσος και το πλήθος των μεταβλητών m . Ο Breiman [255] απέδειξε πως τα RFs δεν κινδυνεύουν από υπερεκπαίδευση όσο αυξάνεται το μέγεθος του δάσους (το πλήθος των δέντρων). Ως εκ τούτου, χρησιμοποιείται συνήθως ένα μεγάλο δάσος και δεν είναι αναγκαία η εύρεση του βέλτιστου πλήθους των δέντρων του δάσους. Ο μοναδικός λόγος για την εύρεση του βέλτιστου πλήθους των δέντρων που αποτελούν το δάσος είναι η μείωση της υπολογιστικής πολυπλοκότητας. Όσο μεγαλύτερο το δάσος, τόσο το RF παρουσιάζει αυξημένη

πολυπλοκότητα. Όσον αφορά στη μεταβλητή m , έχει διαπιστωθεί πως όταν το δάσος χρησιμοποιείται για ταξινόμηση, τότε μια ικανοποιητική τιμή της μεταβλητής αυτής είναι η τετραγωνική ρίζα του πλήθους των χαρακτηριστικών.

Η μέθοδος παρουσιάζει πολλά πλεονεκτήματα [199]:

- ✚ Το σφάλμα γενίκευσης είναι αρκετά περιορισμένο από τη στιγμή που αναπτύσσεται ένας πολύ μεγάλος αριθμός δέντρων.
- ✚ Η τυχαία επιλογή των χαρακτηριστικών μειώνει τη σχέση των μεγάλων και μη κλαδεμένων δένδρων, κάτι που κάνει την όλη μέθοδο αρκετά αμερόληπτη.
- ✚ Η χρήση του σφάλματος οοb επιτρέπει την εξαγωγή συμπερασμάτων για τη γενίκευση του μοντέλου χωρίς να είναι αναγκαία η χρήση ενός ξεχωριστού συνόλου ελέγχου.
- ✚ Δεν κινδυνεύουν από υπερεκπαίδευση.
- ✚ Διαχειρίζονται με φυσικό τρόπο δείγματα με ελλιπείς τιμές.
- ✚ Εξάγουν πραγματικές εκτιμήσεις εκ των υστέρων πιθανοτήτων.
- ✚ Μπορούν να εκπαιδευτούν αποτελεσματικά από σύνολα με μη ισορροπημένες κλάσεις.



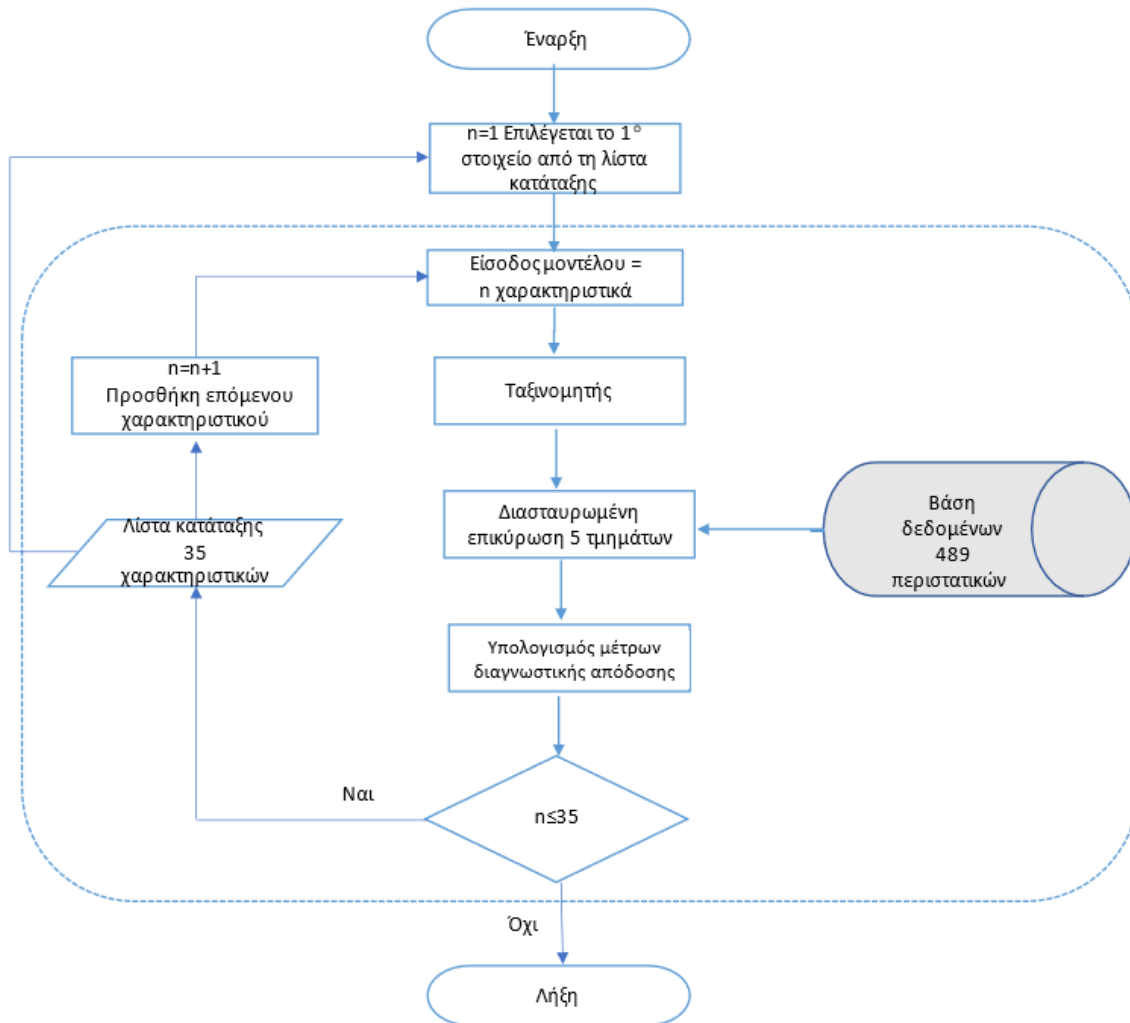
Εικόνα 24: Η λειτουργία του αλγορίθμου Random Forest [256]

Όσον αφορά στη διαχείριση των ελλειπουσών τιμών, εφαρμόζεται η τεχνική των εναλλακτικών διαδρομών (Surrogate Splits). Πρακτικά, όλα τα δέντρα εκπαιδεύονται υπολογίζοντας τις εναλλακτικές διαδρομές. Όταν ένα δέντρο διαθέτει έναν κόμβο ο οποίος ανταποκρίνεται σε χαρακτηριστικό που λείπει, τότε το δέντρο εξάγει το αποτέλεσμα του μέσω των Surrogate Splits. Επομένως, η διαχείριση των ελλειπουσών τιμών γίνεται στα δέντρα και όχι από το δάσος [199].

Τέλος, η ερευνητική ομάδα των Delgado et al.[257] χαρακτήρισε τον RF ως έναν από τους καλύτερους και αποδοτικότερους ταξινομητές αφού διεξήγαγε μια εκτενή συγκριτική μελέτη σχετικά με την απόδοση διαφόρων ταξινομητών και τα RFs επέδειξαν την καλύτερη απόδοση σε σύγκριση με σχεδόν 200 άλλους ταξινομητές, κατά μέσο όρο, σε περισσότερα από 100 σύνολα δεδομένων.

6.4 Αποτελέσματα Μοντέλων Ταξινόμησης

Ως είσοδος και στα δύο μοντέλα ταξινόμησης χρησιμοποιήθηκαν τα 35 χαρακτηριστικά του διανύσματος εισόδου με τη σειρά κατάταξης που προέκυψε μετά από την εφαρμογή των τεχνικών κατάταξης που αναφέρθηκαν στο υποκεφάλαιο 6.2 και ως έξοδος χρησιμοποιήθηκε το αποτέλεσμα της ιστολογικής εξέτασης που προέκυψε από την παθολογική ανάλυση του δείγματος του χειρουργείου και υποδηλώνει την πραγματική κατάσταση των γυναικών. Ξεκινώντας από το χαρακτηριστικό που βρίσκεται πρώτο στην εκάστοτε λίστα κατάταξης, υπολογίσαμε τα μέτρα της διαγνωστικής απόδοσης του μοντέλου προσθέτοντας ένα προς ένα τα χαρακτηριστικά, ξεκινώντας από το πιο σημαντικό και καταλήγοντας στο λιγότερο σημαντικό, και σε κάθε βήμα υπολογίσαμε την ακρίβεια (accuracy), την ευαισθησία (sensitivity) και την ειδικότητα (specificity) της διάγνωσης όπως περιγράφεται από το διάγραμμα ροής που παρουσιάζεται στην Εικόνα 25. Αναλυτικότερα, ως πρώτο στοιχείο ($n=1$), για την είσοδο στο μοντέλο ταξινόμησης, επιλέγεται το 1^ο (σημαντικότερο) στοιχείο της λίστας κατάταξης. Στη συνέχεια το μοντέλο χρησιμοποιεί τα 489 περιστατικά προκειμένου να εκπαιδευτεί και επαληθευτεί με τη μέθοδο επαναδειγματοληψίας της 5 πλής διασταυρούμενης επικύρωσης (5 - fold cross validation) για πρώτη φορά. Μετά τη δοκιμή του μοντέλου καταγράφεται η ακρίβεια της διάγνωσης που πραγματοποιεί και εφόσον το πλήθος των χαρακτηριστικών είναι μικρότερο του συνόλου ($n \leq 35$), πραγματοποιείται εν νέου κατασκευή του μοντέλου χρησιμοποιώντας ως είσοδο το πρώτο και το δεύτερο χαρακτηριστικό της λίστας. Η ίδια διαδικασία επαναλαμβάνεται έως ότου χρησιμοποιηθούν όλα τα χαρακτηριστικά του διανύσματος εισόδου. Πρόκειται ουσιαστικά για την κατασκευή των ταξινομητών με 35 διαφορετικά σύνολα χαρακτηριστικών, διαφορετικού μήκους και διαφορετικού περιεχομένου.



Εικόνα 25: Διάγραμμα ροής της κατασκευής των μοντέλων ταξινόμησης

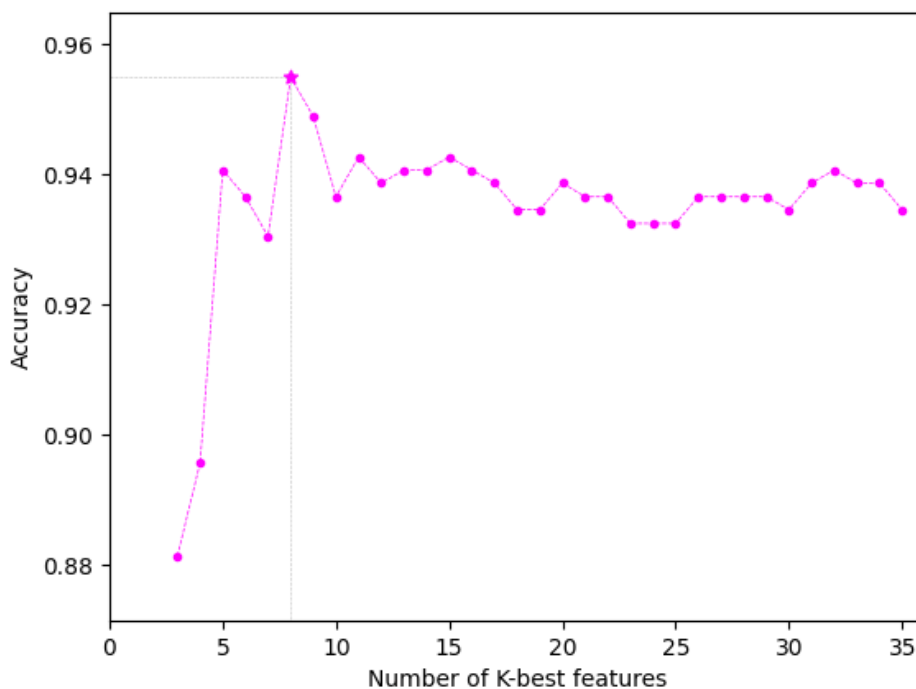
Η ίδια διαδικασία κατασκευής του μοντέλου ακολουθήθηκε και για τους δύο ταξινομητές. Ο ταξινομητής Naive Bayes αξιολογήθηκε με τη χρήση δυο διαφορετικών λιστών κατάταξης, όπως αυτές προέκυψαν στα υποκεφάλαια 6.2.1 και 6.2.2. Από την άλλη, ο ταξινομητής Random Forest έχει ενσωματωμένη την αξιολόγηση της σημαντικότητας των χαρακτηριστικών μέσω του δείκτη Gini Importance, οπότε για την κατασκευή του έγινε χρήση της δικής του λίστας κατάταξης αλλά και των δύο λιστών που προέκυψαν από την εφαρμογή των τεχνικών Information Gain και mRMR.

6.4.1 Αποτελέσματα Ταξινόμησης Αλγορίθμου Naive Bayes

Αναλυτικότερα, ο Naive Bayes δοκιμάστηκε χρησιμοποιώντας τις λίστες κατάταξης της μεθόδου Information Gain και της μεθόδου mRMR (βλ. Πίνακας 13). Τα μοντέλα κατασκευάστηκαν με τη διαδικασία που αναπαρίσταται στο διάγραμμα ροής (Εικόνα 25). Για κάθε διαφορετικό σύνολο των δεδομένων εισόδου, και μετά την επικύρωση του μοντέλου με 5 πλή διασταυρούμενη επικύρωση, εξήγαμε

τα μέτρα της διαγνωστικής απόδοσης που μας ενδιέφεραν προκειμένου να αξιολογήσουμε το μοντέλο. Συγκεκριμένα υπολογίστηκε διαδοχικά τόσο η ακρίβεια όσο και η ευαισθησία και η ειδικότητα του μοντέλου ταξινόμησης και για τις δυο λίστες κατάταξης. Μετά την δοκιμή και των δύο λιστών κατάταξης, καταλήξαμε στη χρήση της λίστας που επιτυγχάνει τη βέλτιστη διαγνωστική απόδοση του μοντέλου με το μικρότερο δυνατό αριθμό χαρακτηριστικών. Τα αποτελέσματα της χρήσης της τεχνικής επιλογής χαρακτηριστικών Information Gain ήταν τα επικρατέστερα, όπως παρατίθενται αναλυτικά στη συνέχεια, ενώ τα αποτελέσματα από τη δοκιμή της λίστας mRMR παρατίθενται στο Παράρτημα Β.

Στην Εικόνα 26 βλέπουμε τη γραφική παράσταση που αντιπροσωπεύει τη διαγνωστική ακρίβεια του ταξινομητή Naive Bayes και για τα 35 διαφορετικά διανύσματα χαρακτηριστικών εισόδου (Εικόνα 25). Κάθε σημείο της γραφικής παράστασης αναπαριστά την ακρίβεια του ταξινομητή έχοντας ως είσοδο το εκάστοτε σύνολο χαρακτηριστικών. Σκοπός μας είναι να εντοπίσουμε εκείνο το σημείο όπου μεγιστοποιείται η διαγνωστική απόδοση του μοντέλου με το μικρότερο δυνατό σύνολο χαρακτηριστικών.



Εικόνα 26: Διαγνωστική ακρίβεια του μοντέλου ταξινόμησης Naive Bayes σύμφωνα με τη λίστα κατάταξης των χαρακτηριστικών εισόδου με χρήση της λίστας Information Gain

Όπως φαίνεται στην ανωτέρω εικόνα (Εικόνα 26), όταν χρησιμοποιείται η λίστα κατάταξης, που προέκυψε με τη μέθοδο του Information Gain, ως είσοδος στο μοντέλο ταξινόμησης Naive Bayes, η ακρίβεια της διάγνωσης μεγιστοποιείται όταν εισάγονται στο μοντέλο ταξινόμησης τα πρώτα 8 χαρακτηριστικά της

λίστας. Με βάση τη λίστα κατάταξης που παρατίθενται στον Πίνακα 13 διαπιστώνουμε ότι τα πρώτα 8 χαρακτηριστικά εισόδου που μεγιστοποιούν τη διαγνωστική ακρίβεια του ταξινομητή Naive Bayes είναι τα χαρακτηριστικά του Πίνακα 14.

Πίνακας 14: Βέλτιστο υποσύνολο χαρακτηριστικών για το μοντέλο ταξινόμησης Naive Bayes

Σειρά κατάταξης Όνομα χαρακτηριστικού
(σημαντικότητα)

1	Tumor Shape
2	Tumor Margins (bn)
3	Arc_Disturbance (bn)
4	New Finding (bn)
5	Vascularity (bn)
6	Lymph Nodes
7	Age
8	Breast size

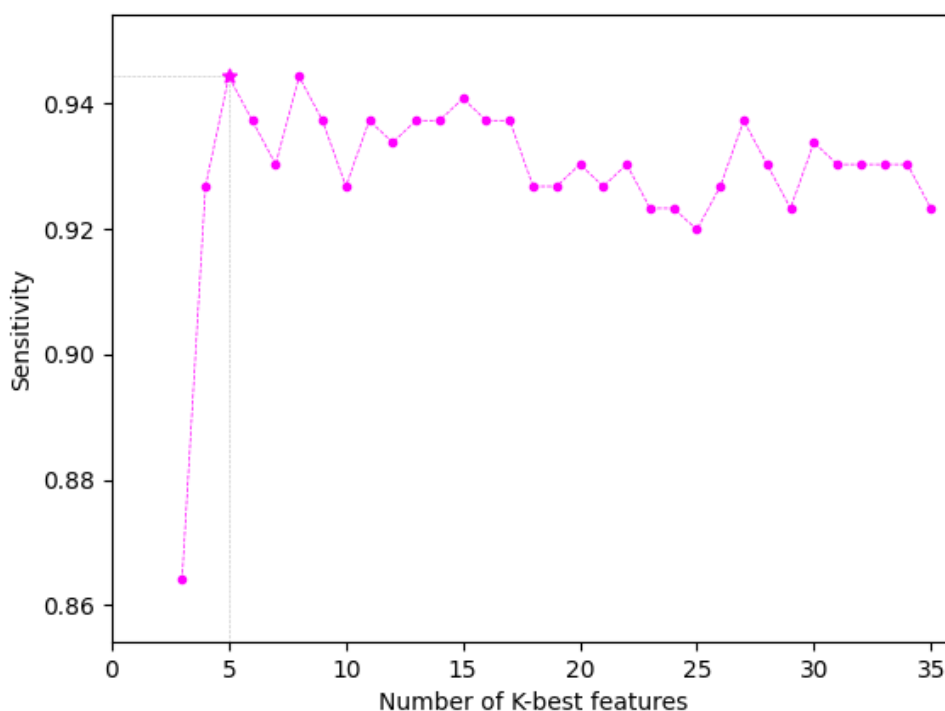
Στη συνέχεια, παρατίθεται ο πίνακας σύγκρισης του προέκυψε κατά τη δοκιμή του ταξινομητή που κατασκευάστηκε με τα προαναφερθέντα 8 σημαντικότερα χαρακτηριστικά.

Πίνακας 15: Πίνακας σύγκρισης του ταξινομητή Naive Bayes για τα 8 καλύτερα χαρακτηριστικά με βάση το Information Gain

Naive Bayes με IG	Θετικό ιστολογικό αποτέλεσμα	Αρνητικό ιστολογικό αποτέλεσμα
Ταξινομήθηκε ως θετικό	271	6
Ταξινομήθηκε ως αρνητικό	16	196

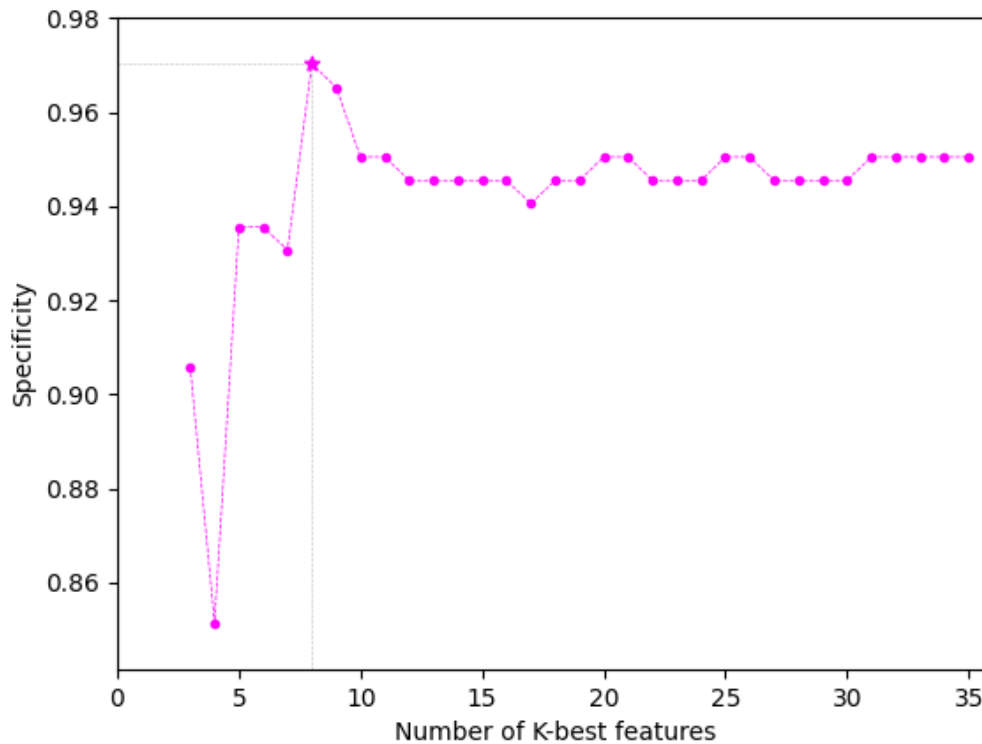
Έπειτα, παρατίθενται τα αποτελέσματα για την ευαισθησία (sensitivity) και την ειδικότητα (specificity) του μοντέλου ταξινόμησης Naive Bayes και για τα 35 διαφορετικά διανύσματα χαρακτηριστικών εισόδου, όπως αυτά προέκυψαν από τη λίστα κατάταξης και σύμφωνα με τη διαδικασία του διαγράμματος ροής (Εικόνα 25). Κάθε σημείο της γραφικής παράστασης αναπαριστά την ευαισθησία και την ειδικότητα του ταξινομητή έχοντας ως είσοδο το εκάστοτε σύνολο χαρακτηριστικών.

Όπως βλέπουμε στην Εικόνα 27, η ευαισθησία ταξινόμησης για τα μόλις 5 πρώτα χαρακτηριστικά της λίστας ανέρχεται στο 94,4%, και εντοπίζουμε ότι διατηρεί την τιμή αυτή για τα 8 επιλεγμένα χαρακτηριστικά που προέκυψαν ως το βέλτιστο υποσύνολο για την κατασκευή του μοντέλου ταξινόμησης (βλ. Πίνακας 14: Βέλτιστο υποσύνολο χαρακτηριστικών για το μοντέλο ταξινόμησης Naive Bayes Πίνακας 14).



Εικόνα 27: Ευαισθησία του μοντέλου ταξινόμησης Naive Bayes σύμφωνα με τη λίστα κατάταξης των χαρακτηριστικών εισόδου με χρήση της λίστας Information Gain

Στη συνέχεια, στο διάγραμμα που ακολουθεί παρατηρούμε ότι η ειδικότητα της ταξινόμησης μεγιστοποιείται για τα 8 πρώτα χαρακτηριστικά της λίστας κατάταξης, αγγίζοντας το 97%, γεγονός που πιστοποιεί την επιλογή των συγκεκριμένων χαρακτηριστικών ως βέλτιστο υποσύνολο χαρακτηριστικών του μοντέλου ταξινόμησης Naive Bayes.



Εικόνα 28: Ειδικότητα του μοντέλου ταξινόμησης Naive Bayes σύμφωνα με τη λίστα κατάταξης των χαρακτηριστικών εισόδου με χρήση της λίστας Information Gain

Συμπερασματικά, αφού λάβαμε υπόψιν τα ανωτέρω γραφήματα και καταλήξαμε στο βέλτιστο υποσύνολο των 8 χαρακτηριστικών (βλ. Πίνακας 14), που προέκυψε από τη λίστα κατάταξης της μεθόδου Information Gain, υπολογίσαμε τα βασικά μέτρα αξιολόγησης της διαγνωστικής απόδοσης για το μοντέλο ταξινόμησης Naive Bayes έχοντας ως είσοδο το συγκεκριμένο σύνολο χαρακτηριστικών. Τα αποτελέσματα παρουσιάζονται στον Πίνακας 16 που ακολουθεί.

Πίνακας 16: Διαγνωστική απόδοση του μοντέλου ταξινόμησης Naive Bayes με χρήση του Information Gain και με το βέλτιστο υποσύνολο χαρακτηριστικών

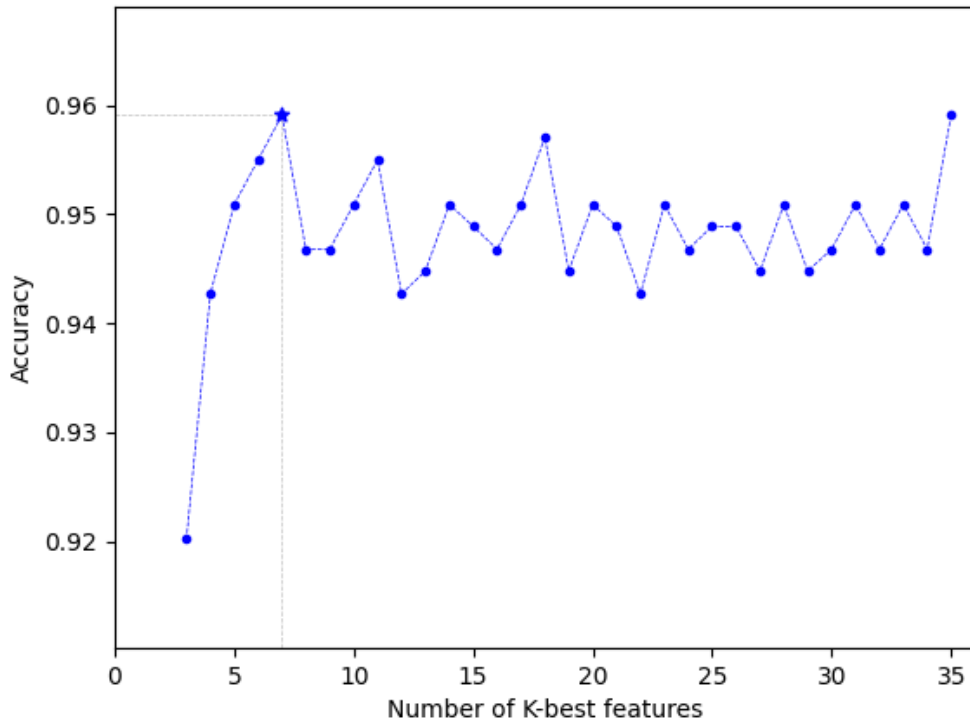
Naive Bayes με χρήση IG	
Accuracy (%)	95,5
Sensitivity (%)	94,4
Specificity (%)	97,0
PPV (%)	97,8

<i>NPV (%)</i>	92,4
<i>Youden's index</i>	0,91
<i>F1 score</i>	0,96

6.4.2 Αποτελέσματα Ταξινόμησης Αλγορίθμου Random Forest

Στη συνέχεια, ο αλγόριθμος ταξινόμησης Random Forest δοκιμάστηκε χρησιμοποιώντας τις λίστες κατάταξης της μεθόδου Information Gain, της μεθόδου mRMR και της, ενσωματωμένης στον αλγόριθμο, μεθόδου Gini Importance (βλ. Πίνακας 13). Τα μοντέλα κατασκευάστηκαν επίσης με τη διαδικασία που αναπαρίσταται στο διάγραμμα ροής (Εικόνα 25). Και σε αυτή την περίπτωση υπολογίστηκε διαδοχικά τόσο η ακρίβεια όσο και η ευαισθησία και η ειδικότητα του μοντέλου ταξινόμησης και για τις τρεις λίστες κατάταξης. Μετά την δοκιμή και των τριών λιστών κατάταξης, καταλήξαμε στη χρήση της λίστας που επιτυγχάνει τη βέλτιστη διαγνωστική απόδοση του μοντέλου με το μικρότερο δυνατό αριθμό χαρακτηριστικών. Τα αποτελέσματα της χρήσης της τεχνικής επιλογής χαρακτηριστικών mRMR ήταν τα επικρατέστερα, όπως παρατίθενται αναλυτικά στη συνέχεια, ενώ τα αποτελέσματα από τη δοκιμή της λίστας Information Gain και Gini importance παρατίθενται στο Παράρτημα Β.

Στην Εικόνα 29 βλέπουμε τη γραφική παράσταση που αντιπροσωπεύει τη διαγνωστική ακρίβεια του ταξινομητή Random Forest και για τα 35 διαφορετικά διανύσματα χαρακτηριστικών εισόδου, όπως αυτά προέκυψαν από τη λίστα κατάταξης του αλγορίθμου mRMR.



Εικόνα 29: Διαγνωστική ακρίβεια του μοντέλου Random Forest σύμφωνα με τη λίστα κατάταξης κατάταξης των χαρακτηριστικών εισόδου με χρήση της λίστας mRMR

Από την Εικόνα 29 διαπιστώνουμε ότι ο ταξινομητής Random Forest μεγιστοποιεί την ακρίβεια διάγνωσης όταν εισέρχεται και το 7^ο στη σειρά κατάταξης χαρακτηριστικό. Με βάση τη λίστα κατάταξης που παρατίθενται στον Πίνακα 13 διαπιστώνουμε ότι τα πρώτα 7 χαρακτηριστικά εισόδου που μεγιστοποιούν τη διαγνωστική ακρίβεια του ταξινομητή είναι τα εξής:

Πίνακας 17: Βέλτιστο υποσύνολο χαρακτηριστικών για το μοντέλο ταξινόμησης Random Forest

Σειρά κατάταξης Όνομα χαρακτηριστικού
(σημαντικότητα)

1	Tumor Margins (bn)
2	Tumor Shape
3	Vascularity (bn)

4	Arc_Disturbance (bn)
5	New Finding (bn)
6	Breast Size
7	Age

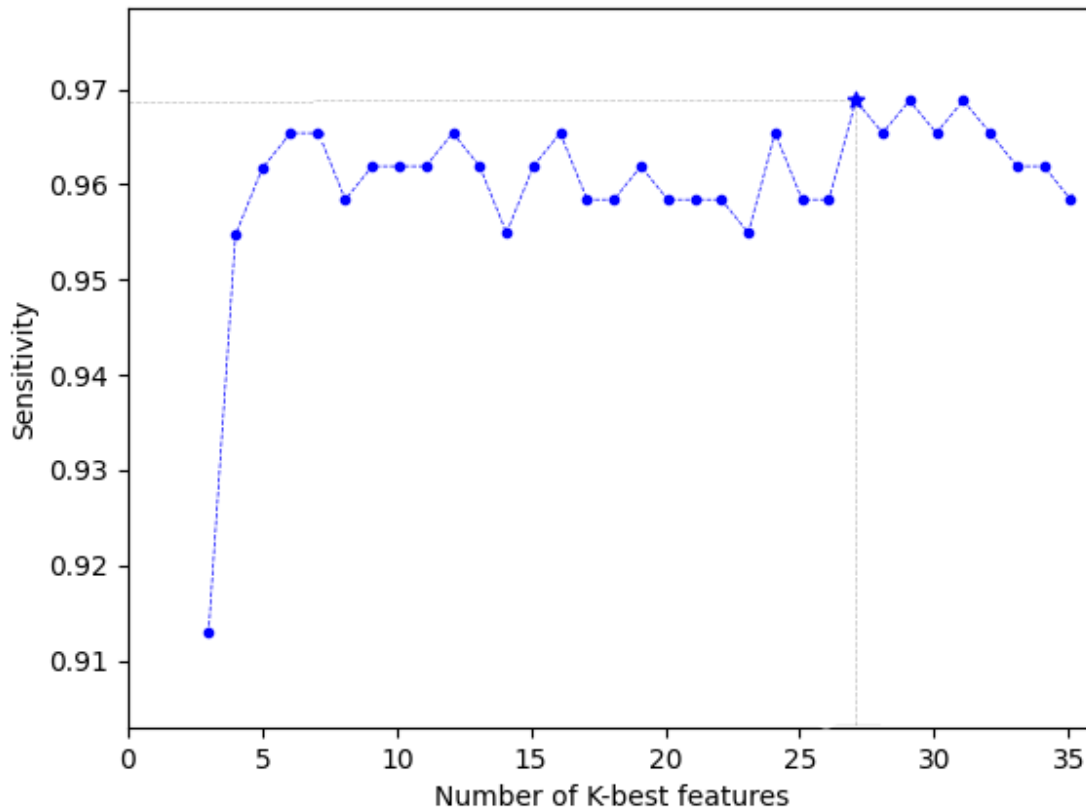
Στον πίνακα που ακολουθεί, παρουσιάζουμε τα αποτελέσματα του πίνακα σύγκρισης και των βασικών μέτρων αξιολόγησης της διαγνωστικής απόδοσης του ταξινομητή Random Forest που προέκυψαν για τα 7 καλύτερα χαρακτηριστικά εισόδου με βάση τα αποτελέσματα του mRMR (Εικόνα 29).

Πίνακας 18: Πίνακας σύγκρισης του ταξινομητή Random Forest για τα καλύτερα 7 χαρακτηριστικά

Random Forest	Θετικό ιστολογικό αποτέλεσμα	Αρνητικό ιστολογικό αποτέλεσμα
Ταξινομήθηκε ως θετικό	277	10
Ταξινομήθηκε ως αρνητικό	10	192

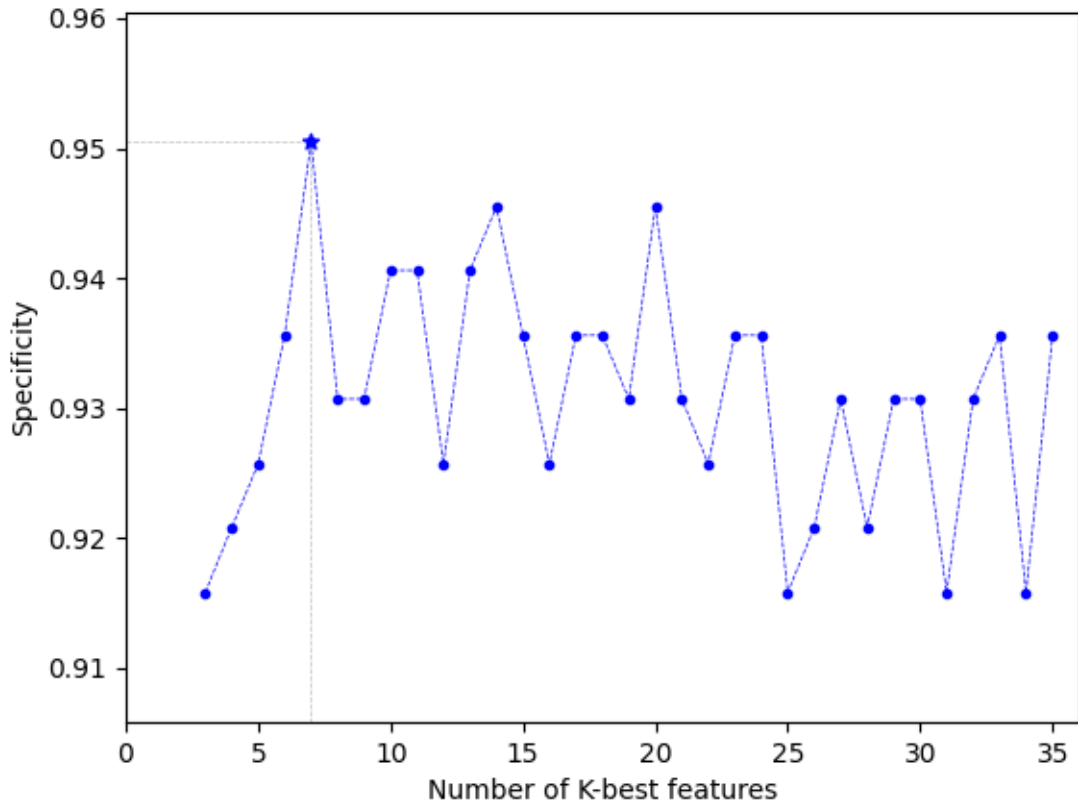
Στη συνέχεια, παρατίθενται τα αποτελέσματα για την ευαισθησία (sensitivity) και την ειδικότητα (specificity) του μοντέλου ταξινόμησης Random Forest και για τα 35 διαφορετικά διανύσματα χαρακτηριστικών εισόδου, όπως αυτά προέκυψαν από τη λίστα κατάταξης του mRMR. Κάθε σημείο της γραφικής παράστασης αναπαριστά την ευαισθησία και την ειδικότητα του ταξινομητή έχοντας ως είσοδο το εκάστοτε σύνολο χαρακτηριστικών.

Όπως βλέπουμε στην Εικόνα 30, η ευαισθησία ταξινόμησης για τα μόλις 5 πρώτα χαρακτηριστικά της λίστας φαίνεται να αγγίζει το 95,5 %, και να φτάνει στην τιμή 96,5 % στα πρώτα 6 και 7 πρώτα χαρακτηριστικά της λίστας. Η ευαισθησία της διάγνωσης του ταξινομητή RF παραμένει σε υψηλές τιμές για όλα τα υποσύνολα χαρακτηριστικών που ακολουθούν ως είσοδοι στο μοντέλο ταξινόμησης, φτάνοντας στη μέγιστη τιμή 96,9% για το υποσύνολο των 27 χαρακτηριστικών.



Εικόνα 30: Ευαισθησία του μοντέλου ταξινόμησης *Random Forest* σύμφωνα με τη λίστα κατάταξης των χαρακτηριστικών εισόδου με χρήση του *mRMR*

Αντίστοιχα, όσον αφορά στην ειδικότητα της διάγνωσης του ταξινομητή, όπως βλέπουμε στο διάγραμμα που ακολουθεί (βλ. Εικόνα 31), το μοντέλο ταξινόμησης εμφανίζει τη μέγιστη ειδικότητα για το υποσύνολο των 7 πρώτων χαρακτηριστικών (βλ. Πίνακας 17) και διατηρεί υψηλές τιμές (>92%) για όλα τα υποσύνολα χαρακτηριστικών που αποτελούν είσοδοι στο μοντέλο.



Εικόνα 31: Ειδικότητα του μοντέλου ταξινόμησης Random Forest σύμφωνα με τη λίστα κατάταξης των χαρακτηριστικών εισόδου με χρήση του mRMR

Συμπερασματικά, αφού λάβαμε υπόψιν τα ανωτέρω γραφήματα και καταλήξαμε στο βέλτιστο υποσύνολο των 7 χαρακτηριστικών (βλ. Πίνακας 17), που προέκυψε από τη λίστα κατάταξης της μεθόδου mRMR, υπολογίσαμε τα βασικά μέτρα αξιολόγησης της διαγνωστικής απόδοσης για το μοντέλο ταξινόμησης Naive Bayes έχοντας ως είσοδο το συγκεκριμένο σύνολο χαρακτηριστικών. Τα αποτελέσματα παρουσιάζονται στον Πίνακα 19 που ακολουθεί.

Πίνακας 19: Διαγνωστική απόδοση του ταξινομητή Random Forest με χρήση του mRMR και με το βέλτιστο υποσύνολο χαρακτηριστικών

	Random Forest
Accuracy (%)	95,9
Sensitivity (%)	96,5

<i>Specificity (%)</i>	95,1
<i>PPV (%)</i>	96,5
<i>NPV (%)</i>	95
<i>Youden's index</i>	0,92
<i>F1 score</i>	0,97

6.5 Συζήτηση - Συμπεράσματα

Λαμβάνοντας υπόψη τα παραπάνω αποτελέσματα (Πίνακας 16 & Πίνακας 19), συγκρίναμε τους ταξινομητές που κατασκευάσαμε, στα σημεία της μέγιστης απόδοσής τους δηλαδή για το μικρότερο σύνολο χαρακτηριστικών όπου εντοπίσαμε ότι μεγιστοποιείται η διαγνωστική τους απόδοση ή αλλιώς το βέλτιστο υποσύνολο χαρακτηριστικών. Ένα από τα συμπεράσματα και σημαντικότερα αποτελέσματα της παρούσας διδακτορικής διατριβής αφορά στο βέλτιστο υποσύνολο χαρακτηριστικών, στο οποίο καταλήξαμε με τη βοήθεια των εργαλείων μηχανικής μάθησης, και παρέχει τη μεγαλύτερη διαγνωστική πληροφορία. Όπως φαίνεται από τον Πίνακα 14 & τον Πίνακα 17, τα δύο υποσύνολα που προέκυψαν έχουν ως επί το πλείστον κοινά χαρακτηριστικά υποδεικνύοντας την καθολική σημασία του συνδυασμού των εν λόγω χαρακτηριστικών για την παραγωγή των βέλτιστων αποτελεσμάτων. Τα αποτελέσματα αυτά μπορούν να οδηγήσουν στην καλύτερη διαίσθηση για το πραγματικό πρόβλημα, δίνοντας έτσι τη δυνατότητα στους ειδικούς του τομέα να το αντιμετωπίσουν πιο αποτελεσματικά στο μέλλον.

Για τον Naive Bayes αξιολογήθηκε το μοντέλο που κατασκευάστηκε με τον εν λόγω αλγόριθμο για το βέλτιστο υποσύνολο των 8 χαρακτηριστικών, όπως αυτά προέκυψαν από τη χρήση του δείκτη Information Gain, και τον Random Forest αντίστοιχα αξιολογήθηκε το μοντέλο με το βέλτιστο σύνολο των 7 χαρακτηριστικών όπως προέκυψαν από τη χρήση του δείκτη mRMR. Τα συγκεντρωτικά αποτελέσματα των μέτρων διαγνωστικής απόδοσης των μοντέλων ταξινόμησης που κατασκευάστηκαν παρουσιάζονται στον Πίνακα 20.

Πίνακας 20: Συγκεντρωτικά αποτελέσματα της διαγνωστικής απόδοσης των επιλεγμένων ταξινομητών

Naive Bayes με χρήση Information Gain	Random Forest με χρήση mRMR
---------------------------------------------	--------------------------------

<i>Accuracy (%)</i>	95,5	95,9
<i>Sensitivity (%)</i>	94,4	96,5
<i>Specificity (%)</i>	97,0	95,1
<i>PPV (%)</i>	97,8	96,5
<i>NPV (%)</i>	92,4	95
<i>Youden's index</i>	0,91	0,92
<i>F1 score</i>	0,96	0,97

Όπως βλέπουμε στον παραπάνω πίνακα, ο ταξινομητής RF έδειξε την υψηλότερη διαγνωστική ακρίβεια 95,9%, δείχνοντας πόσο αποτελεσματικά λειτουργεί στο συνδυασμό δυαδικών, κατηγορικών και αριθμητικών μεταβλητών. Ο RF έδειξε τις υψηλότερες τιμές στα μέτρα ευαισθησίας (96,2%), NPV (95%), Youden's index (0,92) και F1 score (0,97). Αντίθετα, οι τιμές της ειδικότητας (95,1%) και της PPV (96,5%) είναι χαμηλότερες για τον ταξινομητή RF σε σχέση με το Naive Bayes. Σε αυτό το σημείο είναι σημαντικό να τονίσουμε ότι είναι εξαιρετικής σημασίας για έναν ταξινομητή να παρουσιάζει ισορροπία στα διαγνωστικά αποτελέσματα, κριτήριο που, όπως φαίνεται από τους πίνακες σύγχυσης των ταξινομητών (βλ. Πίνακας 15 & Πίνακας 18), πληρείται καλύτερα από τον Random Forest καθώς ο αριθμός των ψευδώς θετικών (10) και ψευδώς αρνητικών αποτελεσμάτων (10) είναι ο ίδιος. Επιπλέον, όπως γνωρίζουμε η ευαισθησία της διάγνωσης ενός ταξινομητή (ή ενός τεστ) εκφράζει την ικανότητα του να αναγνωρίζει σωστά τους ασθενείς δηλαδή αντικατοπτρίζει το ποσοστό των πραγματικών ασθενών που σωστά έχουν προσδιοριστεί από τον ταξινομητή ως ασθενείς. Δίνουμε ιδιαίτερη σημασία στο εν λόγω διαγνωστικό μέτρο καθώς είναι ύψιστης σημασίας να μη «χάνεται», από το σύστημα μας, κάποια περίπτωση υψηλού κινδύνου που ελλοχεύει κακοήθεια. Σημαντικοί στατιστικοί δείκτες είναι επίσης ο δείκτης Youden (0,92) και το F1 score (0,97). Ο δείκτης Youden ενσωματώνει τις πληροφορίες της ευαισθησίας και της ειδικότητας, με την υψηλότερη τιμή να υποδηλώνει σημαντική διαγνωστική απόδοση και ως προς τα δύο αποτελέσματα (θετικά και αρνητικά). Όσον αφορά στο F1 score (0,97), είναι ένας τρόπος συνδυασμού της PPV και της ευαισθησίας του μοντέλου και ορίζεται ως ο αρμονικός μέσος όρος αυτών των δύο τιμών, με την υψηλότερη τιμή να υποδηλώνει την εξαιρετική ικανότητα του μοντέλου να αναγνωρίζει τα πραγματικά θετικά αποτελέσματα. Για όλους τους ανωτέρω λόγους, καταλήξαμε στην επιλογή του ταξινομητή Random Forest για τη δημιουργία του καινοτόμου μοντέλου CDSS.

Ο ταξινομητής προσφέρει εξαιρετικά μειωμένα ψευδώς αρνητικά και ψευδώς θετικά αποτελέσματα σε σύγκριση με το διαγνωστικό εργαλείο της μαστογραφίας. Συγκεκριμένα, με βάση τον πίνακα σύγκρισης του RF (Πίνακας 18), διαπιστώνουμε ότι υπήρχαν μόνο 10 ψευδώς θετικά και 10 ψευδώς αρνητικά αποτελέσματα στο σύνολο των 489 περιπτώσεων που εξετάστηκαν με τη μέθοδο 5-fold cross validation. Στον πίνακα που ακολουθεί (Πίνακας 21) παραθέτουμε τα αποτελέσματα των μέτρων της διαγνωστικής απόδοσης της μαστογραφίας, όπως αυτά προέκυψαν με τη χρήση διαφορετικών κατώφλιων θετικότητας (βλ. Πίνακας 12) και τα αποτελέσματα της διαγνωστικής απόδοσης του επικρατέστερου ταξινομητή (RF) που προέκυψε από τις ανωτέρω διαδικασίες.

Πίνακας 21: Συγκριτικός πίνακας για τη διαγνωστική απόδοση του ταξινομητή RF σε σχέση με τη μαστογραφία*

	Μαστογραφία (κατώφλι BI- RADS 4a)	Μαστογραφία (κατώφλι BI- RADS 4b)	Μαστογραφία (κατώφλι BI- RADS 4c)	Random Forest
<i>Accuracy (%)</i>	78,0	81,0	74,8	95,9
<i>Sensitivity (%)</i>	77,9	73,1	59,9	96,5
<i>Specificity (%)</i>	78,1	94,4	98,6	95,1
<i>PPV (%)</i>	88,5	95,4	98,5	96,5
<i>NPV (%)</i>	68,9	68,7	59	95
<i>Youden's index</i>	0,56	0,58	0,59	0,92
<i>F1 score</i>	0,82	0,83	0,74	0,97

*Για τον υπολογισμό των μέτρων της διαγνωστικής απόδοσης της μαστογραφίας δεν χρησιμοποιήθηκαν τα περιστατικά με BI-RADS 0 στη μαστογραφία καθώς δεν παρέχουν καμία πληροφορία εκτίμησης κινδύνου.

Όπως βλέπουμε, το μοντέλο ταξινόμησης Random Forest επέδειξε μεγαλύτερες τιμές σε ακρίβεια (Accuracy) και ευαισθησία (Sensitivity) με συντριπτική διαφορά σε σχέση με τη μαστογραφία. Η μόνη τιμή στην οποία υπερίσχυσε η μαστογραφία είναι η ειδικότητα (Specificity) με ποσοστό 98,5% (κατώφλι BI-RADS 4c) έναντι του 95,1% του RF. Εντούτοις, παρατηρώντας την αντίστοιχη τιμή της ευαισθησίας,

σε αυτή την περίπτωση (59,9%) διαπιστώνεται εξαιρετικά μικρή σε σύγκριση με την ευαισθησία του ταξινομητή (96,5%), γεγονός που υποδηλώνει την αδυναμία της μαστογραφίας να εντοπίζει τους καρκίνους σε αντίθεση με τον ταξινομητή RF. Η διαγνωστική αδυναμία της μαστογραφίας στον εντοπισμό των πραγματικών περιστατικών με κακοήθεια διαφαίνεται από όλους τους αντίστοιχους στατιστικούς δείκτες και αναδεικνύει την μεγάλη προστιθέμενη αξία του μοντέλου ταξινόμησης που παρουσιάζεται στην παρούσα διδακτορική διατριβή στη διαγνωστική διαδικασία.

Κεφάλαιο 7^ο

Το πρόβλημα BI-RADS 0 και η Αντιμετώπισή του μέσω των εργαλείων Μηχανικής Μάθησης

7.1 Το Πρόβλημα BI-RADS 0

Όπως προαναφέρθηκε, σύμφωνα με την κλίμακα αξιολόγησης των μαστογραφικών ευρημάτων [98], η κατηγορία BI-RADS 0 αποδίδεται στις περιπτώσεις όπου η εξέταση θεωρείται ανεπαρκής γεγονός που δεν επιτρέπει στους ακτινολόγους να εξάγουν κάποιο ασφαλές συμπέρασμα. Παρά το γεγονός ότι η σύσταση που ακολουθεί την εν λόγω αξιολόγηση περιλαμβάνει επιπρόσθετες απεικονιστικές εκτιμήσεις και εξετάσεις, πληθώρα δημοσιευμένων ερευνών αναφέρει ότι η κατηγορία BI-RADS 0 χρησιμοποιείται σε πολλές περιπτώσεις με ασυνέπεια κατά τη διαγνωστική διαδικασία, αφού συχνά αντιμετωπίζεται ως τελική κατηγορία αξιολόγησης [258].

Στη μεγάλη έρευνα των Porlack et al. [259] αποδεικνύεται η χρήση της κατηγορίας BI-RADS 0 ως τελική κατηγορία της διαγνωστικής διαδικασίας παρά τις προαναφερθείσες κατευθυντήριες οδηγίες [98]. Στις περιπτώσεις αυτές οι ακτινολόγοι συχνά προβαίνουν σε βιοψία, όπως στη μελέτη των Geller et al. [258] όπου για το 20% των περιπτώσεων με BI-RADS 0 προτάθηκε περαιτέρω κλινική διερεύνηση και διενέργεια βιοψίας. Παρά το γεγονός ότι οι διεθνείς κατευθυντήριες γραμμές προτείνουν τη διενέργεια πρόσθετων απεικονιστικών εξετάσεων στις περιπτώσεις ενός αποτελέσματος BI-RADS 0 [98], προκειμένου να γίνεται πιο έγκυρη διάγνωση, πολλές φορές την εν λόγω αξιολόγηση ακολουθεί η διενέργεια βιοψίας [258].

Στη μεγάλη έρευνα των Rosenberg et al. [260], που περιελάμβανε τον έλεγχο 136.540 μαστογραφιών γυναικών από το Μεξικό, αναφέρεται ότι βρέθηκε ποσοστό 10,7% (121 από 1.132) του συνόλου των γυναικών με διαγνωσμένο καρκίνου του μαστού, στις οποίες είχε αποδοθεί η αξιολόγηση BI-RADS 0 στη μαστογραφία τους κατά τον προσυμπτωματικό έλεγχο. Το γεγονός αυτό ενισχύει την άποψη ότι πρόκειται για μια ύπουλη κατηγορία καθώς υποκρύπτει τον κίνδυνο ύπαρξης καρκίνου του μαστού.

Ένα ακόμη κρίσιμης σημασίας ζήτημα επισημαίνεται στη μεγάλη έρευνα των Geller et al. [258], όπου υπογραμμίζεται ότι η αξιολόγηση της μαστογραφίας μιας ασθενούς με την ειδική κατηγορία BI-RADS 0 επηρεάζει σημαντικά τη συμμόρφωση των ασθενών με τις οδηγίες που αφορούν στον επανέλεγχό τους (follow up). Είναι γεγονός ότι ενδεχόμενα λάθη στην καθοδήγηση, μετά το εν λόγω αποτέλεσμα στη

μαστογραφία, θα μπορούσε να οδηγήσει σε μη έγκαιρο επανέλεγχο με επιβαρυντικές συνέπειες για την υγεία των ασθενών. Οι Geller et al. [258] τονίζουν επίσης ότι στις περιπτώσεις κλινικής ασάφειας θα ήταν χρήσιμο να υπάρχει μια πιο σαφής και ρητή τεκμηρίωση με βάση την κλίμακα BI-RADS, με την έννοια ότι ίσως θα ήταν πιο συνεπές με την κλινική διαδικασία να γίνεται προσπάθεια απόδοσης του περιστατικού σε κάποια άλλη από τις κατηγορίες αξιολόγησης BI-RADS με την βοήθεια περαιτέρω απεικονιστικών διαδικασιών. Παρ' όλα αυτά υπάρχουν πολλές περιπτώσεις που κάτι τέτοιο δεν είναι εφικτό.

Η χρήση περαιτέρω απεικονιστικών τεχνικών είναι ένας αποτελεσματικός τρόπος για τη βελτίωση της διαγνωστικής απόδοσης στις περιπτώσεις των μαστογραφικά απροσδιόριστων αλλοιώσεων του μαστού. Όπως φαίνεται στη μελέτη των Chunyan Yi et al. [261], προκειμένου να διασαφηνιστούν οι περιπτώσεις με αξιολόγηση BI-RADS 0 στη μαστογραφία, το 17,28% των ασθενών υποβλήθηκε σε πρόσθετη εξέταση MRI, το 22,61% των ασθενών σε δύο ή περισσότερες εξετάσεις παρακολούθησης (follow-up), ενώ το 70,71% των ασθενών υπεβλήθη σε βιοψίες όπου αποδείχθηκε καλοήθης ή φυσιολογικό εύρημα. Ωστόσο, οι προαναφερθείσες διαδικασίες ενδέχεται να αυξήσουν το άγχος και το κόστος είτε του συστήματος υγείας είτε των ίδιων των ασθενών.

Ένας σημαντικός παράγοντας που σύμφωνα με τους Chunyan Yi et al. [261], επηρεάζει σημαντικά την αξιολόγηση των ευρημάτων της μαστογραφίας αφορά στην εμπειρία των ακτινολόγων, ειδικά όταν πρόκειται για τις ασαφείς περιπτώσεις BI-RADS 0 δεδομένου ότι δεν υπάρχει σαφές πρότυπο για την αξιολόγηση αυτών των περιστατικών. Η κατάχρηση της αξιολόγησης του BI-RADS 0 (ιδίως από αρχάριους ή άπειρους ακτινολόγους) έχει αποδειχθεί ότι αυξάνει την ψυχολογική και ενδεχομένως την οικονομική επιβάρυνση των ασθενών. Όπως διαφαίνεται από τα ανωτέρω προβλήματα που μπορεί να προκύψουν από την αξιολόγηση ευρημάτων με BI-RADS 0, υπάρχει ανάγκη για διερεύνηση των περιπτώσεων BI-RADS 0 μέσω πολλαπλών αναγνώσεων από ειδικούς ακτινολόγους γεγονός που αποτελεί σημαντικό και δύσκολο πρόβλημα στην ανάλυση κατά τη διαγνωστική διαδικασία [262].

Συνοψίζοντας, από όλα τα παραπάνω διαφαίνεται μια σειρά προβλημάτων που ενέχει η αξιολόγηση των BI-RADS 0 περιπτώσεων, καθώς η χρήση της κατηγορίας αυτής ως τελικής κατηγορίας αξιολόγησης οδηγεί στην ανάγκη πληθώρας επιπρόσθετων απεικονιστικών διαδικασιών ή και βιοψιών γεγονός που επιβαρύνει οικονομικά το σύστημα Υγείας και επηρεάζει αρνητικά την ψυχολογία των γυναικών. Επιπλέον, φαίνεται ότι η ερμηνεία της εν λόγω κατηγορίας επηρεάζεται σημαντικά από την εμπειρία των ακτινολόγων γεγονός που επιφέρει την ανάγκη για επαναλαμβανόμενες αξιολογήσεις. Προκειμένου να αντιμετωπίσουν το πρόβλημα των αδιευκρίνιστων μαστογραφικών ευρημάτων, κάποιες ερευνητικές ομάδες προέβησαν στη δοκιμή συνδυασμού απεικονιστικών μεθόδων προκειμένου να οδηγηθούν σε εγκυρότερα διαγνωστικά συμπεράσματα. Οι Zanello PA et al. [263] στη μελέτη τους διαπίστωσαν ότι η

χρήση του υπερηχογραφήματος σε συνδυασμό με τη μαστογραφία βελτιώνει σημαντικά το ποσοστό ανίχνευσης του καρκίνου ενώ έδωσαν ιδιαίτερα σημαντική έμφαση στη προστιθέμενη αξία του υπερηχογραφήματος στην αξιολόγηση των περιστατικών με BI-RADS 0 στη μαστογραφία τους. Σε μια πιο πρόσφατη έρευνα, των Lee et al [264], σε περιπτώσεις πυκνών μαστών με αξιολόγηση BI-RADS 0, διαπιστώθηκε ότι η DBT επέδειξε μεγαλύτερη διαγνωστική απόδοση σε σχέση με τον υπέρηχο μαστών, επίσης ο συνδυασμός DBT και μαστογραφίας οδήγησε σε βελτίωση της ανίχνευσης των απροσδιόριστων βλαβών του μαστού με αξιολόγηση BI-RADS 0, 3 και 4. Συγκεκριμένα, η ειδικότητα και η PPV για το υπερηχογράφημα ήταν 0,539 και 0,288 αντίστοιχα ενώ για την DBT ήταν 0, 813 και 0,5 αντίστοιχα. Επιπλέον η DBT παρουσίασε υψηλότερη διαγνωστική ακρίβεια (0,843) από εκείνη του υπερηχογραφήματος μαστού (0,611).

Έχουν γίνει πολλές προσπάθειες για την υποβοήθηση της διάγνωσης με τη βοήθεια αλγορίθμων AI [265], [266], προκειμένου να αντιμετωπιστεί το πρόβλημα που περιγράφηκε ανωτέρω και αφορά στην δυσκολία ερμηνείας κάποιων μαστογραφικών ευρημάτων εκ μέρους των ακτινολόγων. Μια σημαντική προσπάθεια για την υποστήριξη της διάγνωσης των ακτινολόγων για τις περιπτώσεις BI-RADS 0 προτάθηκε πολύ πρόσφατα από την ερευνητική ομάδα των Chunyan Yi et al. [261], και αφορά σε ένα AI σύστημα επεξεργασίας εικόνας με εφαρμογή στη προσυμπτωματική και διαγνωστική μαστογραφία. Αξίζει να σημειωθεί ότι το εν λόγω σύστημα είχε κάποιους σημαντικούς περιορισμούς αφού η διαταραχή αρχιτεκτονικής και η ασυμμετρία δεν συμπεριλήφθηκαν ρητά στη διαδικασία εκπαίδευσης και το σύστημα AI εκπαιδεύτηκε μόνο με τρέχουσες μαστογραφίες και δεν έλαβε υπόψη προηγούμενες μαστογραφίες, συμπτώματα του μαστού ή άλλα κλινικά δεδομένα. Κατά συνέπεια, όλες αυτές οι σημαντικές πληροφορίες δεν είναι ορατές για τους συμμετέχοντες ακτινολόγους κατά την ανάλυση των μαστογραφιών. Επιπλέον στη μελέτη συμπεριελήφθησαν ακτινολόγοι με μικρή ή μεσαία εμπειρία. Τέλος, η ομάδα ασθενών επιλέχθηκε από έναν αμιγώς ομοιογενή πληθυσμό γυναικών από την Ασία ενός και μόνο ιδρύματος και τα ευρήματά τους ενδέχεται να μην έχουν άμεση εφαρμογή σε έναν πληθυσμό με μεγαλύτερη φυλετική ποικιλομορφία. Παρά το γεγονός ότι η πρόσφατη εργασία των McKinney et al. [267] έδειξε ότι ένα AI σύστημα που εκπαιδεύτηκε σε μαστογραφίες που συλλέχθηκαν από κέντρα μαστού του Ηνωμένου Βασιλείου μπόρεσε να δοκιμαστεί και να αποδώσει αποτελεσματικά σε δεδομένα που συλλέχθηκαν από κλινικές των ΗΠΑ, η γενίκευση ενός AI μοντέλου εξακολουθεί να αποτελεί μια πρόκληση για την επιστημονική κοινότητα.

Παρά τους προαναφερθέντες περιορισμούς, το σύστημα των Chunyan Yi et al. [261], παρουσίασε αυξημένη διαγνωστική απόδοση στους μαστούς πυκνότητας D, κατά acr, και αξιολόγησης BI-RADS 0 γεγονός που βοήθησε τους ακτινολόγους στην εγκυρότερη αξιολόγηση των ύποπτων αλλοιώσεων. Είναι

σημαντικό γεγονός ότι το σύστημα των Chunyan Yi et al. [261], μπόρεσε να ανιχνεύσει σωστά ορισμένες από τις κακοήθεις βλάβες παρότι είχαν παρόμοια εμφάνιση με τις καλοήθεις μάζες.

Επιπλέον, διαπιστώθηκε ότι το προτεινόμενο AI σύστημα, συγκρινόμενο με μικρής και μεσαίας εμπειρίας ακτινολόγους, εμφάνισε μεγαλύτερο ποσοστό επιτυχίας στη διάγνωση των κακοηθειών σε περιστατικά BI-RADS 0, όπου συγκεκριμένα η ακρίβεια της διάγνωσης ήταν 0,80, η ευαισθησία 0,83, η ειδικότητα 0,79 και η PPV 0,22. Το προτεινόμενο σύστημα διαπιστώθηκε ότι οδήγησε σε σημαντική μείωση των περιττών εξετάσεων παρακολούθησης των ασθενών.

Το γεγονός ότι το σύστημα των Chunyan Yi et al. [261] εμφάνισε σχετικά χαμηλή ευαισθησία και πολύ χαμηλή PPV οδήγησε στο να χάνεται ένα σημαντικό μέρος των κακοηθειών καθώς διαπιστώθηκε αδυναμία εντοπισμού της αρχιτεκτονικής διαταραχής του μαστικού ιστού και των αλλοιώσεων κοντά στο τοίχωμα του μαστού. Ένα ακόμα πρόβλημα που αντιμετώπισε το εν λόγω AI εντοπίστηκε στην ταξινόμηση αδενικών ιστών με ακανόνιστη μορφή ως κακοηθών, έχοντας ως συνέπεια ψευδώς θετικά αποτελέσματα. Οι Chunyan Yi et al. [261] διαπίστωσαν ότι ένα AI σύστημα μπορεί να βοηθήσει τους ακτινολόγους με μικρή ή μεσαία εμπειρία και να οδηγήσει σε σημαντική μείωση των περιττών εξετάσεων παρακολούθησης των ασθενών, καθώς παρά την βελτίωση της απόδοσης που διαπιστώνεται στην ένταξη πολλών διαφορετικών εξετάσεων στη διαγνωστική διαδικασία, πολλές φορές προκαλούν πρόσθετο στρες και επιβαρυντική ψυχολογία στους ασθενείς, ενώ δεν αποτελούν μια οικονομικά αποδοτική λύση για τα συστήματα υγείας.

Συνολικά, είναι επιτακτική η ανάγκη περαιτέρω ανάπτυξης ενός συστήματος AI για τη βελτίωση της λήψης έγκυρων κλινικών αποφάσεων για τα απροσδιόριστα μαστογραφικά ευρήματα που κατατάσσονται ως BI-RADS 0 και όπως τόνισαν οι Chunyan Yi et al. [261], είναι σημαντικό να συμπεριληφθεί η διαταραχή αρχιτεκτονικής των μαστών στα μελλοντικά συστήματα ώστε να βελτιωθεί η διαγνωστική απόδοση. Ιδιαίτερα για τις περιπτώσεις ακτινολόγων με μικρότερη εμπειρία, ένα αποδοτικό AI σύστημα μπορεί να αποδειχθεί εξαιρετικά επωφελές ακόμα και να συμβάλει στην ανακούφιση της έλλειψης ειδικών κλινικών ιατρών σε αγροτικές ή λιγότερο ανεπτυγμένες περιοχές

Όπως καταλαβαίνουμε, αποδεικνύεται κρίσιμης σημασίας η κατάλληλη διαχείριση των περιστατικών με αξιολόγηση BI-RADS 0 καθώς ο ενδεχόμενος αποκλεισμός τους, σε περιπτώσεις που δεν ακολουθείται περαιτέρω ακτινολογική παρακολούθηση, θα μπορούσε να οδηγήσει στο να χαθεί ένα σημαντικό ποσοστό καρκίνων, όπως διαφαίνεται και στην έρευνα των Rosenberg et al. [260], αλλά και εξαχθούν παραπλανητικά και μεροληπτικά αποτελέσματα κατά τον μαστογραφικό έλεγχο. Ως εκ τούτου, οι ακτινολόγοι θα πρέπει να έχουν επίγνωση των συνεπειών της αξιολόγησης των μαστογραφικών

ευρημάτων με BI-RADS 0 αφού η εν λόγω πρακτική θα μπορούσε να αυξήσει το ποσοστό ψευδώς αρνητικών αποτελεσμάτων γεγονός που ενέχει σημαντικούς κινδύνους για τις ασθενείς.

Συνοψίζοντας, διαπιστώσαμε ότι η προσθήκη επιπλέον απεικονιστικών μεθόδων όπως είναι ο υπέρηχος και η DBT στη διαγνωστική διαδικασία, βελτιώνει σε κάποιο βαθμό τα ποσοστά της διαγνωστικής απόδοσης ωστόσο δεν λύνει το πρόβλημα της αντιμετώπισης της κατηγορίας BI-RADS 0 καθώς έχει αποδειχθεί ότι τα ποσοστά της ευαισθησίας και της PPV παραμένουν χαμηλά. Επιπλέον η χρήση των AI αλγορίθμων μέσω της επεξεργασίας εικόνας μπορεί να βελτιώσει τη διαγνωστική απόδοση διευκολύνοντας τις αποφάσεις των μη έμπειρων και μέτριας εμπειρίας ακτινολόγων, ωστόσο έχει μεγάλα περιθώρια βελτίωσης προκειμένου να μη χάνονται οι επικίνδυνες περιπτώσεις κακοήθειας που υποκρύπτονται στην κατηγορία BI-RADS 0. Εξ όσον γνωρίζουμε δεν έχει δημιουργηθεί κάποιο CDSS που να κατατάσσει με υψηλή ακρίβεια τις περιπτώσεις με BI-RADS 0 στη μαστογραφία σε καλοήθεις ή κακοήθεις περιπτώσεις.

7.2 Διαγνωστική Απόδοση του Ταξινομητή Random Forest για τα Περιστατικά γυναικών με αποτέλεσμα BI-RADS 0 στη Μαστογραφία τους

Στο παρόν υποκεφάλαιο, θα βασιστούμε στο σύστημα που αναπτύξαμε στην υποενότητα 6.4.2, όπου δημιουργήσαμε ένα CDSS βασισμένο στη χρήση του ταξινομητή RF αφού διαπιστώθηκε η διαγνωστική του υπεροχή με βάση τα σημαντικότερα στατιστικά μέτρα αξιολόγησης της απόδοσης. Συγκεκριμένα, θα διερευνήσουμε την αποτελεσματικότητα του ταξινομητή RF που προέκυψε με τα 7 καλύτερα (σύμφωνα με τη λίστα κατάταξης mRMR) χαρακτηριστικά εισόδου (βλ. Πίνακας 17) για τη διάγνωση περιπτώσεων με αξιολόγηση BI-RADS 0 στη μαστογραφία.

Όπως αναφέρθηκε προηγουμένως, η κατηγορία αυτή εγείρει πολλούς προβληματισμούς όσον αφορά τη διαχείρισή της στην ιατρική κοινότητα. Η διερεύνηση της διαγνωστικής απόδοσης ενός CDSS σε αυτές τις δύσκολες περιπτώσεις θα ενίσχυε σημαντικά τις ιατρικές αποφάσεις και ενέργειες. Αξίζει να σημειωθεί ότι η κατηγορία αυτή αντιπροσωπεύει περίπου το 25 % των γυναικών που συμμετείχαν στην έρευνα, καθώς αντιστοιχεί σε 120 από τις συνολικά 489 περιπτώσεις. Συνεπώς, έχει ιδιαίτερη σημασία να αξιολογηθεί η διαγνωστική απόδοση του CDSS σε περιπτώσεις για τις οποίες η μαστογραφία δεν μπορεί να παράσχει καμία πληροφορία για την εκτίμηση της πιθανότητας κακοήθειας.

Σε αυτή την περίπτωση χωρίσαμε τις περιπτώσεις σε δύο κατηγορίες, ένα σύνολο εκπαίδευσης και ένα σύνολο δοκιμής. Όπως αναφέρθηκε παραπάνω, επιλέξαμε να αξιολογήσουμε την απόδοση του CDSS σε περιπτώσεις με BI-RADS 0, οπότε δημιουργήσαμε ένα σύνολο δοκιμής που αποτελείται αποκλειστικά από γυναίκες που είχαν ως αξιολόγηση στη μαστογραφία το BI-RADS 0. Πιο συγκεκριμένα, επιλέξαμε τυχαία το 20% όλων των περιπτώσεων (97 περιπτώσεις) με κοινό χαρακτηριστικό το BI-RADS 0 στη

μαστογραφία για να σχηματίσουμε το σύνολο δοκιμής και δημιουργήσαμε ένα σύνολο εκπαίδευσης με όλες τις υπόλοιπες περιπτώσεις (392 περιπτώσεις όλων των κατηγοριών BI-RADS).

Επιλέξαμε να αξιολογήσουμε την απόδοση του αλγορίθμου Random Forest (RF) [253], καθώς έχει αποδειχθεί από τα αποτελέσματα της προηγούμενης μελέτης μας ότι είναι ένας αξιόπιστος και σταθερός ταξινομητής. Η διαπίστωσή μας υποστηρίζεται επίσης από διεθνείς μελέτες, όπως αυτή των Fernández-Delgado et al. [257], οι οποίοι αναρωτήθηκαν αν «χρειαζόμαστε εκατοντάδες ταξινομητές για την επίλυση προβλημάτων ταξινόμησης πραγματικών προβλημάτων;» και κατέληξαν ότι χρειαζόμαστε μόνο τον RF.

Ο πίνακας σύγκρισης που προέκυψε από τη δοκιμή του ταξινομητή RF χρησιμοποιήθηκε για την αξιολόγηση της διαγνωστικής απόδοσης του μοντέλου στον εντοπισμό των κακοηθειών που υποκρύπτονται στην κατηγορία BI-RADS 0 (βλ. Πίνακας 22).

Πίνακας 22: Πίνακας σύγκρισης του ταξινομητή Random Forest σε σύνολο δοκιμής αποκλειστικά από την κατηγορία BI-RADS 0

Random Forest	Θετικό ιστολογικό αποτέλεσμα	Αρνητικό ιστολογικό αποτέλεσμα
Ταξινομήθηκε ως θετικό	49	4
Ταξινομήθηκε ως αρνητικό	3	41

Δεδομένου ότι το CDSS μας έχει ως στόχο την ταξινόμηση των περιστατικών σε δύο κατηγορίες, θα χρησιμοποιήσουμε τα σημαντικότερα στατιστικά μέτρα για την αξιολόγηση της διαγνωστικής του απόδοσης. Τα συγκεντρωτικά αποτελέσματα για τα ανωτέρω μέτρα διαγνωστικής απόδοσης παρουσιάζονται στον Πίνακα 23.

Πίνακας 23: Διαγνωστική απόδοση του αλγορίθμου RF σε σύνολο δοκιμής αποκλειστικά από την κατηγορία BI-RADS 0

	Random Forest
<i>Accuracy (%)</i>	92,7
<i>Sensitivity (%)</i>	94,2
<i>Specificity (%)</i>	91,1
<i>PPV (%)</i>	92,4
<i>NPV (%)</i>	93,1

Youden's index

0,85

F1 score

0,93

Τα παραπάνω αποτελέσματα (Πίνακας 23) δείχνουν ότι ο αλγόριθμος RF πέτυχε πολύ υψηλή ακρίβεια 92,7%, αποδεικνύοντας την ικανότητα του CDSS να προβλέπει αρκετά επιτυχημένα την πραγματική κατάσταση μιας εξαιρετικά δύσκολης κατηγορίας γυναικών, εκείνων με την τελική αξιολόγηση BI-RADS 0 στη μαστογραφία τους. Ο ταξινομητής RF παρουσίασε επίσης τις υψηλές τιμές στα μέτρα ευαισθησίας (94,2%), ειδικότητας (91,1%), PPV (92,4%) και NPV (93,1%), δίνοντας πολύ χαμηλά ψευδώς αρνητικά και ψευδώς θετικά αποτελέσματα. Συγκεκριμένα, με βάση τον πίνακα σύγκρισης του RF σε αυτή την περίπτωση (Πίνακας 22), προκύπτουν 49 αληθώς θετικά και 41 αληθώς αρνητικά αποτελέσματα σε 97 περιπτώσεις που δοκιμάστηκαν συνολικά χρησιμοποιώντας το στρατηγικά επιλεγμένο σύνολο δοκιμών. Όπως εξηγήθηκε παραπάνω, επιλέχθηκαν τυχαία περιστατικά γυναικών αποκλειστικά από την κατηγορία BI-RADS 0, προκειμένου να αξιολογηθεί ο επικρατέστερος ταξινομητής μας στις πιο δύσκολες περιπτώσεις, αυτές που αποτελούν «μαύρο κουτί» για τους ειδικούς ιατρούς.

7.3 Συζήτηση - Συμπεράσματα

Ανακεφαλαιώνοντας, η κατηγορία αξιολόγησης BI-RADS 0 θεωρείται η πιο δύσκολη κατηγορία αξιολόγησης της μαστογραφίας, ως προς την εξαγωγή συμπερασμάτων, και επιφέρει πολλούς προβληματισμούς ως προς τη διαχείρισή της καθώς δεν παρέχει καμία πληροφορία τους κλινικούς. Πολλές φορές το εν λόγω αποτέλεσμα ακολουθείται από επιπρόσθετες απεικονιστικές εξετάσεις και βιοψίες προκειμένου να αποσαφηνιστούν ενδεχόμενοι κίνδυνοι και να αντιμετωπιστούν όσο το δυνατόν πιο έγκαιρα. Το προτεινόμενο CDSS, θέλοντας να εξαλείψει το επιπρόσθετο ψυχολογικό και οικονομικό βάρος που επιφέρουν αυτές οι διαδικασίες, χρησιμοποιεί χαρακτηριστικά της απλής καθημερινής κλινικής πρακτικής και αποτελέσματα από τη μαστογραφία και τον υπέρηχο, δηλαδή μη επεμβατικές και οικονομικά αποδοτικές διαγνωστικές εξετάσεις προκειμένου να κατατάξει τις γυναίκες με BI-RADS 0 στη μαστογραφία σε κάποια από τις δυο κλάσεις εξόδου παρέχοντας έτσι μια αξιόπιστη εκτίμηση κινδύνου καρκίνου του μαστού. Το CDSS αποδείχθηκε ότι παρουσίασε υψηλές τιμές στα μέτρα διαγνωστικής απόδοσης στο σύνολο δοκιμής των γυναικών με BI-RADS 0 στη μαστογραφία τους ξεπερνώντας το φράγμα της ασάφειας και καθιστώντας το ένα εργαλείο με καθολική εφαρμογή και ουσιαστική προστιθέμενη αξία στη διαδικασία διάγνωσης του καρκίνου του μαστού.

Κεφάλαιο 8^ο

Συμπεράσματα και Μελλοντική Έρευνα

8.1 Συζήτηση και Συμπεράσματα

Ο καρκίνος του μαστού είναι μια πολύπλοκη ασθένεια και πολύ μικρό ποσοστό των περιστατικών οφείλεται σε κληρονομικούς παράγοντες [79]. Όπως παρουσιάστηκε στο Κεφάλαιο 2 της παρούσας διατριβής, η παθολογία του μαστού οδηγεί σε πολλές διαφορετικές κατηγορίες παθήσεων, άλλες κακοήθεις και άλλες καλοήθεις. Η αιτιολογία της νόσου δεν έχει διευκρινιστεί με σαφήνεια από την επιστημονική κοινότητα, εντούτοις μια πληθώρα παραγόντων κινδύνου έχουν ενοχοποιηθεί, με βάση τη διεθνή βιβλιογραφία, για τη συμβολή τους στην ανάπτυξη καρκίνου του μαστού, όπως αναλύεται διεξοδικά στο υποκεφάλαιο 2.6. Οι διαγνωστικές μέθοδοι για την ανίχνευση της νόσου περιλαμβάνουν πληθώρα απεικονιστικών εξετάσεων και εξετάσεις του βιολογικού υλικού που προκύπτει από τη διεξαγωγή βιοψιών.

Η εξατομικευμένη διαχείριση των γυναικών με κίνδυνο ανάπτυξης καρκίνου του μαστού αποτελεί μια ιδιαίτερος σημαντική αλλά ταυτόχρονα δύσκολη και πολύπλοκη υπόθεση. Τα εργαλεία που είναι διαθέσιμα για την εργασία αυτή δεν εξυπηρετούν στον απαιτούμενο βαθμό: από τη μία πλευρά οι βασικές απεικονιστικές εξετάσεις όπως είναι η μαστογραφία και ο υπέρηχος μαστών εμφανίζουν περιορισμένη διαγνωστική απόδοση (είτε χαμηλή ευαισθησία, είτε χαμηλή ειδικότητα) [79], ενώ από την άλλη η MRI είναι μια εξέταση με υψηλή διαγνωστική απόδοση ωστόσο, σύμφωνα με τις παγκόσμιες κατευθυντήριες οδηγίες, δε συνίσταται για χρήση στην καθημερινή κλινική πράξη καθώς πρόκειται για μια εξειδικευμένη και δαπανηρή, για τα συστήματα υγείας, εξέταση [93].

Επιπλέον η άμεση παραπομπή σε βιοψία κάποιες φορές συνοδεύεται με εμφάνιση τόσο ψυχολογικών όσο και παθολογικών προβλημάτων για την ασθενή καθώς πρόκειται για μια επεμβατική διαδικασία [6]. Τα ιδιαίτερα χαρακτηριστικά της νόσου και τα μειονεκτήματα που παρουσιάζουν τα διαθέσιμα εργαλεία αντιμετώπισής της, καθιστούν απολύτως αναγκαία την υποστήριξη της απόφασης για την έγκαιρη και έγκυρη ανίχνευση και την αποτελεσματική διαχείριση των γυναικών που βρίσκονται σε κίνδυνο ανάπτυξης καρκίνου του μαστού.

Σε αυτό το πλαίσιο και λαμβάνοντας υπόψη αφενός την πολυπλοκότητα που χαρακτηρίζει μια σχετική κλινική απόφαση και αφετέρου τις ιδιαιτερότητες του προβλήματος, η δημιουργία ενός ευφυούς συστήματος υποστήριξης κλινικής απόφασης παρουσιάζει αυξημένη χρησιμότητα στην καθημερινή πρακτική διαχείρισης της συγκεκριμένης νόσου. Η χρήση των υπολογιστών (και της μηχανικής μάθησης) στη διάγνωση, πρόβλεψη και πρόγνωση ασθενειών αποτελεί δείγμα μιας αυξανόμενης τάσης προς την εξατομικευμένη προληπτική ιατρική. Αυτή η τάση είναι σημαντική, όχι μόνο για τους ασθενείς (όσον αφορά στον τρόπο ζωής και τις αποφάσεις ποιότητας ζωής), αλλά και για τους ιατρούς (στη λήψη απόφασης), για τους οικονομολόγους της υγείας και τους σχεδιαστές στρατηγικών (στην υλοποίηση μεγάλης κλίμακας στρατηγικών πρόληψης ή θεραπείας του καρκίνου).

Προσδοκώντας στην κάλυψη αυτής της ανάγκης, η παρούσα διατριβή παρουσιάζει την ανάπτυξη αλγορίθμων υπολογιστικής νοημοσύνης για τη δημιουργία ενός ευφυούς Συστήματος υποστήριξης κλινικής απόφασης για την έγκαιρη διάγνωση και την εξατομικευμένη διαχείριση γυναικών με κίνδυνο ανάπτυξης καρκίνου του μαστού. Μετά από εκτεταμένη βιβλιογραφική ανασκόπηση, διαπιστώθηκε ότι οι μέθοδοι που έχουν αναπτυχθεί για την εκτίμηση του κινδύνου και την έγκαιρη διάγνωση του καρκίνου του μαστού, χρησιμοποιούν γενετικές πληροφορίες, δεδομένα από επεμβατικές εξετάσεις, όπως η CNB και η FNA, δεδομένα από την ανάλυση εικόνων μαστογραφίας, υπερήχου και MRI, πληροφορίες από εξέταση αιματολογικού υλικού και χαρακτηριστικά που λαμβάνονται από τα προσωπικά ιατρικά αρχεία των γυναικών, όπως αναλυτικά παρουσιάζονται στο Κεφάλαιο 4 της παρούσας διατριβής.

Τα χαρακτηριστικά που συλλέχθηκαν, χρησιμοποιήθηκαν και αξιολογήθηκαν, στο προτεινόμενο CDSS, αντιπροσωπεύουν το μεγαλύτερο μέρος των παραγόντων κινδύνου που αναλύθηκαν διεξοδικά στο υποκεφάλαιο 2.6 και σύμφωνα με τις γνώσεις μας δεν έχει δημοσιευθεί μέχρι σήμερα κάποιο αντίστοιχο CDSS που να συνδυάζει τόσους πολλούς παράγοντες κινδύνου ανάπτυξης και εντοπισμού της νόσου αξιοποιώντας δεδομένα που συλλέγονται στην καθημερινή κλινική πράξη. Στο υποκεφάλαιο 5.1 παρουσιάζονται αναλυτικά οι λόγοι επιλογής των συγκεκριμένων χαρακτηριστικών με βάση τη σημασία τους στην ανάπτυξη του καρκίνου του μαστού αλλά και τη συμβολή τους για την έγκαιρη διάγνωση καρκινικών αλλοιώσεων (π.χ. μορφολογικά χαρακτηριστικά αλλοιώσεων και μαστών), όπως αυτά προέκυψαν με τα την εκτενή μελέτη της διεθνούς βιβλιογραφίας.

Εν κατακλείδι, εξ όσων γνωρίζουμε, δεν έχει δημοσιευθεί μέχρι σήμερα στη βιβλιογραφία ένα ισοδύναμο ευφές σύστημα με αυτό που προτείνεται στην παρούσα μελέτη, δηλαδή ένα σύστημα που συνδυάζει μη επεμβατικά και οικονομικά αποδοτικά διαγνωστικά εργαλεία μαζί με τα χαρακτηριστικά του ατομικού ιατρικού φακέλου, με στόχο την ενίσχυση της διαγνωστικής απόδοσης

Συνοψίζοντας, στη διδακτορική διατριβή συμπεριελήφθησαν 489 γυναίκες εκ των οποίων οι 287 είχαν παθολογοανατομική έκθεση υπέρ κακοήθειας και οι 202 υπέρ καλοήθειας. Όλες οι περιπτώσεις γυναικών υποβλήθηκαν σε χειρουργική επέμβαση μαστού, είτε σε συντηρητική επέμβαση μαστού (Ουκεκτομή) είτε σε Μαστεκτομή. Όπως αναφέρθηκε και προηγουμένως, ο υπό μελέτη πληθυσμός αφορά πληθυσμό παραπομπής σε χειρουργική επέμβαση μαστού (πληθυσμός διαλογής - στόχος). Για κάθε ένα περιστατικό χρησιμοποιήθηκε ως μεταβλητή εξόδου (ή μεταβλητή στόχος) το αποτέλεσμα της ιστολογικής εξέτασης όπως αυτό προέκυψε από την παθολογοανατομική έκθεση, καθώς αυτό εκφράζει την πραγματική κατάσταση των μαστών κάθε γυναίκας (Gold Standard).

Για τους σκοπούς της παρούσας έρευνας, ερευνήθηκε η διαγνωστική απόδοση των ταξινομητών Naive Bayes και Random Forest αφού προηγήθηκε εφαρμογή των τεχνικών επιλογής χαρακτηριστικών με βάση τη σημαντικότητά τους. Ο λόγος που επιλέχθηκε ο Naive Bayes είναι ότι παρά τον αφελή σχεδιασμό και τις υπεραπλουστευμένες υποθέσεις ανεξαρτησίας των χαρακτηριστικών, έχει αποδειχθεί ότι ο εν λόγω αλγόριθμος αποδίδει ιδιαίτερα καλά σε πολλά πραγματικά πολύπλοκα προβλήματα ταξινόμησης, ενώ ταυτόχρονα είναι ανθεκτικός σε «παραβιάσεις» της παραδοχής ανεξαρτησίας των χαρακτηριστικών. Ο Naive Bayes είναι ένας μη παραμετρικός ταξινομητής, λειτουργεί άμεσα και δεν απαιτεί ιδιαίτερο σχεδιασμό για την κατασκευή του [252]. Από την άλλη πλευρά, ο Random Forest είναι ένας πολύ αξιόπιστος και σταθερός ταξινομητής και όπως διαπιστώνεται στην εκτενή μελέτη των Delgado et al. [257], αναδείχθηκε ως ο αποτελεσματικότερος ταξινομητής όταν συναγωνίστηκε με σχεδόν 200 άλλους ταξινομητές, κατά μέσο όρο, σε περισσότερα από 100 σύνολα δεδομένων.

Συγκεκριμένα, για την κατασκευή των μοντέλων ταξινόμησης χρησιμοποιήθηκαν τρεις διαφορετικές τεχνικές επιλογής χαρακτηριστικών, με βάση το δείκτη Information Gain, με βάση τη μέθοδο mRMR και με βάση το δείκτη σημαντικότητας gini importance που είναι ενσωματωμένος στον αλγόριθμο Random Forest. Η ίδια διαδικασία κατασκευής των μοντέλων ακολουθήθηκε και για τους δύο ταξινομητές, σύμφωνα με το διάγραμμα ροής (βλ. Εικόνα 25). Ο ταξινομητής Naive Bayes αξιολογήθηκε με τη χρήση δυο διαφορετικών λιστών κατάταξης, όπως αυτές προέκυψαν στα υποκεφάλαια 6.2.1 και 6.2.2. Από την άλλη, ο ταξινομητής Random Forest έχει ενσωματωμένη την αξιολόγηση της σημαντικότητας των χαρακτηριστικών μέσω του δείκτη Gini importance, οπότε για την κατασκευή του έγινε χρήση της δικής του λίστας κατάταξης αλλά και των δύο λιστών που προέκυψαν από την εφαρμογή των τεχνικών Information Gain και mRMR.

Αφού συγκρίναμε την απόδοση των διαφορετικών μοντέλων ταξινόμησης που προέκυψαν με τις διαφορετικές λίστες κατάταξης, καταλήξαμε ότι ο Random Forest παρουσίασε την υψηλότερη διαγνωστική απόδοση με τη χρήση της λίστας που προέκυψε από την εφαρμογή της μεθόδου mRMR. Συγκεκριμένα, ο

ταξινομητής επέδειξε την υψηλότερη διαγνωστική ακρίβεια όταν χρησιμοποιήθηκαν τα εξής χαρακτηριστικά ως διάνυσμα εισόδου στο μοντέλο: Τα περιθώρια της αλλοίωσης (Tumor Margins), το σχήμα της αλλοίωσης (Tumor Shape), την ύπαρξη αγγείωσης (Vascularity), τη διαταραχή της αρχιτεκτονικής των μαστών (Arc_Disturbance), την ύπαρξη νέου ευρήματος (New finding), το μέγεθος του στήθους γυναίκας (Breast Size) και την ηλικία της (Age). Τα χαρακτηριστικά αυτά αποτέλεσαν το βέλτιστο υποσύνολο χαρακτηριστικών καθώς ο συνδυασμός τους φαίνεται ότι παρέχει τη μέγιστη διαγνωστική πληροφορία.

Ο ταξινομητής RF έδειξε την υψηλότερη ακρίβεια 95,9%, δείχνοντας πόσο αποτελεσματικά λειτουργεί στο συνδυασμό δυαδικών, κατηγορικών και αριθμητικών μεταβλητών και επιβεβαιώνοντας τη διεθνή βιβλιογραφία όπου συχνά αναφέρεται ως ο πιο αξιόπιστος και σταθερός αλγόριθμος [257]. Ο RF έδειξε τις υψηλότερες τιμές στα μέτρα ευαισθησίας (96,5%), NPV (95%), Youden's index (0,92) και F1 score (0,97), υποδεικνύοντας την καταλληλότητά του για τη δημιουργία του καινοτόμου μοντέλου CDSS καθώς προσφέρει εξαιρετικά χαμηλά ποσοστά ψευδώς αρνητικών και ψευδώς θετικών αποτελεσμάτων και εντοπίζει με πολύ μεγάλη επιτυχία τα πραγματικά θετικά περιστατικά (ασθενείς). Το γεγονός ότι το Σύστημα υποστήριξης κλινικής απόφασης (CDSS), το οποίο βασίζεται στον ταξινομητή RF, μπορεί να διακρίνει με υψηλή ακρίβεια τις περιπτώσεις πραγματικών ασθενών και υγιών γυναικών σημαίνει ότι μπορεί να διαδραματίσει διπλά ευεργετικό ρόλο για τα συστήματα υγειονομικής περίθαλψης. Αφενός, έχει τη δυνατότητα να εντοπίζει έγκαιρα τις περιπτώσεις γυναικών με αυξημένο κίνδυνο καρκίνου του μαστού, με αποτέλεσμα την έγκαιρη παρέμβαση από τους κλινικούς ιατρούς και την προετοιμασία τόσο της ασθενούς όσο και του συστήματος υγείας για την εξατομικευμένη διαχείριση κάθε περιστατικού. Από την άλλη, με την παροχή σημαντικών πληροφοριών για περιστατικά των υγιών γυναικών, οδηγεί στη μείωση του περιττού κόστους αποφεύγοντας την υπερδιάγνωση και την υπερθεραπεία.

Ένας ακόμη τρόπος προκειμένου να αξιολογήσουμε την προστιθέμενη αξία του προτεινόμενου CDSS για την υποβοήθηση της διάγνωσης του καρκίνου του μαστού, είναι να ανατρέξουμε στις 3 διαφορετικές περιπτώσεις αξιολόγησης διαγνωστικής απόδοσης της μαστογραφίας που εξηγήθηκαν παραπάνω (βλ. Πίνακας 9, Πίνακας 10 & Πίνακας 11) και να συγκρίνουμε τις τιμές στον Πίνακα 12 με τα αποτελέσματα του RF αλγορίθμου (βλ. Πίνακας 21). Ο ταξινομητής Random Forest θεωρήθηκε εξαιρετικά πιο αποτελεσματικός σε σχέση με τη μαστογραφία, καθώς αντιληφθήκαμε ότι όλα τα μέτρα διαγνωστικής απόδοσης που χρησιμοποιήθηκαν, είναι προφανώς υψηλότερα μετά τη χρήση του ταξινομητή RF σε σύγκριση με το διαγνωστικό εργαλείο της μαστογραφίας με βάση τη βαθμολογία BI-RADS.

Σημειώνοντας την υψηλή διαγνωστική απόδοση του προτεινόμενου μοντέλου, είναι σημαντικό να τονίσουμε το γεγονός ότι χρησιμοποιήσαμε το σύνολο δεδομένων 489 γυναικών, συμπεριλαμβανομένων

εκείνων με βαθμολογία BI-RADS 0 στη μαστογραφία τους. Όπως σημειώθηκε παραπάνω, οι 120 περιπτώσεις με BI-RADS 0 εξαιρέθηκαν από τους πίνακες σύγκρισης της μαστογραφίας και δεν ελήφθησαν υπόψη στον υπολογισμό των διαγνωστικών της μέτρων. Ωστόσο, αυτές οι περιπτώσεις δεν εξαιρέθηκαν από το μοντέλο μας, καθώς επιλέξαμε να διερευνήσουμε τη συμβολή τους στη διαγνωστική διαδικασία παρά το γεγονός ότι αντιπροσωπεύουν μια δύσκολη και «ανεπαρκή κατηγορία». Τονίζεται επίσης ότι ο ταξινομητής RF αξιολογήθηκε ξεχωριστά ως προς την απόδοσή του στην κατηγορία γυναικών με BI-RADS 0 στη μαστογραφία, προκειμένου να διαπιστωθεί η ικανότητά του να αξιολογεί σωστά τις περιπτώσεις αυτές όταν το σύνολο δοκιμής του ταξινομητή προέρχεται αποκλειστικά από αυτή την κατηγορία περιπτώσεων. Το CDSS αποδείχθηκε ότι παρουσίασε υψηλές τιμές στα μέτρα διαγνωστικής απόδοσης στο σύνολο δοκιμής των γυναικών με BI-RADS 0 στη μαστογραφία τους ξεπερνώντας το φράγμα της ασάφειας και καθιστώντας το ένα εργαλείο με καθολική εφαρμογή και ουσιαστική προστιθέμενη αξία στη διαδικασία διάγνωσης του καρκίνου του μαστού.

Επιπλέον, είναι σημαντικό το γεγονός ότι το μοντέλο μας να μπορεί να βοηθήσει τους κλινικούς ιατρούς να εξάγουν συμπεράσματα ως προς τη σημαντικότητα των πληροφοριών που λαμβάνουν, με βάση τα χαρακτηριστικά που συλλέγονται από μη επεμβατικές και χαμηλού κόστους διαγνωστικές διαδικασίες και τον ατομικό ιατρικό φάκελο των γυναικών. Η διάκριση των πιο σημαντικών χαρακτηριστικών ως προς το αποτέλεσμα της διαδικασίας, είναι πιθανό να οδηγήσει στην καλύτερη διαίσθηση για το πραγματικό πρόβλημα, δίνοντας έτσι τη δυνατότητα στους ειδικούς του τομέα να το αντιμετωπίσουν πιο αποτελεσματικά.

Συνοψίζοντας, μέσω της παρούσας έρευνας προτείνεται μια βιώσιμη και οικονομικά αποδοτική λύση για τα εθνικά συστήματα υγειονομικής περίθαλψης που μπορεί να συμβάλει σε σημαντική μείωση του κόστους όσον αφορά τη διαχείριση των γυναικών που κινδυνεύουν να αναπτύξουν καρκίνο του μαστού, επιτυγχάνοντας παράλληλα αξιοσημείωτη αύξηση της εγκυρότητας της διάγνωσης.

8.2 Μελλοντική Έρευνα

Στις μέρες μας, οι ομοσπονδιακές βάσεις δεδομένων (federated databases) έχουν κερδίσει μεγάλη προσοχή λόγω της όλο και αυξανόμενης ανάγκης επεξεργασίας ευαίσθητων προσωπικών δεδομένων ασθενών σε διάφορα κλινικά κέντρα, όπου οι υφιστάμενες νομοθεσίες και οι απαιτήσεις προστασίας δεδομένων εμποδίζουν την ανταλλαγή δεδομένων εκτός των τοπικών βάσεων δεδομένων (data silos).

Είναι γεγονός ότι η εγκατάσταση τοπικού εξοπλισμού (διακομιστών - server) εντός των χώρων του νοσοκομείου ώστε να καταστεί δυνατή η ανάπτυξη ροών εργασίας ανάλυσης δεδομένων στα ευαίσθητα δεδομένα των ασθενών (μέσω της δημιουργίας ασφαλών εικονικών ιδιωτικών δικτύων) δεν είναι ούτε βιώσιμη ούτε εφικτή λύση λόγω των περιορισμών ISO ή νομικών φραγμών. Για την αντιμετώπιση αυτών

των αναγκών, στη βιβλιογραφία έχουν προταθεί ομοσπονδιακά περιβάλλοντα [268]–[273] που ακολουθούν την προσέγγιση «ας φέρουμε την ανάλυση στα δεδομένα» (“bring the analysis to the data”), σύμφωνα με την οποία τα ευαίσθητα δεδομένα των ασθενών αποθηκεύονται σε ασφαλείς ιδιωτικές βάσεις δεδομένων εντός μιας ομοσπονδιακής υποδομής νέφους (federated cloud infrastructure). Κάθε πάροχος δεδομένων (από οποιονδήποτε χώρο ή και χώρα) μπορεί να αποκτήσει πρόσβαση σε ένα ιδιωτικό αποθετήριο νέφους (cloud repository), όπου μπορεί να αποθηκεύσει ευαίσθητα ψευδωνυμοποιημένα δεδομένα ασθενών σε ασφαλείς ομοσπονδιακές βάσεις δεδομένων για επεξεργασία.

Προς αυτή την κατεύθυνση, σχεδιάζουμε να αναπτύξουμε μια ομοσπονδιακή έκδοση του προτεινόμενου CDSS, ώστε να μπορεί να εκτελεστεί εντός ομοσπονδιακών περιβαλλόντων για την υποστήριξη της εκπαίδευσης σε βάσεις δεδομένων με κλινικά δεδομένα από διαφορετικές χώρες, όπου η εγκατάσταση τοπικού εξοπλισμού ή η ενσωμάτωση των δεδομένων δεν είναι εφικτή.

Οι συμβατικοί ML αλγόριθμοι, όπως είναι ο Multinomial NaiveBayes, οι SGD (stochastic gradient descent) ταξινομητές, όπως είναι ο SVM, οι παλινδρομητές (γραμμικοί και λογιστικοί) και τα δέντρα βαθμιδωτής ενίσχυσης (radient boosting trees - GBTs) [270], [271], [273] θα επεκταθούν για να υποστηρίξουν τη σταδιακή μάθηση για πολυτροπικές πηγές δεδομένων που θα βρίσκονται σε πολλαπλές ομοσπονδιακές βάσεις δεδομένων. Για το σκοπό αυτό, θα αναπτυχθεί μια απλή μηχανή υπολογιστικού νέφους για να: (i) να αρχικοποιεί τα μοντέλα ML για την πρόβλεψη του καρκίνου του μαστού, ii) να παρακολουθεί τη διαδικασία ενημέρωσης των βαρών του μοντέλου ML στις διαθέσιμες βάσεις δεδομένων εκπαίδευσης και iii) να επιτρέπει τη διαδικασία επικύρωσης του μοντέλου ML σε μία ή περισσότερες ομοσπονδιακές βάσεις δεδομένων.

Για το σκοπό αυτό, θα αναπτυχθεί μια ασφαλής υπηρεσία πρόσβασης σε δεδομένα που θα επιτρέπει την πρόσβαση στους αλγορίθμους ML κάθε ομοσπονδιακής βάσης δεδομένων. Θα εφαρμοστούν μέθοδοι τυχαίας υποδειγματοληψίας με μεθόδους αντικατάστασης για την αντιμετώπιση της ανισορροπίας των κλάσεων κατά τη διαδικασία εκπαίδευσης σε όλες τις ομοσπονδιακές βάσεις δεδομένων. Επιπλέον, θα σχεδιαστούν και θα εφαρμοστούν υβριδικές συναρτήσεις απωλειών για τη μείωση των φαινομένων υπερπροσαρμογής (overfitting) κατά τη διαδικασία ενημέρωσης του βάρους του μοντέλου σε κάθε ομοσπονδιακή βάση δεδομένων. Επιπλέον, θα εφαρμοστεί ανάλυση επεξήγησης των ML μοντέλων (π.χ. με τη μέθοδο προσθετικής επεξήγησης Shapley (Shapley additive explanation - SHAP) [274]) για να προκύψουν ομοσπονδιακά και επεξηγήσιμα ML μοντέλα και ερμηνεύσιμα μονοπάτια αποφάσεων.

Κλείνοντας μελλοντικός στόχος της παρούσας διδακτορικής διατριβής είναι μέσω της εφαρμογής μεθόδων εναρμόνισης πολυτροπικών δεδομένων (Multimodal data harmonization) να ξεπεραστούν οι δομικές ετερογένειες που υπάρχουν στα κλινικά δεδομένα (προερχόμενα από τον ιατρικό φάκελο ή τις

απεικονιστικές εξετάσεις) σε διαφορετικές τοπικά εγκατεστημένες βάσεις δεδομένων (data silos). Με αυτόν τον τρόπο το προτεινόμενο CDSS θα επεκταθεί για να υποστηρίξει την εφαρμογή σε εναρμονισμένα (harmonized) κλινικά δεδομένα τα οποία θα αποθηκεύονται με ασφάλεια σε ομοσπονδιακά περιβάλλοντα.

Παράρτημα Α

Βασικά Διαγνωστικά Μέτρα Απόδοσης

Στο παρόν Παράρτημα επεξηγούνται συνοπτικά βασικές έννοιες σχετικά με την αξιολόγηση της απόδοσης των διαγνωστικών εξετάσεων (τεστ) [275].

Ανάλογα με τη σχέση των αποτελεσμάτων ενός διαγνωστικού τεστ και της πραγματικής κατάστασης του ατόμου που ελέγχεται, ορίζονται οι εξής έννοιες όσον αφορά στο αποτέλεσμα του τεστ:

- ✚ Αληθώς θετικό (true positive - TP): Ασθενής που αναγνωρίστηκε σωστά από το τεστ ως ασθενής
- ✚ Ψευδώς θετικό (false positive - FP): Υγιής που αναγνωρίστηκε λανθασμένα από το τεστ ως ασθενής
- ✚ Αληθώς αρνητικό (true negative - TN): Υγιής που αναγνωρίστηκε σωστά από το τεστ ως υγιής
- ✚ Ψευδώς αρνητικό (false negative - FN): Ασθενής που αναγνωρίστηκε λανθασμένα από το τεστ ως υγιής

Με βάση τις ανωτέρω έννοιες μπορεί να κατασκευαστεί ένας πίνακας ο οποίος αποτυπώνει την κατανομή των περιστατικών όπως αυτά ταξινομήθηκαν από το τεστ. Ο πίνακας αυτός ονομάζεται Πίνακας Σύγχυσης και παρουσιάζεται παρακάτω. Από τον Πίνακα Σύγχυσης εν συνεχεία υπολογίζονται τα διάφορα διαγνωστικά μέτρα απόδοσης, όπως παρουσιάζεται παρακάτω.

Πίνακας Σύγχυσης

Ο Πίνακας Σύγχυσης (Confusion Matrix) είναι ένα πολύ δημοφιλές μέτρο που χρησιμοποιείται κατά την επίλυση προβλημάτων ταξινόμησης. Μπορεί να εφαρμοστεί στη δυαδική ταξινόμηση καθώς και σε προβλήματα ταξινόμησης πολλαπλών κατηγοριών. Ένα παράδειγμα πίνακα σύγχυσης για δυαδική ταξινόμηση παρουσιάζεται στον παρακάτω πίνακα:

Πίνακας 24: Πίνακας σύγχυσης για δυαδική ταξινόμηση.

Αποτέλεσμα διαγνωστικού Τεστ		
Πραγματική	Negative	Positive

κατάσταση	Negative	TN	FP
	Positive	FN	TP

Ακρίβεια (Accuracy)

Το πιο απλό μέτρο για την έκφραση της απόδοσης ενός τεστ είναι η ακρίβεια. Το μέτρο αυτό είναι το συνολικό ποσοστό των ορθών αποτελεσμάτων του τεστ στο σύνολο του υπό μελέτη πληθυσμού και υπολογίζεται από τον κάτωθι τύπο:

$$\left\{ \text{Ακρίβεια} = \frac{TP + TN}{T} = \text{όπου } T = \text{Συνολικό πλήθος περιστατικών} \right\}$$

Ευαισθησία (Sensitivity)

Εκφράζει την ικανότητα του τεστ να αναγνωρίζει σωστά τους ασθενείς. Μετρά το ποσοστό των αληθώς θετικών αποτελεσμάτων (TP) του τεστ στο σύνολο των πραγματικών θετικών αποτελεσμάτων (TD). Πρακτικώς αντικατοπτρίζει το ποσοστό των πραγματικών ασθενών που σωστά έχουν προσδιοριστεί από το τεστ ως ασθενείς. Η ευαισθησία υπολογίζεται από τον Πίνακα Σύγκρισης με βάση την παρακάτω σχέση:

$$\left\{ \text{Ευαισθησία} = \frac{TP}{TP + FN} = \frac{\text{Πλήθος αληθώς θετικών}}{\text{Πλήθος όλων όσων έχουν την ασθένεια}} \right\}$$

Ειδικότητα (Specificity)

Εκφράζει την ικανότητα του τεστ να αναγνωρίζει σωστά εκείνους που δεν είναι ασθενείς. Μετρά το ποσοστό των αληθώς αρνητικών αποτελεσμάτων (TN) του τεστ στο σύνολο των πραγματικών αρνητικών αποτελεσμάτων (TND). Πρακτικώς αντικατοπτρίζει το ποσοστό των πραγματικά υγίων που σωστά έχουν προσδιοριστεί από το τεστ ως υγιείς.

Η ειδικότητα υπολογίζεται από τον Πίνακα Σύγκρισης με βάση την παρακάτω σχέση:

$$\left\{ \text{Ειδικότητα} = \frac{TN}{TN + FP} = \frac{\text{Πλήθος αληθώς αρνητικών}}{\text{Πλήθος όλων όσων δεν έχουν την ασθένεια}} \right\}$$

Μέσω των μέτρων της ειδικότητας και της ευαισθησίας μπορούν να προσδιοριστούν διάφορες άλλες ποσότητες σχετικές με την απόδοση, όπως:

- ✚ Ποσοστό ψευδώς θετικών αποτελεσμάτων (False Positive Rate – FPR) = $FP / (FP + TN) = 1 - \text{ειδικότητα}$
- ✚ Ποσοστό ψευδώς αρνητικών αποτελεσμάτων (False Negative Rate – FNR) = $FN / (TP + FN) = 1 - \text{ευαισθησία}$
- ✚ Ποσοστό αληθώς θετικών αποτελεσμάτων (True Positive Rate – TPR) = ευαισθησία
- ✚ Ποσοστό αληθώς αρνητικών αποτελεσμάτων (True Negative Rate – TNR) = ειδικότητα

Αρνητική Προγνωστική Αξία (Negative Predictive Value - NPV)

Εκφράζει το ποσοστό των ατόμων που βρέθηκαν από το τεστ ως υγιείς και δεν έχουν στην πραγματικότητα την ασθένεια (είναι στην πραγματικότητα υγιείς) [276]. Το συγκεκριμένο μέτρο απαντά στο ερώτημα «άν για ένα άτομο το τεστ βγάλει αρνητικό αποτέλεσμα, ποια η πιθανότητα να μην έχει την ασθένεια;». Δείχνει λοιπόν ουσιαστικά τον βαθμό εμπιστοσύνης που μπορούμε να έχουμε σε ένα αρνητικό αποτέλεσμα του τεστ. Η μαθηματική απόδοση του συγκεκριμένου μέτρου περιγράφεται από τη σχέση:

$$\left\{ \text{Αρνητική Προγνωστική Αξία} = \frac{TN}{TN + FN} = \frac{\text{Πλήθος αληθώς αρνητικών}}{\text{Πλήθος αρνητικών αποτελεσμάτων}} \right\}$$

Θετική Προγνωστική Αξία (Positive Predictive Value - PPV)

Η θετική προγνωστική αξία εκφράζει το ποσοστό των περιπτώσεων που δίνουν θετικά αποτελέσματα της διαγνωστικής εξέτασης και είναι ασθενείς [276]. Είναι η αναλογία των ασθενών που πραγματικά διαγνώστηκαν ως θετικοί προς όλους εκείνους που είχαν θετικά αποτελέσματα δοκιμών (συμπεριλαμβανομένων των υγιών ατόμων που διαγνώστηκαν εσφαλμένα ως ασθενείς). Αυτό το χαρακτηριστικό μπορεί να προβλέψει πόσο πιθανό είναι κάποιος να είναι πραγματικά ασθενής, σε περίπτωση θετικού αποτελέσματος δοκιμής. Η μαθηματική απόδοση του συγκεκριμένου μέτρου περιγράφεται από τη σχέση:

$$\left\{ \text{Θετική Προγνωστική Αξία} = \frac{TP}{TP + FP} = \frac{\text{Πλήθος αληθώς θετικών}}{\text{Πλήθος θετικών αποτελεσμάτων}} \right\}$$

Youden's Index

Ο δείκτης προτάθηκε από τον W.J. Youden το 1950 [277], ως ένας τρόπος σύνοψης της απόδοσης μιας διαγνωστικής δοκιμής. Η τιμή του κυμαίνεται από 0 έως 1 (συμπεριλαμβανομένου) και έχει μηδενική τιμή όταν μια διαγνωστική δοκιμή δίνει το ίδιο ποσοστό θετικών αποτελεσμάτων για ομάδες με και χωρίς τη νόσο, όταν δηλαδή η δοκιμή είναι άχρηστη. Η τιμή 1 υποδηλώνει ότι δεν υπάρχουν ψευδώς θετικά ή ψευδώς αρνητικά αποτελέσματα, δηλαδή η δοκιμή είναι τέλεια. Ο δείκτης δίνει ίση βαρύτητα στις ψευδώς θετικές και ψευδώς αρνητικές τιμές, οπότε όλες οι δοκιμές με την ίδια τιμή του δείκτη δίνουν το ίδιο ποσοστό συνολικών λανθασμένων αποτελεσμάτων.

$$Youden's\ Index = \text{Ευαισθησία} + \text{Ειδικότητα} - 1$$

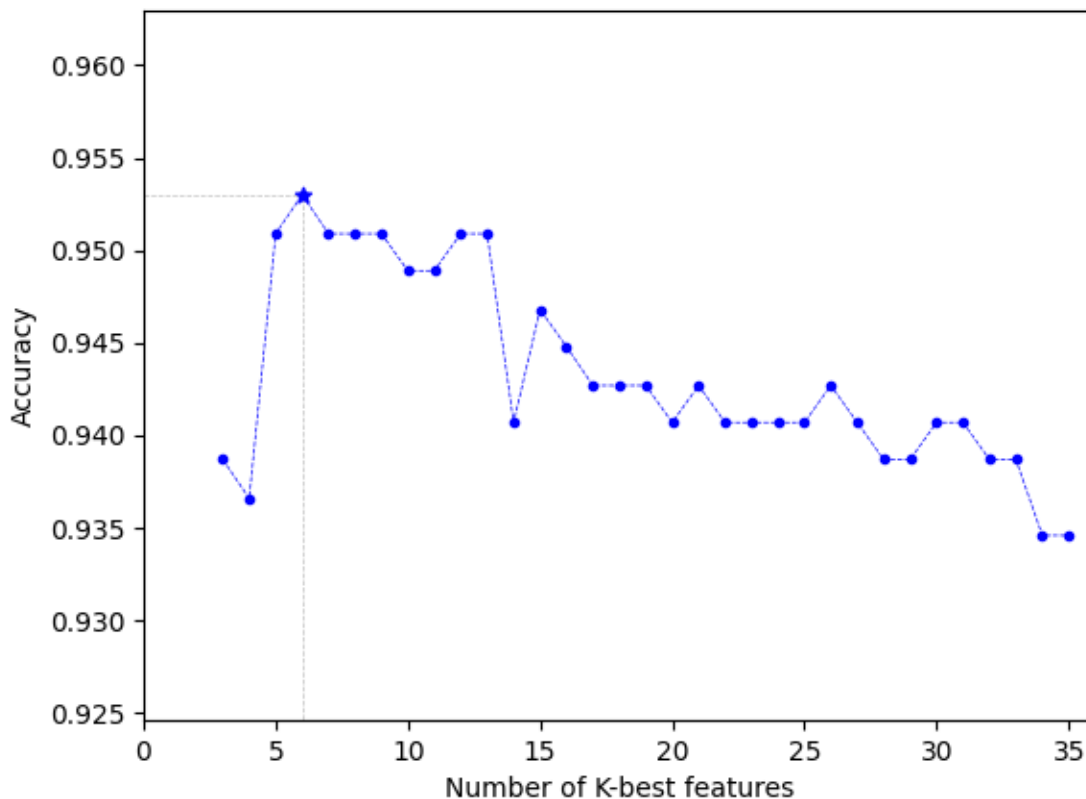
F1 score (F1 σκορ)

Η PPV και η ευαισθησία είναι τα δύο δομικά στοιχεία του F1 σκορ. Ο στόχος του F1 σκορ είναι να συνδυάσει τις μετρικές της PPV και της ευαισθησίας σε μια ενιαία μετρική. Ταυτόχρονα, το F1 σκορ έχει σχεδιαστεί για να λειτουργεί καλά σε μη ισορροπημένα δεδομένα. Ορίζεται ως ο αρμονικός μέσος όρος της PPV και της ευαισθησίας. Ο τύπος του F1 σκορ παρουσιάζεται εδώ:

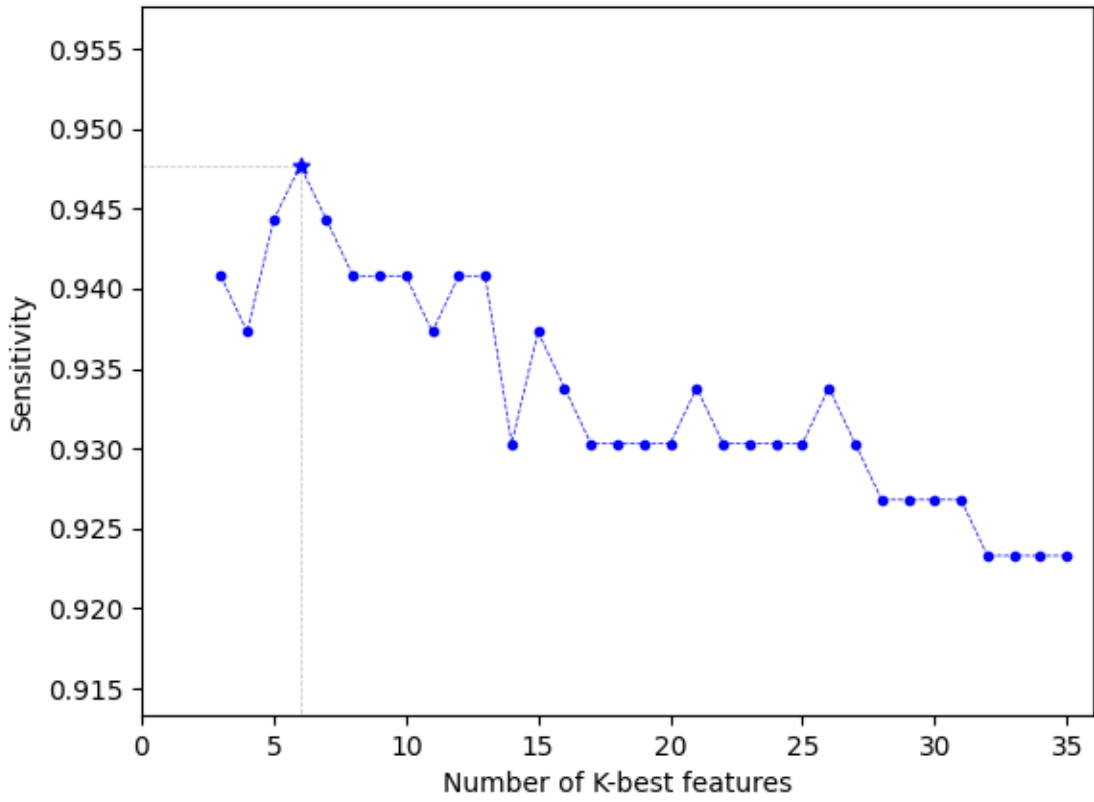
$$F1\ score = 2 * \frac{PPV * \text{ευαισθησία}}{PPV + \text{ευαισθησία}}$$

Παράρτημα Β

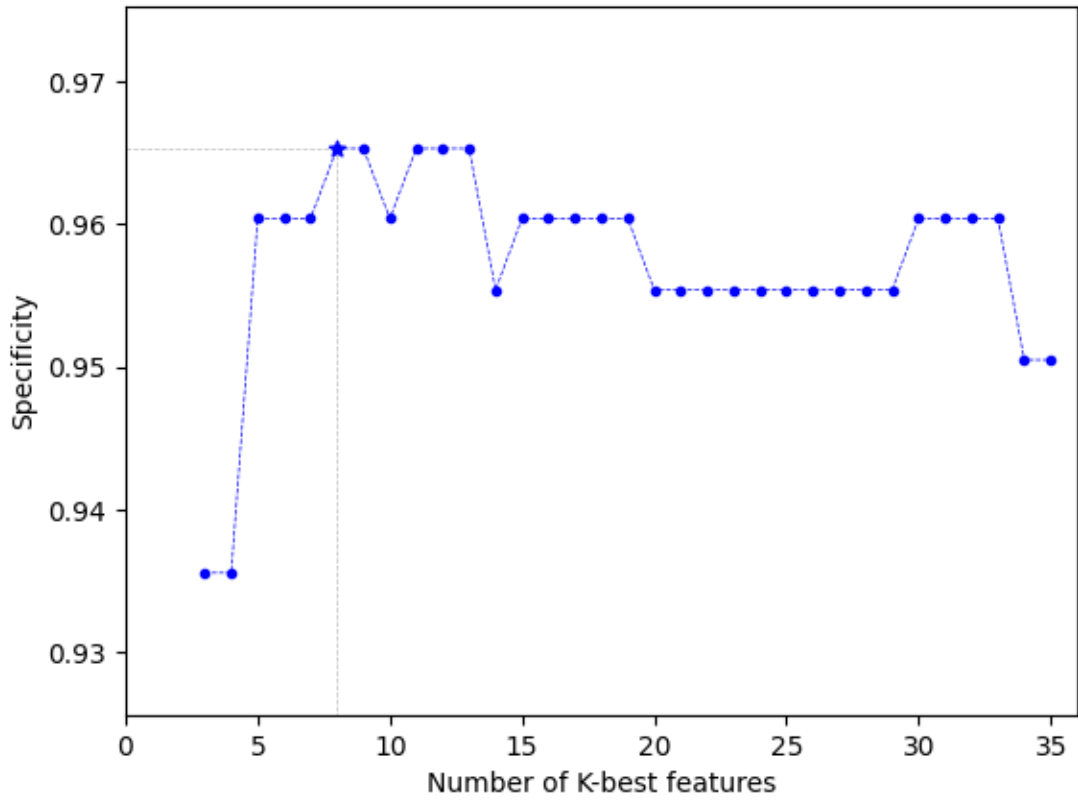
Αποτελέσματα των Ταξινομητών με τη χρήση Διαφορετικών Λιστών Κατάταξης Χαρακτηριστικών



Εικόνα 32: Διαγνωστική ακρίβεια του μοντέλου ταξινόμησης Naive Bayes σύμφωνα με τη λίστα κατάταξης των χαρακτηριστικών εισόδου με χρήση της μεθόδου mRMR

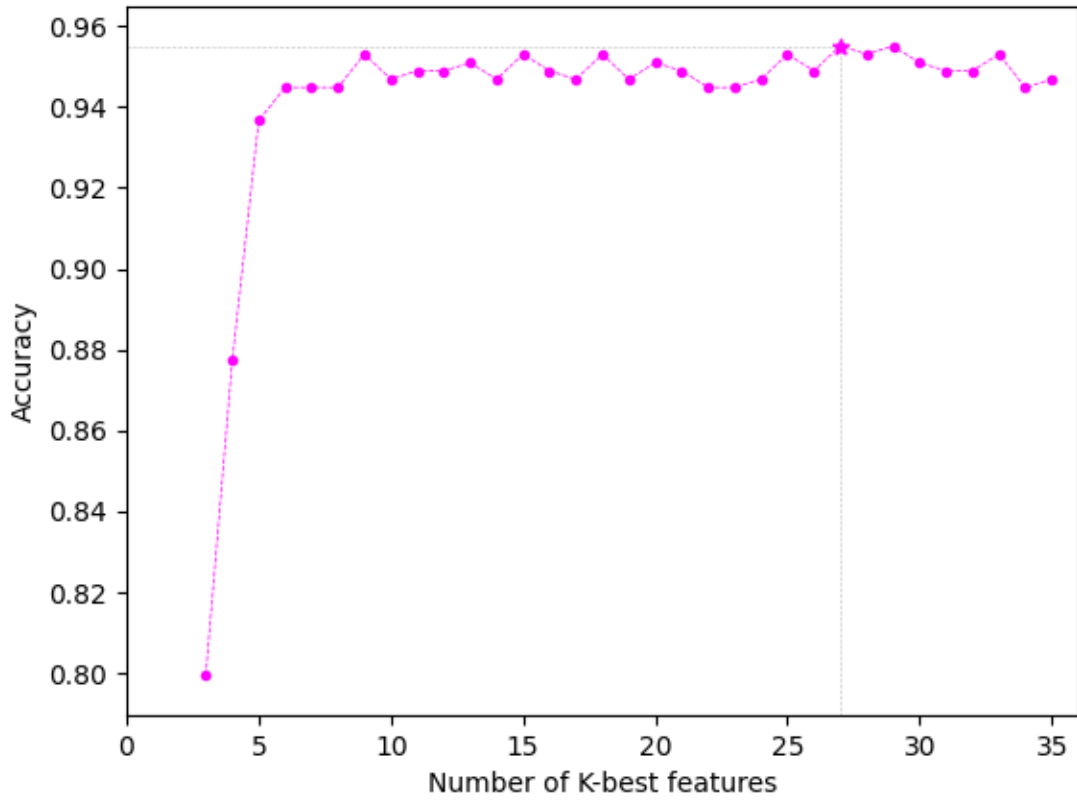


Εικόνα 33: Ευαισθησία του μοντέλου ταξινόμησης Naive Bayes σύμφωνα με τη λίστα κατάταξης των χαρακτηριστικών εισόδου με χρήση της μεθόδου mRMR

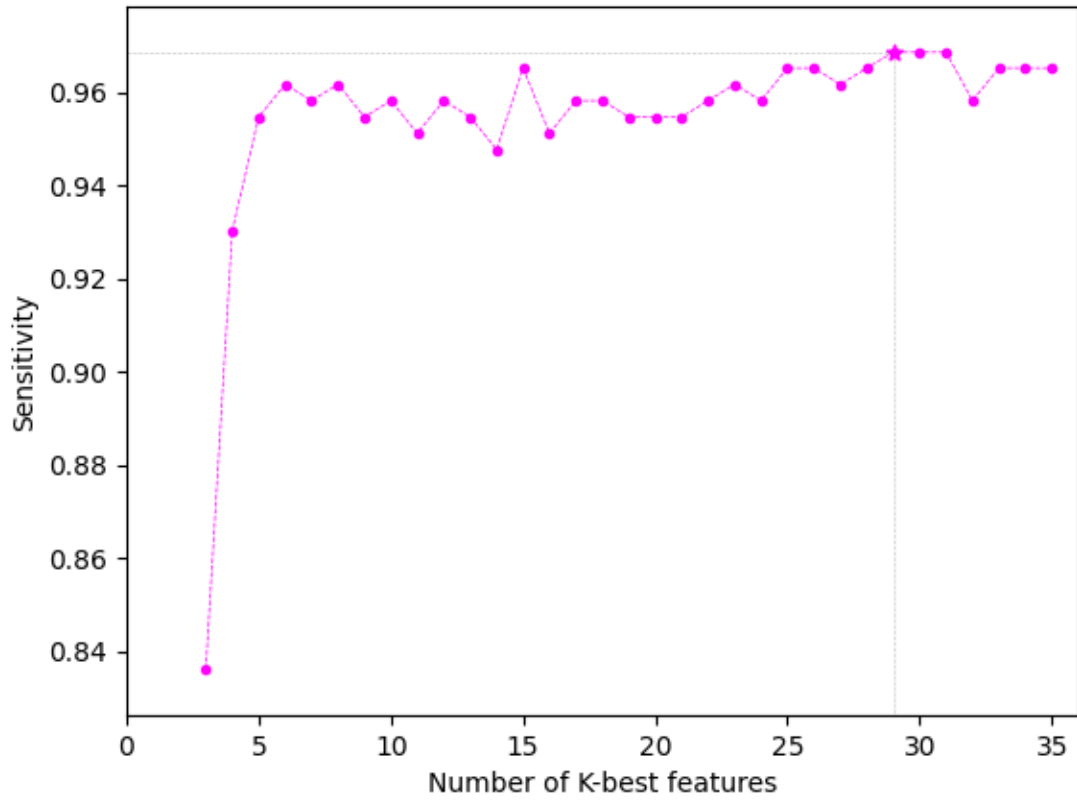


Εικόνα 34: Ειδικότητα του μοντέλου ταξινόμησης Naive Bayes σύμφωνα με τη λίστα κατάταξης των χαρακτηριστικών εισόδου με χρήση της μεθόδου mRMR

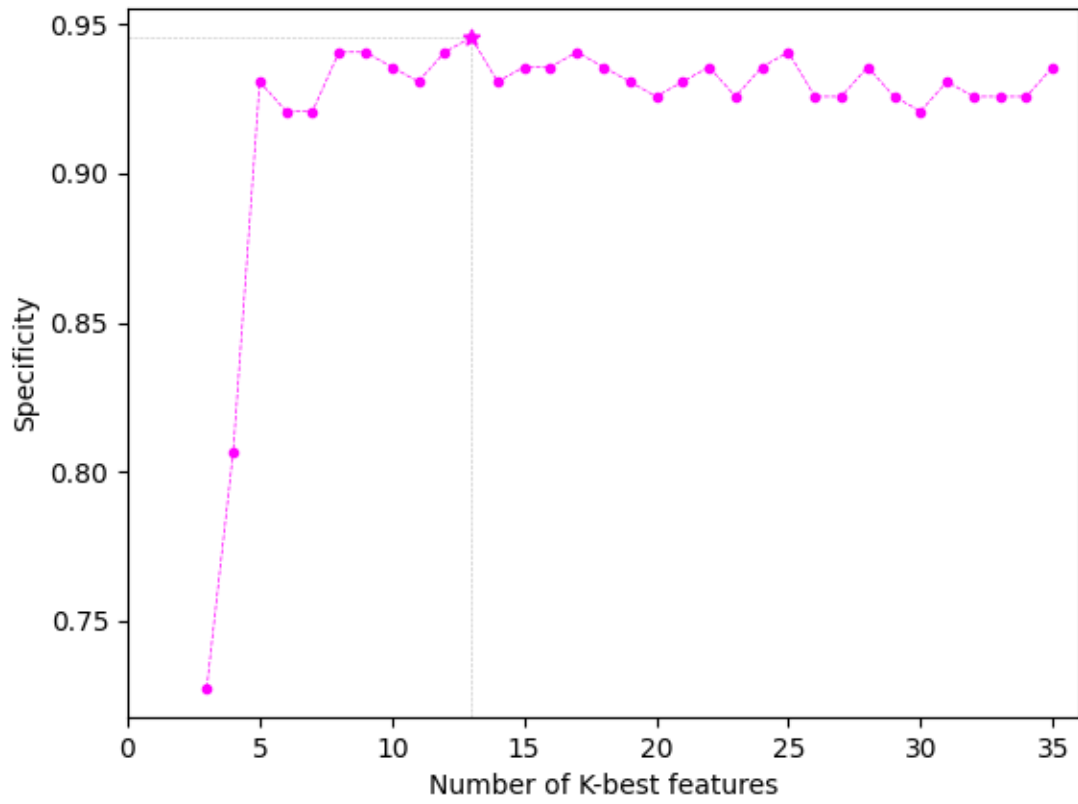
Random Forest



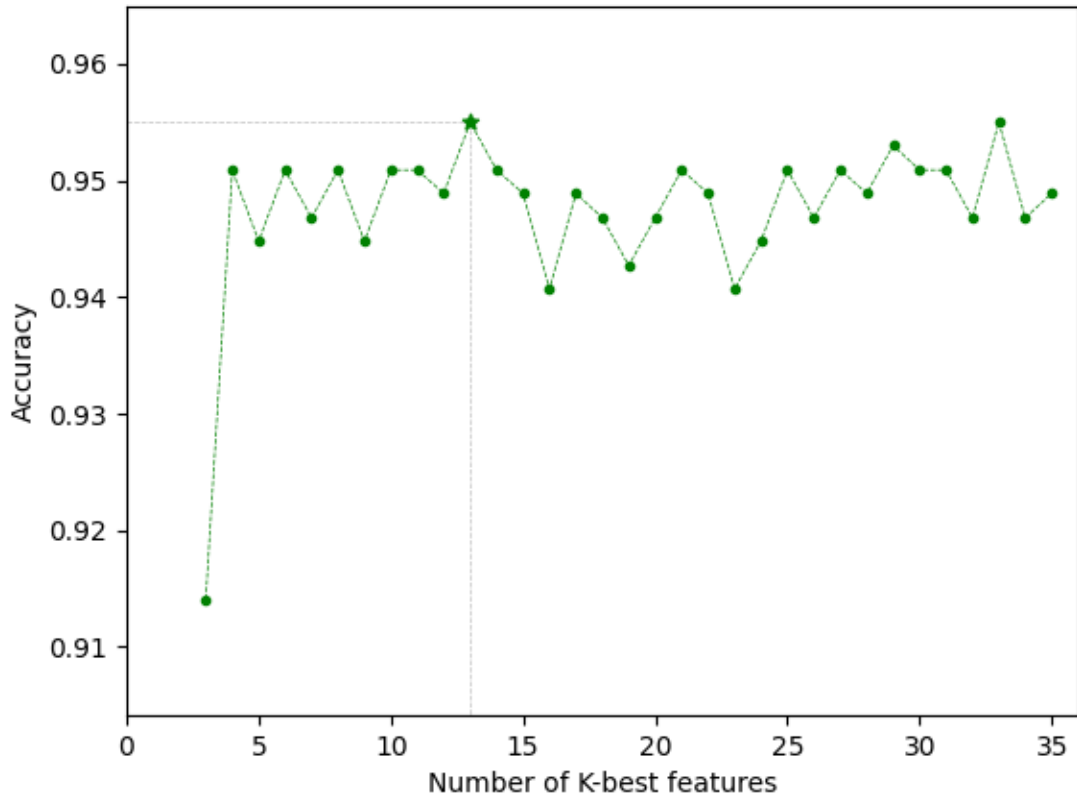
Εικόνα 35: Διαγνωστική ακρίβεια του μοντέλου ταξινόμησης *Random Forest* σύμφωνα με τη λίστα κατάταξης των χαρακτηριστικών εισόδου με χρήση της μεθόδου *Information Gain*



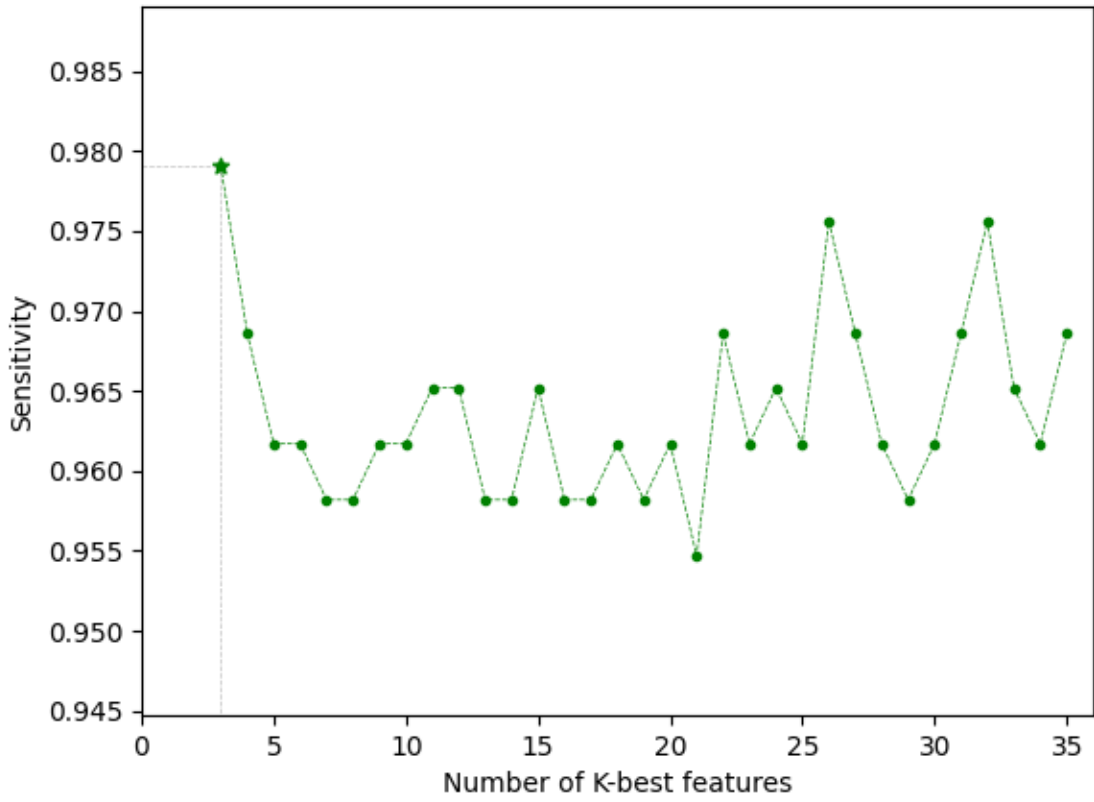
Εικόνα 36: Ευαισθησία του μοντέλου ταξινόμησης *Random Forest* σύμφωνα με τη λίστα κατάταξης των χαρακτηριστικών εισόδου με χρήση της μεθόδου *Information Gain*



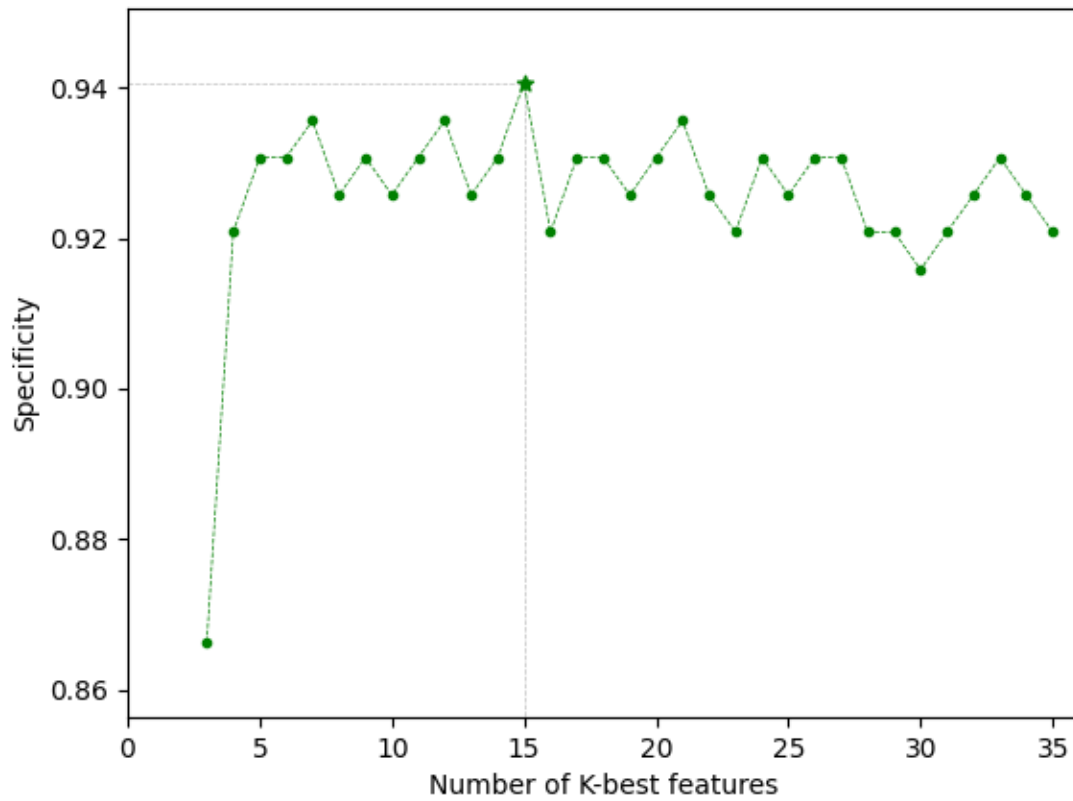
Εικόνα 37: Ειδικότητα του μοντέλου ταξινόμησης *Random Forest* σύμφωνα με τη λίστα κατάταξης των χαρακτηριστικών εισόδου με χρήση της μεθόδου *Information Gain*



Εικόνα 38: Διαγνωστική ακρίβεια του μοντέλου ταξινόμησης *Random Forest* σύμφωνα με τη λίστα κατάταξης των χαρακτηριστικών εισόδου με χρήση του *gini importance*



Εικόνα 39: Ευαισθησία του μοντέλου ταξινόμησης Random Forest σύμφωνα με τη λίστα κατάταξης των χαρακτηριστικών εισόδου με χρήση του gini importance



Εικόνα 40: Ειδικότητα του μοντέλου ταξινόμησης *Random Forest* σύμφωνα με τη λίστα κατάταξης των χαρακτηριστικών εισόδου με χρήση του *gini importance*

Βιβλιογραφία

- [1] H. Sung *et al.*, “Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries,” *CA. Cancer J. Clin.*, vol. 71, no. 3, pp. 209–249, May 2021, doi: 10.3322/CAAC.21660.
- [2] “Cancer Today.” <https://gco.iarc.fr/today/home> (accessed May 28, 2021).
- [3] H. J. Schünemann *et al.*, “Breast cancer screening and diagnosis: A synopsis of the european breast guidelines,” *Ann. Intern. Med.*, vol. 172, no. 1, pp. 46–56, Jan. 2020, doi: 10.7326/M19-2125.
- [4] N. Aristokli, I. Polycarpou, S. C. Themistocleous, D. Sophocleous, and I. Mamais, “Comparison of the diagnostic performance of Magnetic Resonance Imaging (MRI), ultrasound and mammography for detection of breast cancer based on tumor type, breast density and patient’s history: A review,” *Radiography*, Feb. 2022, doi: 10.1016/J.RADI.2022.01.006.
- [5] E. F. Long and P. A. Ganz, “Cost-effectiveness of Universal BRCA1/2 Screening: Evidence-Based Decision Making,” *JAMA Oncol.*, vol. 1, no. 9, pp. 1217–1218, Dec. 2015, doi: 10.1001/JAMAONCOL.2015.2340.
- [6] J. T. Casaubon, S. Tomlinson-Hansen, and J.-P. Regan, “Fine Needle Aspiration Of Breast Masses,” *StatPearls*, Aug. 2021, Accessed: Oct. 07, 2021. [Online]. Available: <https://www.ncbi.nlm.nih.gov/books/NBK470268/>.
- [7] Z. Momenimovahed and H. Salehiniya, “Epidemiological characteristics of and risk factors for breast cancer in the world,” *Breast Cancer (London)*, vol. 11, p. 151, 2019, doi: 10.2147/BCTT.S176070.
- [8] K. Maajani, A. Jalali, S. Alipour, M. Khodadost, H. R. Tohidinik, and K. Yazdani, “The Global and Regional Survival Rate of Women With Breast Cancer: A Systematic Review and Meta-analysis,” *Clin. Breast Cancer*, vol. 19, no. 3, pp. 165–177, Jun. 2019, doi: 10.1016/J.CLBC.2019.01.006.
- [9] R. L. Siegel, K. D. Miller, and A. Jemal, “Cancer statistics, 2019,” *CA. Cancer J. Clin.*, vol. 69, no. 1, pp. 7–34, Jan. 2019, doi: 10.3322/caac.21551.
- [10] L. Iacoviello, M. Bonaccio, G. de Gaetano, and M. B. Donati, “Epidemiology of breast cancer, a paradigm of the ‘common soil’ hypothesis,” *Seminars in Cancer Biology*, vol. 72. Academic Press, pp. 4–10, Jul. 01, 2020, doi: 10.1016/j.semcancer.2020.02.010.
- [11] “Disability-adjusted life years (DALYs).” <https://www.who.int/data/gho/indicator-metadata-registry/imr-details/158> (accessed Jun. 28, 2022).
- [12] E. Patsea, L. Kaklamanis, and A. Batistatou, “The first report of a 5-year period cancer registry in Greece (2009–2013): a pathology-based cancer registry,” *Virchows Arch. 2018 4724*, vol. 472, no. 4, pp. 677–682, Jan. 2018, doi: 10.1007/S00428-017-2287-8.
- [13] R. Jesinger, “Breast anatomy for the interventionalist,” *Tech. Vasc. Interv. Radiol.*, vol. 17, no. 1, pp. 3–9, Mar. 2014, doi: 10.1053/J.TVIR.2013.12.002.
- [14] I. Χατζημπούγιας, *Στοιχεία ανατομικής του ανθρώπου*, 4η έκδ. GM Design, 2007.
- [15] H. Ellis and V. Mahadevan, “Anatomy and physiology of the breast,” *Surg.*, vol. 31, no. 1, pp. 11–

- 14, Jan. 2013, doi: 10.1016/J.MPSUR.2012.10.018.
- [16] K. P. McGuire, “Breast Anatomy and Physiology,” *Breast Dis.*, pp. 1–9, 2019, doi: 10.1007/978-3-030-04606-4_1.
- [17] T. G. Troupis, A. Michalinos, G. Skandalakis, and P. Skandalakis, “Physiology and Developmental Stages of the Breast,” *Breast Cancer Manag. Surg.*, pp. 11–17, 2018, doi: 10.1007/978-3-319-56673-3_2.
- [18] M. Guray and A. A. Sahin, “Benign Breast Diseases: Classification, Diagnosis, and Management; Benign Breast Diseases: Classification, Diagnosis, and Management,” *Breast Cancer Oncol.*, vol. 11, pp. 435–449, 2006, doi: 10.1634/theoncologist.11-5-435.
- [19] M. Muttarak, “Breast calcifications: which are malignant?,” *Singapore Med J*, vol. 50(9) : 90, 2009.
- [20] E. Mallon, P. Osin, N. Nasiri, I. Blain, B. Howard, and B. Gusterson, “The Basic Pathology of Human Breast Cancer,” *J. Mammary Gland Biol. Neoplasia* 2000 52, vol. 5, no. 2, pp. 139–163, 2000, doi: 10.1023/A:1026439204849.
- [21] “Ductal Carcinoma In Situ (DCIS).” <https://www.breastcancer.org/types/ductal-carcinoma-in-situ> (accessed Jun. 28, 2022).
- [22] “Invasive Ductal Carcinoma (IDC).” <https://www.breastcancer.org/types/invasive-ductal-carcinoma> (accessed Jun. 28, 2022).
- [23] “Invasive Lobular Carcinoma (ILC).” <https://www.breastcancer.org/types/invasive-lobular-carcinoma> (accessed Jun. 28, 2022).
- [24] “Lobular Carcinoma in Situ (LCIS) and Atypical Lobular Hyperplasia.” <https://www.breastcancer.org/benign-breast-conditions/lobular-carcinoma-in-situ-atypical-lobular-hyperplasia> (accessed Jun. 28, 2022).
- [25] V. Chippa and H. Barazi, “Inflammatory Breast Cancer,” *Breast Compr. Manag. Benign Malig. Dis.*, pp. 832–838.e2, Jul. 2021, Accessed: Sep. 01, 2021. [Online]. Available: <https://www.ncbi.nlm.nih.gov/books/NBK564324/>.
- [26] “Paget Disease of the Breast.” <https://www.breastcancer.org/types/paget-disease> (accessed Jun. 28, 2022).
- [27] “Understanding Your Pathology Report.” <https://www.breastcancer.org/pathology-report> (accessed Jun. 28, 2022).
- [28] A. Prat *et al.*, “Clinical implications of the intrinsic molecular subtypes of breast cancer,” *The Breast*, vol. 24, pp. S26–S35, Nov. 2015, doi: 10.1016/J.BREAST.2015.07.008.
- [29] “Breast Cancer Stages.” <https://www.breastcancer.org/pathology-report/breast-cancer-stages> (accessed Jun. 28, 2022).
- [30] M. M. Koo, C. von Wagner, G. A. Abel, S. McPhail, G. P. Rubin, and G. Lyratzopoulos, “Typical and atypical presenting symptoms of breast cancer and their associations with diagnostic intervals: Evidence from a national audit of cancer diagnosis,” *Cancer Epidemiol.*, vol. 48, pp. 140–146, Jun. 2017, doi: 10.1016/J.CANEP.2017.04.010.
- [31] “Συμπτώματα - Ελληνική Γυναικολογική Εταιρεία Παθήσεων - ΕΓΕΠΑΜ.” <https://www.egepam.gr/για-το-κοινό/καρκίνος-μαστού/συμπτώματα/> (accessed Jun. 28, 2022).

- [32] “Breast Cancer Symptoms and Signs | CTCA.” <https://www.cancercenter.com/cancer-types/breast-cancer/symptoms> (accessed Jun. 28, 2022).
- [33] “Εταιρεία Ογκολόγων Παθολόγων Ελλάδας - ΕΟΠΕ.” <https://www.hesmo.gr/el/> (accessed Jun. 28, 2022).
- [34] C. Gajdos, P. I. Tartter, I. J. Bleiweiss, C. Bodian, and S. T. Brower, “Stage 0 to stage III breast cancer in young women,” *J. Am. Coll. Surg.*, vol. 190, no. 5, pp. 523–529, May 2000, doi: 10.1016/S1072-7515(00)00257-X.
- [35] A. Toriola and G. Colditz, “Trends in breast cancer incidence and mortality in the United States: Implications for prevention,” *Breast Cancer Research and Treatment*, vol. 138, no. 3. Breast Cancer Res Treat, pp. 665–673, Apr. 2013, doi: 10.1007/s10549-013-2500-7.
- [36] C. DeSantis, R. Siegel, P. Bandi, and A. Jemal, “Breast cancer statistics, 2011,” *CA. Cancer J. Clin.*, vol. 61, no. 6, pp. 408–418, Nov. 2011, doi: 10.3322/caac.20134.
- [37] B. C. Yankaskas *et al.*, “Performance of First Mammography Examination in Women Younger Than 40 Years,” *JNCI J. Natl. Cancer Inst.*, vol. 102, no. 10, pp. 692–701, May 2010, doi: 10.1093/JNCI/DJQ090.
- [38] H. D. Nelson *et al.*, “Risk factors for breast cancer for women aged 40 to 49 years: A systematic review and meta-analysis a systematic review and meta-analysis,” *Ann. Intern. Med.*, vol. 156, no. 9, pp. 635–648, May 2012, doi: 10.7326/0003-4819-156-9-201205010-00006.
- [39] R. Verma, R. L. Bowen, S. E. Slater, F. Mihaimed, and J. L. Jones, “Pathological and epidemiological factors associated with advanced stage at diagnosis of breast cancer,” *British Medical Bulletin*, vol. 103, no. 1. Br Med Bull, pp. 129–145, Sep. 2012, doi: 10.1093/bmb/lds018.
- [40] W.-C. Chie *et al.*, “Age at Any Full-term Pregnancy and Breast Cancer Risk,” *Am. J. Epidemiol.*, vol. 151, no. 7, pp. 715–722, Apr. 2000, doi: 10.1093/OXFORDJOURNALS.AJE.A010266.
- [41] M. Lambe, C. Hsieh, H. Chan, A. Ekblom, D. Trichopoulos, and H.-O. Adami, “Parity, age at first and last birth, and risk of breast cancer: A population-based study in Sweden,” *Breast Cancer Res. Treat.* 1996 383, vol. 38, no. 3, pp. 305–311, 1996, doi: 10.1007/BF01806150.
- [42] P. F. Slepicka, S. L. Cyrill, and C. O. dos Santos, “Pregnancy and Breast Cancer: Pathways to Understand Risk and Prevention,” *Trends Mol. Med.*, vol. 25, no. 10, pp. 866–881, Oct. 2019, doi: 10.1016/J.MOLMED.2019.06.003.
- [43] G. Erlandsson, S. M. Montgomery, S. Cnattingius, and A. Ekblom, “Abortions and breast cancer: Record-based case-control study,” *Int. J. Cancer*, vol. 103, no. 5, pp. 676–679, Feb. 2003, doi: 10.1002/IJC.10877.
- [44] G. K. Reeves *et al.*, “Breast cancer risk in relation to abortion: Results from the EPIC study,” *Int. J. Cancer*, vol. 119, no. 7, pp. 1741–1745, Oct. 2006, doi: 10.1002/IJC.22001.
- [45] M. S. R. Shawon, M. Eriksson, and J. Li, “Body size in early life and risk of breast cancer,” *Breast Cancer Res.*, vol. 19, no. 1, Jul. 2017, doi: 10.1186/S13058-017-0875-9.
- [46] G. A. Colditz, “Decline in breast cancer incidence due to removal of promoter: Combination estrogen plus progestin,” *Breast Cancer Research*, vol. 9, no. 4. Breast Cancer Res, Jul. 26, 2007, doi: 10.1186/bcr1736.
- [47] V. Beral, G. Reeves, D. Bull, and J. Green, “Breast cancer risk in relation to the interval between menopause and starting hormone therapy,” *J. Natl. Cancer Inst.*, vol. 103, no. 4, pp. 296–305, Feb.

- 2011, doi: 10.1093/jnci/djq527.
- [48] S. C *et al.*, “Trends in breast cancer by race and ethnicity: update 2006,” *CA. Cancer J. Clin.*, vol. 56, no. 3, pp. 168–183, May 2006, doi: 10.3322/CANJCLIN.56.3.168.
- [49] Z. Q. Tao, A. Shi, C. Lu, T. Song, Z. Zhang, and J. Zhao, “Breast Cancer: Epidemiology and Etiology,” *Cell Biochem. Biophys.*, vol. 72, no. 2, pp. 333–338, Jun. 2015, doi: 10.1007/s12013-014-0459-6.
- [50] C. DeSantis, R. Siegel, P. Bandi, and A. Jemal, “Breast cancer statistics, 2011,” *CA. Cancer J. Clin.*, vol. 61, no. 6, pp. 408–418, Nov. 2011, doi: 10.3322/CAAC.20134.
- [51] H. Baer, S. Tworoger, S. Hankinson, and W. Willett, “Body fatness at young ages and risk of breast cancer throughout life,” *Am. J. Epidemiol.*, vol. 171, no. 11, pp. 1183–1194, 2010, doi: 10.1093/AJE/KWQ045.
- [52] J. R. Palmer, L. L. Adams-Campbell, D. A. Boggs, L. A. Wise, and L. Rosenberg, “A prospective study of body size and breast cancer in black women,” *Cancer Epidemiol. Biomarkers Prev.*, vol. 16, no. 9, pp. 1795–1802, Sep. 2007, doi: 10.1158/1055-9965.EPI-07-0336.
- [53] H. J. Baer *et al.*, “Body fatness during childhood and adolescence and incidence of breast cancer in premenopausal women: A prospective cohort study,” *Breast Cancer Res.*, vol. 7, no. 3, Feb. 2005, doi: 10.1186/BCR998.
- [54] E. V. Bandera *et al.*, “Body size in early life and breast cancer risk in African American and European American women,” *Cancer Causes Control*, vol. 24, no. 12, p. 2231, 2013, doi: 10.1007/S10552-013-0302-1.
- [55] M. Ahlgren, M. Melbye, J. Wohlfahrt, and T. Sørensen, “Growth patterns and the risk of breast cancer in women,” *N. Engl. J. Med.*, vol. 351, no. 16, pp. 1619–1626, Oct. 2004, doi: 10.1056/NEJMOA040576.
- [56] J. Li, K. Humphreys, L. Eriksson, K. Czene, J. Liu, and P. Hall, “Effects of childhood body size on breast cancer tumour characteristics,” *Breast Cancer Res.*, vol. 12, no. 2, p. R23, Apr. 2010, doi: 10.1186/BCR2564.
- [57] G. Fagherazzi, G. Guillas, M. C. Boutron-Ruault, F. Clavel-Chapelon, and S. Mesrine, “Body shape throughout life and the risk for breast cancer at adulthood in the French E3N cohort,” *Eur. J. Cancer Prev.*, vol. 22, no. 1, pp. 29–37, Jan. 2013, doi: 10.1097/CEJ.0B013E328355EC04.
- [58] J. Wit, H. Balen, G. Kamp, and W. Oostdijk, “Benefit of postponing normal puberty for improving final height,” *Eur. J. Endocrinol.*, vol. 151 Suppl, no. SUPPL. 1, Aug. 2004, doi: 10.1530/EJE.0.151S041.
- [59] E. Couto *et al.*, “Hormone therapy use and mammographic density in postmenopausal Norwegian women,” *Breast Cancer Res. Treat.*, vol. 132, no. 1, pp. 297–305, Feb. 2012, doi: 10.1007/S10549-011-1810-X.
- [60] N. Boyd, G. Lockwood, J. Byng, ... D. T.-C. E., and U. 1998, “Mammographic densities and breast cancer risk.,” *AACR*, 2007, Accessed: Sep. 10, 2021. [Online]. Available: <https://cebp.aacrjournals.org/content/7/12/1133.short>.
- [61] K. Rojas and A. Stuckey, “Breast Cancer Epidemiology and Risk Factors,” *Clin. Obstet. Gynecol.*, vol. 59, no. 4, pp. 651–672, 2016, doi: 10.1097/GRF.0000000000000239.
- [62] L. Jansen, R. Backstein, and M. Brown, “Breast size and breast cancer: a systematic review,” *J.*

- Plast. Reconstr. Aesthet. Surg.*, vol. 67, no. 12, pp. 1615–1623, 2014, doi: 10.1016/J.BJPS.2014.10.001.
- [63] A. Markkula *et al.*, “Given breast cancer, does breast size matter? Data from a prospective breast cancer cohort,” *Cancer Causes Control*, vol. 23, no. 8, pp. 1307–1316, Aug. 2012, doi: 10.1007/S10552-012-0008-9.
- [64] C. Hsieh and D. Trichopoulos, “Breast size, handedness and breast cancer risk,” *Eur. J. Cancer*, vol. 27, no. 2, pp. 131–135, 1991, doi: 10.1016/0277-5379(91)90469-T.
- [65] C. Sellahewa, P. Nightingale, and A. R. Carmichael, “Women with large breasts are at an increased risk of advanced breast cancer,” *Int. Semin. Surg. Oncol.*, vol. 5, p. 16, Jun. 2008, doi: 10.1186/1477-7800-5-16.
- [66] R. J. MacInnis, D. R. English, A. M. Haydon, J. L. Hopper, D. M. Gertig, and G. G. Giles, “Body size and composition and risk of rectal cancer (Australia),” *Cancer Causes Control*, vol. 17, no. 10, pp. 1291–1297, 2006, doi: 10.1007/s10552-006-0074-y.
- [67] E. John, A. Phipps, and M. Sangaramoorthy, “Body size, modifying factors, and postmenopausal breast cancer risk in a multiethnic population: the San Francisco Bay Area Breast Cancer Study,” *Springerplus*, vol. 2, no. 1, pp. 1–20, 2013, doi: 10.1186/2193-1801-2-239.
- [68] J. L. Hopper *et al.*, “Age-specific breast cancer risk by body mass index and familial risk: prospective family study cohort (ProF-SC),” *Breast Cancer Res. 2018 201*, vol. 20, no. 1, pp. 1–11, Nov. 2018, doi: 10.1186/S13058-018-1056-1.
- [69] Y. Guo *et al.*, “Genetically Predicted Body Mass Index and Breast Cancer Risk: Mendelian Randomization Analyses of Data from 145,000 Women of European Descent,” *PLOS Med.*, vol. 13, no. 8, p. e1002105, Aug. 2016, doi: 10.1371/JOURNAL.PMED.1002105.
- [70] E. I. Papageorgiou, Jayashree Subramanian, A. Karmegam, and N. Papandrianos, “A risk management model for familial breast cancer: A new application using Fuzzy Cognitive Map method,” *Comput. Methods Programs Biomed.*, vol. 122, no. 2, pp. 123–135, Nov. 2015, doi: 10.1016/J.CMPB.2015.07.003.
- [71] N. Mavaddat *et al.*, “Prediction of breast cancer risk based on profiling with common genetic variants,” *J. Natl. Cancer Inst.*, vol. 107, no. 5, May 2015, doi: 10.1093/JNCI/DJV036.
- [72] G. Dite, R. MacInnis, A. Bickerstaffe, J. Dowty, and R. Allman, “Breast Cancer Risk Prediction Using Clinical Models and 77 Independent Risk-Associated SNPs for Women Aged Under 50 Years: Australian Breast Cancer Family Registry,” *Cancer Epidemiol. Biomarkers Prev.*, vol. 25, no. 2, pp. 359–365, Feb. 2016, doi: 10.1158/1055-9965.EPI-15-0838.
- [73] D. F. Easton *et al.*, “Gene-Panel Sequencing and the Prediction of Breast-Cancer Risk,” <http://dx.doi.org/10.1056/NEJMSr1501341>, vol. 372, no. 23, pp. 2243–2257, Jun. 2015, doi: 10.1056/NEJMSR1501341.
- [74] T. Nguyen-Dumont *et al.*, “Hi-Plex for high-throughput mutation screening: application to the breast cancer susceptibility gene PALB2,” *BMC Med. Genomics*, vol. 6, no. 1, p. 48, 2013, doi: 10.1186/1755-8794-6-48.
- [75] J. L. Hopper, “Genetics for population and public health,” *Int. J. Epidemiol.*, vol. 46, no. 1, pp. 8–11, Feb. 2017, doi: 10.1093/IJE/DYX008.
- [76] A. C. Antoniou *et al.*, “The BOADICEA model of genetic susceptibility to breast and ovarian cancers: updates and extensions,” *Br. J. Cancer 2008 988*, vol. 98, no. 8, pp. 1457–1466, Mar.

- 2008, doi: 10.1038/sj.bjc.6604305.
- [77] “The BRCA1 and BRCA2 Genes | CDC.” https://www.cdc.gov/genomics/disease/breast_ovarian_cancer/genes_hboc.htm (accessed Jun. 28, 2022).
- [78] G. Dite *et al.*, “Testing for Gene-Environment Interactions Using a Prospective Family Cohort Design: Body Mass Index in Early and Later Adulthood and Risk of Breast Cancer,” *Am. J. Epidemiol.*, vol. 185, no. 6, pp. 487–500, Mar. 2017, doi: 10.1093/AJE/KWW241.
- [79] E. R. Manahan *et al.*, “Consensus Guidelines on Genetic Testing for Hereditary Breast Cancer from the American Society of Breast Surgeons,” *Ann. Surg. Oncol.*, vol. 26, no. 10, pp. 3025–3031, Oct. 2019, doi: 10.1245/S10434-019-07549-8/TABLES/1.
- [80] I. Godet and D. M. Gilkes, “BRCA1 and BRCA2 mutations and treatment strategies for breast cancer,” *Integr. cancer Sci. Ther.*, vol. 4, no. 1, 2017, doi: 10.15761/ICST.1000228.
- [81] K. B. Kuchenbaecker *et al.*, “Risks of Breast, Ovarian, and Contralateral Breast Cancer for BRCA1 and BRCA2 Mutation Carriers,” *JAMA*, vol. 317, no. 23, pp. 2402–2416, Jun. 2017, doi: 10.1001/JAMA.2017.7112.
- [82] F. Ahmed *et al.*, “Mutations in Human Interferon $\alpha 2b$ Gene and Potential as Risk Factor Associated with Female Breast Cancer,” *Cancer Biother. Radiopharm.*, vol. 31, no. 6, pp. 199–208, Aug. 2016, doi: 10.1089/CBR.2016.2046.
- [83] K. Yari, Z. Rahimi, M. T. Moradi, and Z. Rahimi, “The MMP-2 -735 C allele is a risk factor for susceptibility to breast cancer,” *Asian Pac. J. Cancer Prev.*, vol. 15, no. 15, pp. 6199–6203, 2014, doi: 10.7314/APJCP.2014.15.15.6199.
- [84] A. Lee, N. Mavaddat, A. Wilcox, ... A. C.-G. in, and undefined 2019, “BOADICEA: a comprehensive breast cancer risk prediction model incorporating genetic and nongenetic risk factors,” *nature.com*, Accessed: Apr. 04, 2022. [Online]. Available: <https://www.nature.com/articles/s41436-018-0406-9>.
- [85] C. Catsburg, V. A. Kirsh, C. L. Soskolne, N. Kreiger, and T. E. Rohan, “Active cigarette smoking and the risk of breast cancer: a cohort study,” *Cancer Epidemiol.*, vol. 38, no. 4, pp. 376–381, Aug. 2014, doi: 10.1016/J.CANEP.2014.05.007.
- [86] H. Han *et al.*, “Hypertension and breast cancer risk: a systematic review and meta-analysis,” *Sci. Reports 2017 71*, vol. 7, no. 1, pp. 1–9, Mar. 2017, doi: 10.1038/srep44877.
- [87] M. Touvier *et al.*, “Cholesterol and breast cancer risk: a systematic review and meta-analysis of prospective studies,” *Br. J. Nutr.*, vol. 114, no. 3, pp. 347–357, Aug. 2015, doi: 10.1017/S000711451500183X.
- [88] J. L. P. Kuijpers, I. Nykličtek, M. W. J. Louwman, T. A. P. Weetman, V. J. M. Pop, and J. W. W. Coebergh, “Hypothyroidism might be related to breast cancer in post-menopausal women,” *Thyroid*, vol. 15, no. 11, pp. 1253–1259, Nov. 2005, doi: 10.1089/thy.2005.15.1253.
- [89] C. Ortega-Olvera *et al.*, “Thyroid hormones and breast cancer association according to menopausal status and body mass index,” *Breast Cancer Res.*, vol. 20, no. 1, pp. 1–14, Aug. 2018, doi: 10.1186/s13058-018-1017-8.
- [90] A. Angelousi *et al.*, “Is there an association between thyroid function abnormalities and breast cancer?,” *Arch. Endocrinol. Metab.*, vol. 61, no. 1, pp. 54–61, Jan. 2017, doi: 10.1590/2359-3997000000191.

- [91] K. L. Maughan, M. A. Lutterbie, and P. S. Ham, "Treatment of breast cancer," *Am. Fam. Physician*, vol. 81, no. 11, pp. 1339–1346, Jun. 2010, doi: 10.1056/nejm199810013391407.
- [92] G. Tinoco, S. Warsch, S. Glück, K. Avancha, and A. J. Montero, "Treating Breast Cancer in the 21st Century: Emerging Biological Therapies," *J. Cancer*, vol. 4, no. 2, p. 117, 2013, doi: 10.7150/JCA.4925.
- [93] J. Zgajnar, "Clinical Presentation, Diagnosis and Staging of Breast Cancer," in *Breast Cancer Management for Surgeons*, Springer International Publishing, 2018, pp. 159–176.
- [94] S. Winters, C. Martin, D. Murphy, and N. K. Shokar, "Breast Cancer Epidemiology, Prevention, and Screening," in *Progress in Molecular Biology and Translational Science*, vol. 151, Elsevier B.V., 2017, pp. 1–32.
- [95] "American Cancer Society | Information and Resources about for Cancer: Breast, Colon, Lung, Prostate, Skin." <https://www.cancer.org/cancer/breast-cancer/risk-and-prevention/genetic-testing.html>. (accessed Jun. 28, 2022).
- [96] A. Vourtsis, "Screening for High-Familial-Risk Women," in *Breast Cancer Management for Surgeons*, Springer International Publishing, 2018, pp. 59–67.
- [97] P. Steyerova, "Imaging of the Breast," in *Breast Cancer Management for Surgeons*, Cham: Springer International Publishing, 2018, pp. 127–145.
- [98] E. A. Sickles, C. M. Appleton, E. S. Burnside, and S. C. Gavenonis, "Breast imaging reporting and data system (BI-RADS) Atlas- Ultrasound 5th edn.," *Am. Coll. Radiol. BI-RADS-Atlas*, pp. 121–132, 2013, [Online]. Available: <https://www.acr.org/-/media/ACR/Files/RADS/BI-RADS/US-Reporting.pdf>.
- [99] A. Diamantis, E. Magiorkinis, and H. Koutselini, "Fine-needle aspiration (FNA) biopsy: Historical aspects," *Folia Histochemica et Cytobiologica*, vol. 47, no. 2. Via Medica, pp. 191–197, 2009, doi: 10.2478/v10042-009-0027-x.
- [100] S. M. Willems, C. H. M. Van Deurzen, and P. J. Van Diest, "Diagnosis of breast lesions: Fine-needle aspiration cytology or core needle biopsy? A review," *Journal of Clinical Pathology*, vol. 65, no. 4. BMJ Publishing Group, pp. 287–292, Apr. 01, 2012, doi: 10.1136/jclinpath-2011-200410.
- [101] M. L. Huang, B. E. Adrada, R. Candelaria, D. Thames, D. Dawson, and W. T. Yang, "Stereotactic Breast Biopsy: Pitfalls and Pearls," *Tech. Vasc. Interv. Radiol.*, vol. 17, no. 1, pp. 32–39, Mar. 2014, doi: 10.1053/J.TVIR.2013.12.006.
- [102] O. Iwuagwu and P. Drew, "Vacuum-assisted biopsy device-diagnostic and therapeutic applications in breast surgery," *Breast*, vol. 13, no. 6, pp. 483–487, 2004, doi: 10.1016/J.BREAST.2004.06.004.
- [103] V. J. Zannis and K. M. Aliano, "The evolving practice pattern of the breast surgeon with disappearance of open biopsy for nonpalpable lesions," *Am. J. Surg.*, vol. 176, no. 6, pp. 525–528, 1998, doi: 10.1016/S0002-9610(98)00265-7.
- [104] M. A. Benito and J. C. Herrero, "Breast Imaging," *Learn. Diagnostic Imaging 100 Essent. Cases*, pp. 1–25, 2008, doi: 10.1007/978-3-540-71207-7_1.
- [105] E. R. Manahan *et al.*, "Consensus Guidelines on Genetic Testing for Hereditary Breast Cancer from the American Society of Breast Surgeons," *Ann. Surg. Oncol.*, vol. 26, no. 10, pp. 3025–3031, Oct. 2019, doi: 10.1245/S10434-019-07549-8/TABLES/1.

- [106] T. B. Bevers *et al.*, “NCCN clinical practice guidelines in oncology: breast cancer screening and diagnosis,” *J. Natl. Compr. Canc. Netw.*, vol. 7, no. 10, pp. 1060–1096, 2009, doi: 10.6004/JNCCN.2009.0070.
- [107] T. Bevers *et al.*, “Screening and Early Detection,” *Abeloff’s Clin. Oncol.*, pp. 375–398.e7, Jan. 2019, doi: 10.1016/B978-0-323-47674-4.00023-2.
- [108] M. Golshan, “Examination Techniques,” in *The Breast*, Elsevier, 2009, pp. 605–611.
- [109] D. Saslow *et al.*, “Clinical Breast Examination: Practical Recommendations for Optimizing Performance and Reporting,” *CA. Cancer J. Clin.*, vol. 54, no. 6, pp. 327–344, Nov. 2004, doi: 10.3322/canjclin.54.6.327.
- [110] M. Jadhav *et al.*, “Performance of CBE and mammography for breast cancer screening in Indian setup,” *Indian J. Public Heal. Res. Dev.*, vol. 8, no. 4, pp. 700–707, Oct. 2017, doi: 10.5958/0976-5506.2017.00418.1.
- [111] J. Mathew and M. Sibbering, “Breast Cancer Screening,” in *Breast Cancer Management for Surgeons*, Cham: Springer International Publishing, 2018, pp. 147–156.
- [112] T. McCready, D. Littlewood, and J. Jenkinson, “Breast self-examination and breast awareness: a literature review,” *J. Clin. Nurs.*, vol. 14, no. 5, pp. 570–578, May 2005, doi: 10.1111/J.1365-2702.2004.01108.X.
- [113] A. B. Miller and C. J. Baines, “The role of clinical breast examination and breast self-examination,” *Prev. Med. (Baltim.)*, vol. 53, no. 3, pp. 118–120, Sep. 2011, doi: 10.1016/J.YPMED.2011.05.001.
- [114] M. Turan *et al.*, “Practical Utility of Diagnostic Clinical Breast Examination in the Diagnosis of Breast Cancer,” *Cureus*, vol. 13, no. 9, Sep. 2021, doi: 10.7759/CUREUS.17662.
- [115] W. E. Barlow *et al.*, “Performance of Diagnostic Mammography for Women With Signs or Symptoms of Breast Cancer,” *JNCI J. Natl. Cancer Inst.*, vol. 94, no. 15, pp. 1151–1159, Aug. 2002, doi: 10.1093/JNCI/94.15.1151.
- [116] H. M. Zonderland, “The role of ultrasound in the diagnosis of breast cancer,” *Semin. Ultrasound, CT MRI*, vol. 21, no. 4, pp. 317–324, Aug. 2000, doi: 10.1016/S0887-2171(00)90026-X.
- [117] F. PP, E. VF, T. LC, and V. RJ, “Diagnostic accuracy of the fine needle aspiration cytology and core needle biopsy as a diagnostic method for breast lesions,” *Rev. Bras. Ginecol. Obstet.*, vol. 33, no. 3, pp. 139–143, Mar. 2011, doi: 10.1590/S0100-72032011000300007.
- [118] K. M. Kelly, J. Dean, W. S. Comulada, and S.-J. Lee, “Breast cancer detection using automated whole breast ultrasound and mammography in radiographically dense breasts,” *Eur. Radiol. 2009*, vol. 20, no. 3, pp. 734–742, Sep. 2009, doi: 10.1007/S00330-009-1588-Y.
- [119] L. Irwig, N. Houssami, and C. Van Vliet, “New technologies in screening for breast cancer: a systematic review of their accuracy,” *Br. J. Cancer*, vol. 90, no. 11, pp. 2118–2122, Jun. 2004, doi: 10.1038/SJ.BJC.6601836.
- [120] S. J. Lord *et al.*, “A systematic review of the effectiveness of magnetic resonance imaging (MRI) as an addition to mammography and ultrasound in screening young women at high risk of breast cancer,” *Eur. J. Cancer*, vol. 43, no. 13, pp. 1905–1917, Sep. 2007, doi: 10.1016/J.EJCA.2007.06.007.
- [121] I. Hadadi, W. Rae, J. Clarke, M. McEntee, and E. Ekpo, “Diagnostic Performance of Adjunctive

- Imaging Modalities Compared to Mammography Alone in Women with Non-Dense and Dense Breasts: A Systematic Review and Meta-Analysis,” *Clin. Breast Cancer*, vol. 21, no. 4, pp. 278–291, Aug. 2021, doi: 10.1016/J.CLBC.2021.03.006.
- [122] N. Peters, I. Borel Rinkes, N. Zuithoff, W. Mali, K. Moons, and P. Peeters, “Meta-analysis of MR imaging in the diagnosis of breast lesions,” *Radiology*, vol. 246, no. 1, pp. 116–124, Jan. 2008, doi: 10.1148/radiol.2461061298.
- [123] J. A. Cintolo-Gonzalez *et al.*, “Breast cancer risk models: a comprehensive overview of existing models, validation, and clinical applications,” *Breast Cancer Res. Treat.*, vol. 164, doi: 10.1007/s10549-017-4247-z.
- [124] M. Gail *et al.*, “Projecting individualized probabilities of developing breast cancer for white females who are being examined annually,” *J. Natl. Cancer Inst.*, vol. 81, no. 24, pp. 1879–1886, Dec. 1989, doi: 10.1093/JNCI/81.24.1879.
- [125] X. Wang, Y. Huang, L. Li, H. Dai, F. Song, and K. Chen, “Assessment of performance of the Gail model for predicting breast cancer risk: A systematic review and meta-analysis with trial sequential analysis,” *Breast Cancer Res.*, vol. 20, no. 1, pp. 1–19, Mar. 2018, doi: 10.1186/S13058-018-0947-5/TABLES/3.
- [126] R. Lindsey *et al.*, “Deep neural network improves fracture detection by clinicians,” *Proc. Natl. Acad. Sci.*, vol. 115, no. 45, pp. 11591–11596, Nov. 2018, doi: 10.1073/PNAS.1806905115.
- [127] W. Barlow *et al.*, “Prospective breast cancer risk prediction model for women undergoing screening mammography,” *J. Natl. Cancer Inst.*, vol. 98, no. 17, pp. 1204–1214, Sep. 2006, doi: 10.1093/JNCI/DJJ331.
- [128] N. F. Boyd, L. J. Martin, M. J. Yaffe, and S. Minkin, “Mammographic density: A harmonally responsive risk factor for breast cancer,” *J. Br. Menopause Soc.*, vol. 12, no. 4, pp. 186–193, Dec. 2006, doi: 10.1258/136218006779160436.
- [129] J. A. Tice, S. R. Cummings, R. Smith-Bindman, L. Ichikawa, W. E. Barlow, and K. Kerlikowske, “Using Clinical Factors and Mammographic Breast Density to Estimate Breast Cancer Risk: Development and Validation of a New Predictive Model,” *Ann. Intern. Med.*, vol. 148, no. 5, p. 337, Mar. 2008, Accessed: Sep. 17, 2021. [Online]. Available: /pmc/articles/PMC2674327/.
- [130] J. Tice, S. Cummings, E. Ziv, and K. Kerlikowske, “Mammographic breast density and the Gail model for breast cancer risk prediction in a screening population,” *Breast Cancer Res. Treat.*, vol. 94, no. 2, pp. 115–122, Nov. 2005, doi: 10.1007/S10549-005-5152-4.
- [131] C. Vachon *et al.*, “The contributions of breast density and common genetic variation to breast cancer risk,” *J. Natl. Cancer Inst.*, vol. 107, no. 5, May 2015, doi: 10.1093/JNCI/DJU397.
- [132] S. E. Singletary, “Rating the Risk Factors for Breast Cancer,” *Ann. Surg.*, vol. 237, no. 4, p. 474, 2003, doi: 10.1097/01.SLA.0000059969.64262.87.
- [133] V. S. Pankratz *et al.*, “Model for Individualized Prediction of Breast Cancer Risk After a Benign Breast Biopsy,” *J. Clin. Oncol.*, vol. 33, no. 8, p. 923, Mar. 2015, doi: 10.1200/JCO.2014.55.4865.
- [134] L. C. Hartmann *et al.*, “Benign Breast Disease and the Risk of Breast Cancer,” <http://dx.doi.org/10.1056/NEJMoa044383>, vol. 353, no. 3, pp. 229–237, Oct. 2009, doi: 10.1056/NEJMOA044383.
- [135] B. Rosner and G. Colditz, “Nurses’ health study: log-incidence mathematical model of breast cancer incidence,” *J. Natl. Cancer Inst.*, vol. 88, no. 6, pp. 359–364, Mar. 1996, doi:

10.1093/JNCI/88.6.359.

- [136] G. A. Colditz and B. Rosner, “Cumulative Risk of Breast Cancer to Age 70 Years According to Risk Factor Status: Data from the Nurses’ Health Study,” *Am. J. Epidemiol.*, vol. 152, no. 10, pp. 950–964, Nov. 2000, doi: 10.1093/AJE/152.10.950.
- [137] B. Rosner, G. Colditz, S. Hankinson, J. Sullivan-Halley, J. Lacey, and L. Bernstein, “Validation of Rosner-Colditz breast cancer incidence model using an independent data set, the California Teachers Study,” *Breast Cancer Res. Treat.*, vol. 142, no. 1, pp. 187–202, Nov. 2013, doi: 10.1007/S10549-013-2719-3.
- [138] S. S. Tworoger *et al.*, “Inclusion of Endogenous Hormone Levels in Risk Prediction Models of Postmenopausal Breast Cancer,” *J. Clin. Oncol.*, vol. 32, no. 28, p. 3111, Oct. 2014, doi: 10.1200/JCO.2014.56.1068.
- [139] E. Petracchi *et al.*, “Risk factor modification and projections of absolute breast cancer risk,” *J. Natl. Cancer Inst.*, vol. 103, no. 13, pp. 1037–1048, Jul. 2011, doi: 10.1093/JNCI/DJR172.
- [140] S. Franceschi *et al.*, “Intake of macronutrients and risk of breast cancer,” *Lancet (London, England)*, vol. 347, no. 9012, pp. 1351–1356, May 1996, doi: 10.1016/S0140-6736(96)91008-9.
- [141] E. Claus, N. Risch, and W. Thompson, “The calculation of breast cancer risk for women with a first degree family history of ovarian cancer,” *Breast Cancer Res. Treat.*, vol. 28, no. 2, pp. 115–120, Jan. 1993, doi: 10.1007/BF00666424.
- [142] E. Amir, O. C. Freedman, B. Seruga, and D. G. Evans, “Assessing Women at High Risk of Breast Cancer: A Review of Risk Assessment Models,” *JNCI J. Natl. Cancer Inst.*, vol. 102, no. 10, pp. 680–691, May 2010, doi: 10.1093/JNCI/DJQ088.
- [143] D. Shattuck-Eidens *et al.*, “BRCA1 sequence analysis in women at high risk for susceptibility mutations: Risk factor analysis and implications for genetic testing,” *J. Am. Med. Assoc.*, vol. 278, no. 15, pp. 1242–1250, Oct. 1997, doi: 10.1001/JAMA.278.15.1242.
- [144] T. S. Frank *et al.*, “Sequence analysis of BRCA1 and BRCA2: Correlation of mutations with family history and ovarian cancer risk,” *J. Clin. Oncol.*, vol. 16, no. 7, pp. 2417–2425, 1998, doi: 10.1200/JCO.1998.16.7.2417.
- [145] D. D. Couch *et al.*, “BRCA1 mutations in women attending clinics that evaluate the risk of breast cancer,” *N. Engl. J. Med.*, vol. 336, no. 20, pp. 1409–1415, May 1997, doi: 10.1056/NEJM199705153362002.
- [146] N. M. Lindor *et al.*, “Predicting BRCA1 and BRCA2 gene mutation carriers: comparison of PENN II model to previous study,” *Fam. Cancer*, vol. 9, no. 4, p. 495, Dec. 2010, doi: 10.1007/S10689-010-9348-3.
- [147] D. Evans *et al.*, “A new scoring system for the chances of identifying a BRCA1/2 mutation outperforms existing models including BRCAPRO,” *J. Med. Genet.*, vol. 41, no. 6, pp. 474–480, 2004, doi: 10.1136/JMG.2003.017996.
- [148] D. Evans *et al.*, “Addition of pathology and biomarker information significantly improves the performance of the Manchester scoring system for BRCA1 and BRCA2 testing,” *J. Med. Genet.*, vol. 46, no. 12, pp. 811–817, Dec. 2009, doi: 10.1136/JMG.2009.067850.
- [149] C. Bellcross, A. Lemke, L. Pape, L. Tess, and L. Meisner, “Evaluation of a breast/ovarian cancer genetics referral screening tool in a mammography population,” *Genet. Med.*, vol. 11, no. 11, pp. 783–789, Nov. 2009, doi: 10.1097/GIM.0B013E3181B9B04A.

- [150] C. B. F. Berrino, “COS, case-only-study on breast cancer before the age of 40. Announcement of a new study on gene-environment interaction in breast cancer. — Italian Ministry of Health,” *Fondazione IRCCS Istituto Nazionale dei Tumori*. <https://moh-it.pure.elsevier.com/en/publications/cos-case-only-study-on-breast-cancer-before-the-age-of-40-announc> (accessed Sep. 30, 2021).
- [151] D. Berry, G. Parmigiani, J. Sanchez, J. Schildkraut, and E. Winer, “Probability of carrying a mutation of breast-ovarian cancer gene BRCA1 based on family history,” *J. Natl. Cancer Inst.*, vol. 89, no. 3, pp. 227–238, Feb. 1997, doi: 10.1093/JNCI/89.3.227.
- [152] G. Parmigiani, D. Berry, and O. Aguilar, “Determining carrier probabilities for breast cancer-susceptibility genes BRCA1 and BRCA2,” *Am. J. Hum. Genet.*, vol. 62, no. 1, pp. 145–158, 1998, doi: 10.1086/301670.
- [153] E. Mazzola, A. Blackford, G. Parmigiani, and S. Biswas, “Recent Enhancements to the Genetic Risk Prediction Model BRCAPRO,” *Cancer Inform.*, vol. 14, no. Suppl 2, pp. 147–157, May 2015, doi: 10.4137/CIN.S17292.
- [154] A. Antoniou *et al.*, “A comprehensive model for familial breast cancer incorporating BRCA1, BRCA2 and other genes,” *Br. J. Cancer*, vol. 86, no. 1, pp. 76–83, 2002, doi: 10.1038/SJ.BJC.6600008.
- [155] A. Antoniou, P. Pharoah, P. Smith, and D. Easton, “The BOADICEA model of genetic susceptibility to breast and ovarian cancer,” *Br. J. Cancer*, vol. 91, no. 8, pp. 1580–1590, Oct. 2004, doi: 10.1038/SJ.BJC.6602175.
- [156] A. C. Antoniou *et al.*, “The BOADICEA model of genetic susceptibility to breast and ovarian cancers: updates and extensions,” *Br. J. Cancer 2008 988*, vol. 98, no. 8, pp. 1457–1466, Mar. 2008, doi: 10.1038/sj.bjc.6604305.
- [157] M. A. Jonker, C. E. Jacobi, W. E. Hoogendoorn, N. J. D. Nagelkerke, G. H. De Bock, and J. C. Van Houwelingen, “Modeling Familial Clustered Breast Cancer Using Published Data,” 2003.
- [158] J. Tyrer, S. Duffy, and J. Cuzick, “A breast cancer prediction model incorporating familial and personal risk factors,” *Stat. Med.*, vol. 23, no. 7, pp. 1111–1130, Apr. 2004, doi: 10.1002/SIM.1668.
- [159] A. Quante, A. Whittemore, T. Shriver, K. Strauch, and M. Terry, “Breast cancer risk assessment across the risk continuum: genetic and nongenetic risk factors contributing to differential model performance,” *Breast Cancer Res.*, vol. 14, no. 6, Nov. 2012, doi: 10.1186/BCR3352.
- [160] A. F. Agarap, “On Breast Cancer Detection: An Application of Machine Learning Algorithms on the Wisconsin Diagnostic Dataset,” *ACM Int. Conf. Proceeding Ser.*, pp. 5–9, Nov. 2017, doi: 10.1145/3184066.3184080.
- [161] S. Huang, C. A. I. Nianguang, P. Penzuti Pacheco, S. Narandes, Y. Wang, and X. U. Wayne, “Applications of support vector machine (SVM) learning in cancer genomics,” *Cancer Genomics and Proteomics*, vol. 15, no. 1. International Institute of Anticancer Research, pp. 41–51, Jan. 01, 2018, doi: 10.21873/cgp.20063.
- [162] “UCI Machine Learning Repository: Breast Cancer Wisconsin (Diagnostic) Data Set.” [https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+\(Diagnostic\)](https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+(Diagnostic)) (accessed Jun. 28, 2022).
- [163] B. Hemmati and H. Jafarkarimi, “Determining the Malignancy of Breast Cancer Using a Combined Neural Network-Support Vector Machine Scheme,” *J. Soft Comput. Decis. Support*

- Syst.*, vol. 5, no. 2, pp. 1–8, Mar. 2018, Accessed: Jun. 04, 2021. [Online]. Available: <http://www.jscdss.com/index.php/files/article/view/161>.
- [164] H. Wang, B. Zheng, S. W. Yoon, and H. S. Ko, “A support vector machine-based ensemble algorithm for breast cancer diagnosis,” *Eur. J. Oper. Res.*, vol. 267, no. 2, pp. 687–699, Jun. 2018, doi: 10.1016/j.ejor.2017.12.001.
- [165] A. F. M. Agarap, “On breast cancer detection: An application of machine learning algorithms on the Wisconsin diagnostic dataset,” in *ACM International Conference Proceeding Series*, Feb. 2018, pp. 5–9, doi: 10.1145/3184066.3184080.
- [166] K. Kawamoto, C. A. Houlihan, E. A. Balas, and D. F. Lobach, “Improving clinical practice using clinical decision support systems: A systematic review of trials to identify features critical to success,” *British Medical Journal*, vol. 330, no. 7494, BMJ Publishing Group, pp. 765–768, Apr. 02, 2005, doi: 10.1136/bmj.38398.500764.8f.
- [167] K. Kawamoto, D. F. Lobach, H. F. Willard, and G. S. Ginsburg, “A national clinical decision support infrastructure to enable the widespread and consistent practice of genomic and personalized medicine,” *BMC Med. Inform. Decis. Mak.*, vol. 9, no. 1, pp. 1–14, Mar. 2009, doi: 10.1186/1472-6947-9-17.
- [168] P. J. Lisboa and A. F. G. Taktak, “The use of artificial neural networks in decision support in cancer: A systematic review,” *Neural Networks*, vol. 19, no. 4, pp. 408–415, May 2006, doi: 10.1016/j.neunet.2005.10.007.
- [169] S. M. Carter, W. Rogers, K. T. Win, H. Frazer, B. Richards, and N. Houssami, “The ethical, legal and social implications of using artificial intelligence systems in breast cancer care,” *Breast*, vol. 49, pp. 25–32, Feb. 2020, doi: 10.1016/j.breast.2019.10.001.
- [170] F. Xu *et al.*, “Effect of an Artificial Intelligence Clinical Decision Support System on Treatment Decisions for Complex Breast Cancer,” *JCO Clin. Cancer Informatics*, no. 4, pp. 824–838, Oct. 2020, doi: 10.1200/cci.20.00018.
- [171] M. Ngadi, A. Amine, and B. Nassih, “A robust approach for mammographic image classification using NSVC algorithm,” in *ACM International Conference Proceeding Series*, Nov. 2016, pp. 44–49, doi: 10.1145/3038884.3038892.
- [172] N. I. R. Yassin, S. Omran, E. M. F. El Houby, and H. Allam, “Machine learning techniques for breast cancer computer aided diagnosis using different image modalities: A systematic review,” *Comput. Methods Programs Biomed.*, vol. 156, pp. 25–45, 2018, doi: 10.1016/j.cmpb.2017.12.012.
- [173] G. F. Stark, G. R. Hart, B. J. Nartowt, and J. Deng, “Predicting breast cancer risk using personal health data and machine learning models,” *PLoS One*, vol. 14, no. 12, p. e0226765, Dec. 2019, doi: 10.1371/JOURNAL.PONE.0226765.
- [174] M. Nilashi, O. Ibrahim, H. Ahmadi, and L. Shahmoradi, “A knowledge-based system for breast cancer classification using fuzzy logic method,” *Telemat. Informatics*, vol. 34, no. 4, pp. 133–144, Jul. 2017, doi: 10.1016/j.tele.2017.01.007.
- [175] E. A. Bayrak, P. Kirci, and T. Ensari, “Comparison of machine learning methods for breast cancer diagnosis,” Apr. 2019, doi: 10.1109/EBBT.2019.8741990.
- [176] C. Banerjee, S. Paul, and M. Ghoshal, “A comparative study of different ensemble learning techniques using Wisconsin breast cancer dataset,” Nov. 2018, doi: 10.1109/ICCECE.2017.8526215.

- [177] H. Asri, H. Mousannif, H. Al Moatassime, and T. Noel, "ScienceDirect The 6th International Symposium on Frontiers in Ambient and Mobile Systems Using Machine Learning Algorithms for Breast Cancer Risk Prediction and Diagnosis," *Procedia Comput. Sci.*, vol. 83, pp. 1064–1069, 2016, doi: 10.1016/j.procs.2016.04.224.
- [178] D. A. Omondiagbe, S. Veeramani, and A. S. Sidhu, "Machine Learning Classification Techniques for Breast Cancer Diagnosis," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 495, no. 1, p. 012033, Apr. 2019, doi: 10.1088/1757-899X/495/1/012033.
- [179] A. Osareh and B. Shadgar, "Machine Learning Techniques to Diagnose Breast Cancer," *2010 5th Int. Symp. Heal. Informatics Bioinforma.*, 2020.
- [180] M. M. Rahman, Y. Ghasemi, E. Suley, Y. Zhou, S. Wang, and J. Rogers, "Machine Learning Based Computer Aided Diagnosis of Breast Cancer Utilizing Anthropometric and Clinical Features," *IRBM*, vol. 42, no. 4, pp. 215–226, Aug. 2021, doi: 10.1016/J.IRBM.2020.05.005.
- [181] M. F. Aslan, Y. Çelik, K. Sabancı, and A. Durdu, "Breast cancer diagnosis by different machine learning methods using blood analysis data," *Orig. Res. Pap. Int. J. Intell. Syst. Appl. Eng. IJISAE*, vol. 6, no. 4, pp. 289–293, 2018, doi: 10.1039/b000000x.
- [182] "UCI Machine Learning Repository." <https://archive.ics.uci.edu/ml/index.php> (accessed Jun. 28, 2022).
- [183] B. Séroussi *et al.*, "Reconciliation of multiple guidelines for decision support: a case study on the multidisciplinary management of breast cancer within the DESIREE project," *AMIA ... Annu. Symp. proceedings. AMIA Symp.*, vol. 2017, pp. 1527–1536, 2017, Accessed: Jun. 11, 2021. [Online]. Available: <http://www.g-i-n.net>.
- [184] J. Bouaud *et al.*, "Implementation of an ontological reasoning to support the guideline-based management of primary breast cancer patients in the DESIREE project," *Artif. Intell. Med.*, vol. 108, p. 101922, Aug. 2020, doi: 10.1016/j.artmed.2020.101922.
- [185] A. M. Alaa *et al.*, "ConfidentCare: A Clinical Decision Support System for Personalized Breast Cancer Screening," *IEEE Trans. Multimed.*, vol. 18, no. 10, pp. 1942–1955, Feb. 2016, doi: 10.1109/TMM.2016.2589160.
- [186] M. Alamelumangai and B. Sathiyabhama, "Personalized Care: A Clinical Decision Support System for Breast Cancer Screening Using Clustering and Classification," *SSRN Electron. J.*, Nov. 2017, doi: 10.2139/SSRN.3134277.
- [187] I.-W. Pan, K. C. Oeffinger, and Y.-C. T. Shih, "Out-of-pocket cost of screening with breast MRI for women at high risk for breast cancer.," *J. Clin. Oncol.*, vol. 38, no. 15_suppl, pp. 2017–2017, May 2020, doi: 10.1200/jco.2020.38.15_suppl.2017.
- [188] A. Casteel and N. L. Bridier, "Describing populations and samples in Doctoral student research," *Int. J. Dr. Stud.*, vol. 16, pp. 339–362, doi: 10.28945/4766.
- [189] E. Y. Kim *et al.*, "Serum concentration of thyroid hormones in abnormal and euthyroid ranges and breast cancer risk: A cohort study," *Int. J. Cancer*, vol. 145, no. 12, pp. 3257–3266, Dec. 2019, doi: 10.1002/IJC.32283.
- [190] M. Simon, M. Tang, and L. Bernstein, "Do thyroid disorders increase the risk of breast cancer?," *AACR*, 2002, Accessed: Oct. 21, 2021. [Online]. Available: <https://cebp.aacrjournals.org/content/11/12/1574.short>.
- [191] M. Bahl, J. A. Baker, E. N. Kinsey, and S. V. Ghate, "Architectural Distortion on Mammography:

- Correlation With Pathologic Outcomes and Predictors of Malignancy,” <http://dx.doi.org/10.2214/AJR.15.14628>, vol. 205, no. 6, pp. 1339–1345, Nov. 2015, doi: 10.2214/AJR.15.14628.
- [192] S. W. Duffy *et al.*, “Mammography screening reduces rates of advanced and fatal breast cancers: Results in 549,091 women,” *Cancer*, vol. 126, no. 13, pp. 2971–2979, Jul. 2020, doi: 10.1002/CNCR.32859.
- [193] S. Waite *et al.*, “Analysis of perceptual expertise in radiology – Current knowledge and a new perspective,” *Front. Hum. Neurosci.*, vol. 13, p. 213, Feb. 2019, doi: 10.3389/FNHUM.2019.00213/BIBTEX.
- [194] J. Krammer *et al.*, “Breast cancer detection and tumor characteristics in BRCA1 and BRCA2 mutation carriers,” *Breast Cancer Res. Treat.*, vol. 163, no. 3, pp. 565–571, Jun. 2017, doi: 10.1007/s10549-017-4198-4.
- [195] L. Domingo *et al.*, “Cross-national comparison of screening mammography accuracy measures in U.S., Norway, and Spain,” *Eur. Radiol.*, vol. 26, no. 8, pp. 2520–2528, Aug. 2016, doi: 10.1007/s00330-015-4074-8.
- [196] C. D. Lehman, R. D. Wellman, D. S. M. Buist, K. Kerlikowske, A. N. A. Tosteson, and D. L. Miglioretti, “Diagnostic accuracy of digital screening mammography with and without computer-aided detection,” *JAMA Intern. Med.*, vol. 175, no. 11, pp. 1828–1837, Nov. 2015, doi: 10.1001/jamainternmed.2015.5231.
- [197] L. Esserman *et al.*, “Improving the accuracy of mammography: Volume and outcome relationships,” *J. Natl. Cancer Inst.*, vol. 94, no. 5, pp. 369–375, Mar. 2002, doi: 10.1093/jnci/94.5.369.
- [198] P. J. García-Laencina, J. L. Sancho-Gómez, and A. R. Figueiras-Vidal, “Pattern classification with missing data: A review,” *Neural Comput. Appl.*, vol. 19, no. 2, pp. 263–282, Sep. 2010, doi: 10.1007/S00521-009-0295-6/TABLES/4.
- [199] Π. Μπούντρης, “Ανάπτυξη Αλγορίθμων Υπολογιστικής Νοημοσύνης και Ευφών Συστημάτων Εκτίμησης Κινδύνου και Υποστήριξης Κλινικής Απόφασης για την Έγκαιρη Διάγνωση και την Εξατομικευμένη Διαχείριση Γυναικών με Ενδοεπιθηλιακές Αλλοιώσεις Τραχήλου Μήτρας,” National Technical University of Athens, 2017.
- [200] G. Wei, J. Zhao, Y. Feng, A. He, and J. Yu, “A novel hybrid feature selection method based on dynamic feature importance,” *Appl. Soft Comput.*, vol. 93, p. 106337, Aug. 2020, doi: 10.1016/J.ASOC.2020.106337.
- [201] J. Cai, J. Luo, S. Wang, and S. Yang, “Feature selection in machine learning: A new perspective,” *Neurocomputing*, vol. 300, pp. 70–79, Jul. 2018, doi: 10.1016/J.NEUCOM.2017.11.077.
- [202] A. Goltsev and V. Gritsenko, “Investigation of efficient features for image recognition by neural networks,” *Neural Networks*, vol. 28, pp. 15–23, Apr. 2012, doi: 10.1016/J.NEUNET.2011.12.002.
- [203] N. Vasconcelos, “Feature selection by maximum marginal diversity: Optimality and implications for visual recognition,” *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 1, 2003, doi: 10.1109/CVPR.2003.1211430.
- [204] E. Rashedi, H. Nezamabadi-Pour, and S. Saryazdi, “A simultaneous feature adaptation and feature selection method for content-based image retrieval systems,” *Knowledge-Based Syst.*, vol. 39, pp. 85–94, Feb. 2013, doi: 10.1016/J.KNOSYS.2012.10.011.

- [205] Y. F. Gao, B. Q. Li, Y. D. Cai, K. Y. Feng, Z. D. Li, and Y. Jiang, "Prediction of active sites of enzymes by maximum relevance minimum redundancy (mRMR) feature selection," *Mol. Biosyst.*, vol. 9, no. 1, pp. 61–69, 2013, doi: 10.1039/C2MB25327E.
- [206] Q. Song, J. Ni, and G. Wang, "A fast clustering-based feature subset selection algorithm for high-dimensional data," *IEEE Trans. Knowl. Data Eng.*, vol. 25, no. 1, pp. 1–14, 2013, doi: 10.1109/TKDE.2011.181.
- [207] T. W. Rauber, F. De Assis Boldt, and F. M. Varejão, "Heterogeneous feature models and feature selection applied to bearing fault diagnosis," *IEEE Trans. Ind. Electron.*, vol. 62, no. 1, pp. 637–646, Jan. 2015, doi: 10.1109/TIE.2014.2327589.
- [208] B. Remeseiro and V. Bolon-Canedo, "A review of feature selection methods in medical applications," *Comput. Biol. Med.*, vol. 112, Sep. 2019, doi: 10.1016/J.COMPBIOMED.2019.103375.
- [209] H. Li, C. J. Li, X. J. Wu, and J. Sun, "Statistics-based wrapper for feature selection: An implementation on financial distress identification with support vector machine," *Appl. Soft Comput.*, vol. 19, pp. 57–67, Jun. 2014, doi: 10.1016/J.ASOC.2014.01.018.
- [210] T. Khoshgoftaar, D. Dittman, R. Wald, and A. Fazelpour, "First order statistics based feature selection: A diverse and powerful family of feature selection techniques," *Proc. - 2012 11th Int. Conf. Mach. Learn. Appl. ICMLA 2012*, vol. 2, pp. 151–157, 2012, doi: 10.1109/ICMLA.2012.192.
- [211] L. Shen and L. Bai, "Information theory for gabor feature selection for face recognition," *EURASIP J. Appl. Signal Processing*, vol. 2006, pp. 1–11, 2006, doi: 10.1155/ASP/2006/30274.
- [212] N. Almugren and H. Alshamlan, "A survey on hybrid feature selection methods in microarray gene expression data for cancer classification," *IEEE Access*, vol. 7, pp. 78533–78548, 2019, doi: 10.1109/ACCESS.2019.2922987.
- [213] R. Kohavi and G. H. John, "Wrappers for feature subset selection," *Artif. Intell.*, vol. 97, no. 1–2, pp. 273–324, Dec. 1997, doi: 10.1016/S0004-3702(97)00043-X.
- [214] J. R. Zhang, J. Zhang, T. M. Lok, and M. R. Lyu, "A hybrid particle swarm optimization–back-propagation algorithm for feedforward neural network training," *Appl. Math. Comput.*, vol. 185, no. 2, pp. 1026–1037, Feb. 2007, doi: 10.1016/J.AMC.2006.07.025.
- [215] D. S. Huang and J. X. Du, "A constructive hybrid structure optimization methodology for radial basis probabilistic neural networks," *IEEE Trans. neural networks*, vol. 19, no. 12, pp. 2099–2115, 2008, doi: 10.1109/TNN.2008.2004370.
- [216] A. Rodriguez and A. Laio, "Clustering by fast search and find of density peaks," *Science (80-.)*, vol. 344, no. 6191, pp. 1492–1496, Jun. 2014, doi: 10.1126/SCIENCE.1242072/SUPPL_FILE/RODRIGUEZ.SM.PDF.
- [217] F. Gorunescu, "Data mining: Concepts, models and techniques," *Intell. Syst. Ref. Libr.*, vol. 12, 2011, doi: 10.1007/978-3-642-19721-5.
- [218] M. A. Hall, "Correlation-based feature selection of discrete and numeric class machine learning," 2000, Accessed: Jan. 24, 2022. [Online]. Available: <https://researchcommons.waikato.ac.nz/handle/10289/1024>.
- [219] I. Kononenko, "Estimating attributes: Analysis and extensions of RELIEF," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 784 LNCS, pp.

- 171–182, 1994, doi: 10.1007/3-540-57868-4_57.
- [220] K. Kira and L. A. Rendell, “A Practical Approach to Feature Selection,” *Mach. Learn. Proc. 1992*, pp. 249–256, Jan. 1992, doi: 10.1016/B978-1-55860-247-2.50037-1.
- [221] L. Song, A. Smola, A. Gretton, K. M. Borgwardt, and J. Bedo, “Supervised feature selection via dependence estimation,” *ACM Int. Conf. Proceeding Ser.*, vol. 227, pp. 823–830, 2007, doi: 10.1145/1273496.1273600.
- [222] H. Liu and L. Yu, “Toward integrating feature selection algorithms for classification and clustering,” *IEEE Trans. Knowl. Data Eng.*, vol. 17, no. 4, pp. 491–502, Apr. 2005, doi: 10.1109/TKDE.2005.66.
- [223] L. Yu and H. Liu, “Efficient Feature Selection via Analysis of Relevance and Redundancy,” *J. Mach. Learn. Res.*, vol. 5, pp. 1205–1224, 2004.
- [224] H. Peng, F. Long, and C. Ding, “Feature selection based on mutual information: Criteria of Max-Dependency, Max-Relevance, and Min-Redundancy,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 8, pp. 1226–1238, Aug. 2005, doi: 10.1109/TPAMI.2005.159.
- [225] R. Battiti, “Using Mutual Information for Selecting Features in Supervised Neural Net Learning,” *IEEE Trans. Neural Networks*, vol. 5, no. 4, pp. 537–550, 1994, doi: 10.1109/72.298224.
- [226] N. Kwak and C. H. Choi, “Improved mutual information feature selector for neural networks in supervised learning,” *Proc. Int. Jt. Conf. Neural Networks*, vol. 2, pp. 1313–1318, 1999, doi: 10.1109/IJCNN.1999.831152.
- [227] J. Novovičová, P. Somol, M. Haindl, and P. Pudil, “Conditional mutual information based feature selection for classification task,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 4756 LNCS, pp. 417–426, 2007, doi: 10.1007/978-3-540-76725-1_44.
- [228] Y. Zhang and Z. Zhang, “Feature subset selection with cumulate conditional mutual information minimization,” *Expert Syst. Appl.*, vol. 39, no. 5, pp. 6078–6088, Apr. 2012, doi: 10.1016/J.ESWA.2011.12.003.
- [229] H. H. Yang and J. Moody, “Data Visualization and Feature Selection: New Algorithms for Nongaussian Data,” *Adv. Neural Inf. Process. Syst.*, vol. 12, 1999.
- [230] F. Fleuret, “Fast Binary Feature Selection with Conditional Mutual Information,” *J. Mach. Learn. Res.*, vol. 5, pp. 1531–1555, 2004, Accessed: Jan. 27, 2022. [Online]. Available: <http://diwww.epfl.ch/>.
- [231] Y. Zhang, S. Li, T. Wang, and Z. Zhang, “Divergence-based feature selection for separate classes,” *Neurocomputing*, vol. 101, pp. 32–42, Feb. 2013, doi: 10.1016/J.NEUCOM.2012.06.036.
- [232] C. Lee and G. G. Lee, “Information gain and divergence-based feature selection for machine learning-based text categorization,” *Inf. Process. Manag.*, vol. 42, no. 1, pp. 155–165, Jan. 2006, doi: 10.1016/J.IPM.2004.08.006.
- [233] I. S. Dhillon, S. Mallela, and R. Kumar, “A Divisive Information-Theoretic Feature Clustering Algorithm for Text Classification,” *J. Mach. Learn. Res.*, vol. 3, pp. 1265–1287, 2003, Accessed: Jan. 28, 2022. [Online]. Available: www.dmoz.org.
- [234] H. Liu, X. Wu, and S. Zhang, “Feature selection using hierarchical feature clustering,” *Int. Conf. Inf. Knowl. Manag. Proc.*, pp. 979–984, 2011, doi: 10.1145/2063576.2063716.

- [235] D. M. Witten and R. Tibshirani, "A framework for feature selection in clustering," *J. Am. Stat. Assoc.*, vol. 105, no. 490, p. 713, Jun. 2010, doi: 10.1198/JASA.2010.TM09415.
- [236] D. Ienco and R. Meo, "Exploration and reduction of the feature space by hierarchical clustering," *Proceedings*, vol. 2, pp. 577–587, 2008, doi: 10.1137/1.9781611972788.53.
- [237] X. Zhao, W. Deng, and Y. Shi, "Feature Selection with Attributes Clustering by Maximal Information Coefficient," *Procedia Comput. Sci.*, vol. 17, pp. 70–79, Jan. 2013, doi: 10.1016/J.PROCS.2013.05.011.
- [238] W. H. Au, K. C. C. Chan, A. K. C. Wong, and Y. Wang, "Attribute clustering for grouping, selection, and classification of gene expression data," *IEEE/ACM Trans. Comput. Biol. Bioinforma.*, vol. 2, no. 2, pp. 83–101, Apr. 2005, doi: 10.1109/TCBB.2005.17.
- [239] Q. Liu, J. Zhang, J. Xiao, H. Zhu, and Q. Zhao, "A Supervised Feature Selection Algorithm through Minimum Spanning Tree Clustering," *Proc. - Int. Conf. Tools with Artif. Intell. ICTAI*, vol. 2014-December, pp. 264–271, Dec. 2014, doi: 10.1109/ICTAI.2014.47.
- [240] J. Martínez Sotoca and F. Pla, "Supervised feature selection by clustering using conditional mutual information-based distances," *Pattern Recognit.*, vol. 43, no. 6, pp. 2068–2081, Jun. 2010, doi: 10.1016/J.PATCOG.2009.12.013.
- [241] Q. Zou, J. Zeng, L. Cao, and R. Ji, "A novel features ranking metric with application to scalable visual and bioinformatics data classification," *Neurocomputing*, vol. 173, pp. 346–354, Jan. 2016, doi: 10.1016/J.NEUCOM.2014.12.123.
- [242] I. Guyon, J. Weston, S. Barnhill, and V. Vapnik, "Gene Selection for Cancer Classification using Support Vector Machines," *Mach. Learn. 2002 461*, vol. 46, no. 1, pp. 389–422, 2002, doi: 10.1023/A:1012487302797.
- [243] J. Zhong, J. Wang, W. Peng, Z. Zhang, and M. Li, "A feature selection method for prediction essential protein," *Tsinghua Sci. Technol.*, vol. 20, no. 5, pp. 491–499, Oct. 2015, doi: 10.1109/TST.2015.7297748.
- [244] K. Michalak and H. Kwaśnicka, "Correlation-based feature selection strategy in classification problems," *Int. J. Appl. Math. Comput. Sci.*, vol. Vol. 16, no. no 4, pp. 503–511, 2006.
- [245] K. Javed, H. A. Babri, and M. Saeed, "Feature selection based on class-dependent densities for high-dimensional binary data," *IEEE Trans. Knowl. Data Eng.*, vol. 24, no. 3, pp. 465–477, 2012, doi: 10.1109/TKDE.2010.263.
- [246] A. El Akadi, A. Amine, A. El Ouardighi, and D. Aboutajdine, "A two-stage gene selection scheme utilizing MRMR filter and GA wrapper," *Knowl. Inf. Syst.*, vol. 26, no. 3, pp. 487–500, 2011, doi: 10.1007/S10115-010-0288-X.
- [247] J. M. Cadenas, M. C. Garrido, and R. Martínez, "Feature subset selection Filter-Wrapper based on low quality data," *Expert Syst. Appl.*, vol. 40, no. 16, pp. 6241–6252, 2013, doi: 10.1016/J.ESWA.2013.05.051.
- [248] E. Odhiambo Omuya, G. Onyango Okeyo, and M. Waema Kimwele, "Feature Selection for Classification using Principal Component Analysis and Information Gain," *Expert Syst. Appl.*, vol. 174, p. 114765, Jul. 2021, doi: 10.1016/J.ESWA.2021.114765.
- [249] P. Bugata and P. Drotar, "On some aspects of minimum redundancy maximum relevance feature selection," vol. 63, no. 112103, p. 15, 2020, doi: 10.1007/s11432-019-2633-y.

- [250] S. Theodoridis and K. Koutroumbas, *Pattern recognition*, 4th ed. Academic Press, 2009.
- [251] H. Zhang, “The Optimality of Naive Bayes,” *Proc. Seventeenth Int. Florida Artif. Intell. Res. Soc. Conf. FLAIRS 2004*, vol. 1, no. 2, pp. 1–6, 2004, doi: 10.1016/j.patrec.2005.12.001.
- [252] D. Lowd and P. Domingos, “Naive Bayes Models for Probability Estimation,” 2005, Accessed: Nov. 07, 2021. [Online]. Available: <http://www.cs.washington.edu/ai/nbe>.
- [253] M. Fernández-Delgado, E. Cernadas, S. Barro, D. Amorim, and A. Fernández-Delgado, “Do we Need Hundreds of Classifiers to Solve Real World Classification Problems?,” *J. Mach. Learn. Res.*, vol. 15, pp. 3133–3181, 2014, Accessed: Nov. 08, 2021. [Online]. Available: <http://www.mathworks.es/products/neural-network>.
- [254] P. Bountris, M. Haritou, A. Pouliakis, P. Karakitsos, and D. Koutsouris, “A decision support system based on an ensemble of random forests for improving the management of women with abnormal findings at cervical cancer screening,” *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS*, vol. 2015–November, pp. 8151–8156, Nov. 2015, doi: 10.1109/EMBC.2015.7320286.
- [255] L. Breiman, “Random Forests,” *Mach. Learn. 2001 451*, vol. 45, no. 1, pp. 5–32, Oct. 2001, doi: 10.1023/A:1010933404324.
- [256] “Machine Learning Random Forest Algorithm - Javatpoint.” <https://www.javatpoint.com/machine-learning-random-forest-algorithm> (accessed Jun. 28, 2022).
- [257] M. Fernández-Delgado, E. Cernadas, S. Barro, D. Amorim, and A. Fernández-Delgado, “Do we Need Hundreds of Classifiers to Solve Real World Classification Problems?,” *J. Mach. Learn. Res.*, vol. 15, pp. 3133–3181, 2014, Accessed: Dec. 20, 2021. [Online]. Available: <http://www.mathworks.es/products/neural-network>.
- [258] B. M. Geller *et al.*, “Use of the American College of Radiology BI-RADS to Report on the Mammographic Evaluation of Women with Signs and Symptoms of Breast Disease,” <https://doi.org/10.1148/radiol.2222010620>, vol. 222, no. 2, pp. 536–542, Feb. 2002, doi: 10.1148/RADIOL.2222010620.
- [259] S. P. Poplack, A. N. Tosteson, M. R. Grove, W. A. Wells, and P. A. Carney, “Mammography in 53,803 women from the New Hampshire mammography network,” *Radiology*, vol. 217, no. 3, pp. 832–840, 2000, doi: 10.1148/RADIOLOGY.217.3.R00DC33832.
- [260] R. D. Rosenberg *et al.*, “Effect of variations in operational definitions on performance estimates for screening Mammography,” *Acad. Radiol.*, vol. 7, no. 12, pp. 1058–1068, Dec. 2000, doi: 10.1016/S1076-6332(00)80057-4.
- [261] C. Yi *et al.*, “The added value of an artificial intelligence system in assisting radiologists on indeterminate BI-RADS 0 mammograms,” *Eur. Radiol.*, vol. 32, no. 3, pp. 1528–1537, Mar. 2022, doi: 10.1007/S00330-021-08275-0/TABLES/5.
- [262] E. G. Klompenhouwer *et al.*, “Arbitration of discrepant BI-RADS 0 recalls by a third reader at screening mammography lowers recall rate but not the cancer detection rate and sensitivity at blinded and non-blinded double reading,” *Breast*, vol. 24, no. 5, pp. 601–607, 2015, doi: 10.1016/J.BREAST.2015.06.004.
- [263] P. A. Zanello *et al.*, “Breast ultrasound diagnostic performance and outcomes for mass lesions using Breast Imaging Reporting and Data System category 0 mammogram,” *Clinics (Sao Paulo)*, vol. 66, no. 3, pp. 443–448, 2011, doi: 10.1590/S1807-59322011000300014.
- [264] W. K. Lee, J. Chung, E. S. Cha, J. E. Lee, and J. H. Kim, “Digital breast tomosynthesis and breast

- ultrasound: Additional roles in dense breasts with category 0 at conventional digital mammography,” *Eur. J. Radiol.*, vol. 85, no. 1, pp. 291–296, Jan. 2016, doi: 10.1016/J.EJRAD.2015.09.026.
- [265] W. Lotter *et al.*, “Robust breast cancer detection in mammography and digital breast tomosynthesis using an annotation-efficient deep learning approach,” *Nat. Med.* 2021 272, vol. 27, no. 2, pp. 244–249, Jan. 2021, doi: 10.1038/s41591-020-01174-9.
- [266] A. Yala, T. Schuster, R. Miles, R. Barzilay, and C. Lehman, “A deep learning model to triage screening mammograms: A simulation study,” *Radiology*, vol. 293, no. 1, pp. 38–46, Aug. 2019, doi: 10.1148/RADIOL.2019182908/ASSET/IMAGES/LARGE/RADIOL.2019182908.TBL4.JPEG.
- [267] S. M. McKinney *et al.*, “International evaluation of an AI system for breast cancer screening,” *Nat.* 2020 5777788, vol. 577, no. 7788, pp. 89–94, Jan. 2020, doi: 10.1038/s41586-019-1799-6.
- [268] J. Xu, B. S. Glicksberg, C. Su, P. Walker, J. Bian, and F. Wang, “Federated Learning for Healthcare Informatics,” *J. Healthc. Informatics Res.*, vol. 5, no. 1, pp. 1–19, Mar. 2021, doi: 10.1007/S41666-020-00082-4/TABLES/2.
- [269] R. S. Antunes, C. André, D. A. Costa, A. Küderle, I. A. Yari, and B. Eskofier, “Federated Learning for Healthcare: Systematic Review and Architecture Proposal,” *ACM Trans. Intell. Syst. Technol.*, Mar. 2021, doi: 10.1145/3501813.
- [270] V. C. Pezoulas *et al.*, “Addressing the clinical unmet needs in primary Sjögren’s Syndrome through the sharing, harmonization and federated analysis of 21 European cohorts,” *Comput. Struct. Biotechnol. J.*, vol. 20, pp. 471–484, Jan. 2022, doi: 10.1016/J.CSBJ.2022.01.002.
- [271] V. C. Pezoulas, T. P. Exarchos, K. D. Kourou, A. G. Tzioufas, S. De Vita, and D. I. Fotiadis, “Utilizing incremental learning for the prediction of disease outcomes across distributed clinical data: a framework and a case study,” *IFMBE Proc.*, vol. 76, pp. 823–831, 2020, doi: 10.1007/978-3-030-31635-8_98.
- [272] A. Qayyum, K. Ahmad, M. Ahsan, ... A. A.-F. preprint arXiv, and undefined 2021, “Collaborative federated learning for healthcare: Multi-modal covid-19 diagnosis at the edge,” *arxiv.org*, Accessed: Mar. 17, 2022. [Online]. Available: <https://arxiv.org/abs/2101.07511>.
- [273] V. C. Pezoulas *et al.*, “Overcoming the barriers that obscure the interlinking and analysis of clinical data through harmonization and incremental learning,” *IEEE Open J. Eng. Med. Biol.*, vol. 1, pp. 83–90, 2020, doi: 10.1109/OJEMB.2020.2981258.
- [274] S. M. Lundberg and S. I. Lee, “A Unified Approach to Interpreting Model Predictions,” *Adv. Neural Inf. Process. Syst.*, vol. 2017-December, pp. 4766–4775, May 2017, doi: 10.48550/arxiv.1705.07874.
- [275] A. G. Lalkhen and A. McCluskey, “Clinical tests: sensitivity and specificity,” *Contin. Educ. Anaesth. Crit. Care Pain*, vol. 8, no. 6, pp. 221–223, Dec. 2008, doi: 10.1093/BJACEACCP/MKN041.
- [276] S. Safari, A. Baratloo, M. Elfil, and A. Negida, “Evidence Based Emergency Medicine Part 2: Positive and negative predictive values of diagnostic tests,” *Emergency*, vol. 3, no. 3, p. 87, 2015, doi: 10.22037/aaem.v3i3.274.
- [277] G. Hughes, “Youden’s index and the weight of evidence,” *Methods Inf. Med.*, vol. 54, no. 2, pp. 198–199, 2015, doi: 10.3414/ME14-04-0003.

Δημοσιεύσεις σε Διεθνή Περιοδικά και Συνέδρια (με χρονολογική σειρά)

1. **M. Tarousi**, P. Bountris, P. Daskalakis, and D. D. Koutsouris, “Development of a Cost-Effective Intelligent Clinical Decision Support System for Breast Cancer Early Diagnosis and Triage,” *Int. J. Online Biomed. Eng.*, vol. 18, no. 05, pp. 43–64, Apr. 2022, doi: 10.3991/IJOE.V18I05.29067.
2. P. G. Tzouras et al., Panagiotis G.Tzouras, L. Mitropoulos, E. Stavropoulou, E. Antoniou, K.Koliou, C.Karolemeas, A.Karaloulis, K. Mitropoulos, **M. Tarousi**, E.I. Vlahogianni and K. Kepaptsolou “Agent-based models for simulating e-scooter sharing services: A review and a qualitative assessment,” *Int. J. Transp. Sci. Technol.*, Mar. 2022, doi: 10.1016/J.IJTST.2022.02.001.
3. K.M. Tsiouris, D. Gatsios, V. Tsakanikas, A. A. Pardalis, I. Kouris, T.Androutsou, **M. Tarousi**, N. V. Sedlar, I. Somarakis, F.Mostajeran, N Filipovic, H. Akker, D. D. Koutsouris and D.I. Fotiadis, “Designing interoperable telehealth platforms: bridging IoT devices with cloud infrastructures,” <https://doi.org/10.1080/17517575.2020.1759146>, vol. 14, no. 8, pp. 1194–1218, Sep. 2020, doi: 10.1080/17517575.2020.1759146.
4. M.Sarafidis, **M. Tarousi**, A. Anastasiou, S. Pitoglou, E. Lampoukas, A. Spetsariasis, G. Matsopoulos and D. D. Koutsouris , “Data Quality Challenges in a Learning Health System,” *Stud. Health Technol. Inform.*, vol. 270, pp. 143–147, Jun. 2020, doi: 10.3233/SHTI200139.
5. **M. Tarousi**, M. Sarafidis, P. Katrakazas, I. Kouris, S. Pitoglou, O. Petropoulou, A. Anastasiou, and D. D. Koutsouris, “A Cloud Based Decision Support System Aimed to Contribute in Policy Making for Natural Disaster Related Incidents,” *Int. J. Reliab. Qual. E-Healthcare*, vol. 8, no. 3, pp. 36–46, May 2019, doi: 10.4018/IJRQEH.2019070103.
6. **M. Tarousi**, M. Sarafidis, A. Anastasiou, I. Kouris, O. Petropoulou, and D. Koutsouris, “A Cloud Computing Platform for the Holistic Approach to the Needs of Children and Their Families in Natural Disasters Incidents,” *Proc. 2nd World Conf. Smart Trends Syst. Secur. Sustain. WorldS4 2018*, pp. 203–208, Jan. 2019, doi: 10.1109/WORLDS4.2018.8611604.
7. Y. Dionyssiotis, E.Vellidou, S. Th. Konstantinidis, P. Sarafis, S. Artemi, K. Stergiopoulou, A. M. Vind, N. Birsanu, S. Dupont, D. D. Koutsouris, and EPoCFiNDS Consensus attendants group (I. Iliadi, **M. Tarousi**, M. Sarafidis, M. G. Radu, M. Hvidt, J. Vendelbo), “Education Program for Carers in Facilities with Neuro Disabled Subjects EPoCFiNDS,” *J. Frailty, Sarcopenia Falls*, vol. 4, no. 2, p. 45, Jun. 2019, doi: 10.22540/JFSF-04-045.
8. P Katrakazas, V. Costarides, **M. Tarousi**, M. Christodoulakis, P. Toumpaniaris, S. Pavlopoulos, D. Iliopoulou and D. D. Koutsouris, “Business Process Modelling for a Greek Hospital’s Medical Equipment Data Center,” *Proc. - IEEE Symp. Comput. Med. Syst.*, vol. 2018-June, pp. 328–332, Jul. 2018, doi: 10.1109/CBMS.2018.00064.

9. P. Katrakazas, M. Tarousi, K. Giokas, and D. D. Koutsouris, “ACESO: Analysis of Cervical Cancer: An Evidence-Based Treatments Optimization,” Proc. - IEEE Symp. Comput. Med. Syst., vol. 2017-June, pp. 332–336, Nov. 2017, doi: 10.1109/CBMS.2017.166.