

**ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ
ΣΧΟΛΗ ΧΡΗΜΑΤΟΟΙΚΟΝΟΜΙΚΗΣ ΚΑΙ ΣΤΑΤΙΣΤΙΚΗΣ**



**ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ ΚΑΙ ΑΣΦΑΛΙΣΤΙΚΗΣ
ΕΠΙΣΤΗΜΗΣ**

**ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ ΣΤΗΝ
ΕΦΑΡΜΟΣΜΕΝΗ ΣΤΑΤΙΣΤΙΚΗ**

**ΜΕΛΕΤΗ ΚΛΙΝΙΚΩΝ / ΔΙΑΓΝΩΣΤΙΚΩΝ ΔΕΙΚΤΩΝ ΚΑΙ
ΔΗΜΟΓΡΑΦΙΚΩΝ ΧΑΡΑΚΤΗΡΙΣΤΙΚΩΝ ΠΟΥ ΣΧΕΤΙΖΟΝΤΑΙ
ΜΕ ΤΗΝ ΕΠΙΒΙΩΣΗ ΑΣΘΕΝΩΝ ΜΕ ΚΑΡΚΙΝΟ ΤΟΥ
ΜΑΣΤΟΥ. ΜΕΛΕΤΗ ΔΕΔΟΜΕΝΩΝ ΤΗΣ ΒΑΣΗΣ SEER.**

ΠΑΠΑΔΗΜΗΤΡΙΟΥ ΧΡΙΣΤΟΦΟΡΟΣ

Διπλωματική Εργασία

που υποβλήθηκε στο Τμήμα Στατιστικής και
Ασφαλιστικής Επιστήμης του Πανεπιστημίου
Πειραιώς ως μέρος των απαιτήσεων για την
απόκτηση του Μεταπτυχιακού Διπλώματος
Ειδίκευσης στην Εφαρμοσμένη Στατιστική.

Πειραιάς 2022

**UNIVERSITY OF PIRAEUS
SCHOOL OF FINANCE AND STATISTICS**



**DEPARTMENT OF STATISTICS AND INSURANCE
SCIENCE**

POSTGRADUATE PROGRAM IN APPLIED STATISTICS

**A STUDY OF THE CLINICAL/DIAGNOSIS MARKERS AND
DEMOGRAPHIC CHARACTERISTICS RELATED TO
SURVIVAL IN PATIENTS WITH BREAST CANCER. SEER
BASED STUDY.**

**BY
PAPADIMITRIOU CHRISTOFOROS**

MSc Dissertation

Submitted to the Department of Statistics and Insurance Science of
the University of Piraeus in partial fulfilment of the requirements
for the degree of Master of Science in Applied Statistics.

PIRE AUS 2022

Η παρούσα Διπλωματική Εργασία εγκρίθηκε ομόφωνα από την Τριμελή Εξεταστική Επιτροπή που ορίσθηκε από τη Συνέλευση του Τμήματος Στατιστικής και Ασφαλιστικής Επιστήμης του Πανεπιστημίου Πειραιώς στην υπ' αριθμ. συνεδρίασή της σύμφωνα με τον Εσωτερικό Κανονισμό Λειτουργίας του Προγράμματος Μεταπτυχιακών Σπουδών στην Εφαρμοσμένη Στατιστική

Τα μέλη της Επιτροπής ήταν:

- Αναπληρωτής Καθηγητής Τζαβελάς Γεώργιος (Επιβλέπων)
- Καθηγήτρια Βερροπούλου Γεωργία
- Καθηγήτρια Διεθνούς Πανεπιστημίου Ελλάδος Λαβδανίτη Μαρία

Η έγκριση της Διπλωματικής Εργασίας από το Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης του Πανεπιστημίου Πειραιώς δεν υποδηλώνει αποδοχή των γνώμων του συγγραφέα.

ΕΥΧΑΡΙΣΤΙΕΣ:

Θα ήθελα να ευχαριστήσω τον επιβλέποντα καθηγητή μου κύριο Τζαβελά που μου έδωσε την ευκαιρία να κάνω αυτή την διπλωματική εργασία καθώς και για την βοήθεια που μου παρείχε καθ' όλη τη διάρκεια της προετοιμασίας της.

Στη συνέχεια θα ήθελα να ευχαριστήσω την οικογένεια μου που με στήριξε σε όλη τη διάρκεια του μεταπτυχιακού αυτού, καθώς με στηρίζει σε κάθε απόφαση που θα πάρω.

Παπαδημητρίου Χριστόφορος

ΠΕΡΙΛΗΨΗ

Ο καρκίνος του μαστού είναι ο συχνότερος καρκίνος του γυναικείου πληθυσμού στις ανεπτυγμένες χώρες του δυτικού κόσμου. Ξεκινά όταν η διαδικασία της ανάπτυξης των κυττάρων δεν επιτελείται φυσιολογικά με αποτέλεσμα νέα κύτταρα να σχηματίζονται ανεξέλεγκτα, ενώ ταυτόχρονα κύτταρα που έχουν υποστεί φθορά με το χρόνο δεν αποπίπτουν. Όταν συμβεί αυτό, τα κύτταρα μπορεί να συσσωρευτούν σχηματίζοντας έναν κακοήγη (καρκινικό) όγκο.

Κύριος στόχος της συγκεκριμένης διπλωματικής εργασίας είναι η μελέτη των κλινικών και διαγνωστικών δεικτών και των δημογραφικών χαρακτηριστικών που σχετίζονται με την επιβίωση των ατόμων που πάσχουν από καρκίνο του μαστού. Για την μελέτη αυτή χρησιμοποιήθηκε ένα δείγμα ατόμων από την Ουάσιγκτον, 5552 παρατηρήσεων που αντλήθηκε από μία βάση δεδομένων του λογισμικού SEER*Stat.

Οι βασικές στατιστικές μέθοδοι που χρησιμοποιήθηκαν για την παρούσα μελέτη είναι η πολλαπλή γραμμική παλινδρόμηση, η λογιστική παλινδρόμηση καθώς και το μοντέλο αναλογικού κινδύνου του Cox. Ενώ το μοντέλο πολλαπλής γραμμικής παλινδρόμησης δεν μπορεί να εφαρμοσθεί καθώς δεν πληρούνται οι απαραίτητες συνθήκες, αντίθετα από την προσαρμογή των μοντέλων της λογιστικής παλινδρόμησης και του μοντέλου του Cox συμπεράναμε πως οι βασικοί κλινικοί και διαγνωστικοί δείκτες και τα δημογραφικά χαρακτηριστικά που παίζουν ρόλο στην επιβίωση των ατόμων που πάσχουν από καρκίνο του μαστού είναι η ηλικία, ο βαθμός ομοιότητας των καρκινικών κυττάρων με υγιή κύτταρα, ο υποδοχέας προγεστερόνης, ο υποδοχέας οιστρογόνου, το μέγεθος και η θέση του όγκου, ο αριθμός των λεμφαδένων ενός ατόμου που έχουν καρκίνο, ο αριθμός των περιφερειακών όγκων που εξετάστηκαν καθώς και ο αριθμός των περιφερειακών όγκων που είναι κακοήθεις.

ABSTRACT

Breast cancer is the most common cancer of the female population in the developed countries of the Western world. It begins when the process of cell growth is not carried out normally, resulting in new cells being formed uncontrollably, while cells that have been damaged over time do not fall off. When this happens, the cells can accumulate to form a malignant (cancerous) tumor.

The main objective of this thesis is to study the clinical and diagnostic indicators and the demographic characteristics related to the survival of people suffering from breast cancer. A sample from Washington of 5552 observations drawn from a SEER*Stat software database was used for this study.

The main statistical methods used in the present study are multiple linear regression, logistic regression and the Cox proportional hazard model. The multiple linear regression cannot be applied since the necessary conditions are not met. The use of logistic regression and Cox models reveals that the key clinical and diagnostic indicators as well as the demographic characteristics that play a role in the survival of breast cancer patients, are age, the similarity of cancer cells with the normal cells, the progesterone receptor, the estrogen receptor, the size and location of the tumor, the number of a person's lymph nodes that have cancer, the number of regional tumors examined as well as the number regional tumors that are positive.

ΠΕΡΙΕΧΟΜΕΝΑ

Κατάλογος Πινάκων	12
Κατάλογος Διαγραμμάτων	19
Κατάλογος Εικόνων	23
ΚΕΦ. 1^ο	24
1.1 Εισαγωγή	24
1.2 Ανατομία του μαστού - Καρκίνος του μαστού	25
1.2.1 Αιτίες εμφάνισης του καρκίνου του μαστού	26
1.2.2 Είδη καρκίνου του μαστού	32
1.2.3 Συμπτώματα	37
1.2.4 Στάδια της νόσου	38
1.2.5 Διαγνωστική διερεύνηση της νόσου	39
1.2.6 Θεραπευτική αγωγή του καρκίνου του μαστού	40
ΚΕΦ. 2^ο	44
2.1 Εισαγωγή	44
2.2 Περιγραφική Ανάλυση	44
2.2.1 Ποσοτικές Μεταβλητές	47
2.2.2 Ποιοτικές Μεταβλητές	52
ΚΕΦ. 3^ο	67
3.1 Εισαγωγή	67
3.1.1 Έλεγχος χ^2	67
3.1.2 Συντελεστής συσχέτισης του Pearson	68

3.1.3	Ο συντελεστής συσχέτισης του Spearman	69
3.1.4	Έλεγχος Mann-Whitney U	70
3.1.5	Kruskal Wallis	71
3.2	Έλεγχος συσχέτισης ποσοτικών μεταβλητών	72
3.3	Έλεγχος συσχέτισης ποιοτικών μεταβλητών	74
3.4	Έλεγχος συσχέτισης ποιοτικών και ποσοτικών μεταβλητών	97
ΚΕΦ. 4ο		123
4.1	Εισαγωγή	123
4.2	Πολλαπλή Παλινδρόμηση	123
4.2.1	Υποθέσεις του πολλαπλού γραμμικού μοντέλου	124
4.3	Η Μέθοδος Stepwise Regression	124
4.3.1	Η μέθοδος Stepwise για τον πληθυσμό ατόμων που πέθαναν από καρκίνο του μαστού - Status = 1 .	125
ΚΕΦ. 5ο		131
5.1	Εισαγωγή	131
5.2	Λογιστική Παλινδρόμηση	131
5.3	Εφαρμογή του μοντέλου της λογιστικής παλινδρόμησης με εξαρτημένη μεταβλητή τη μεταβλητή "Status".	132
ΚΕΦ. 6ο		137
6.1	Εισαγωγή	137
6.2	Μοντέλο Κινδύνων του Cox	137
6.3	Θνησιμότητα από τον καρκίνο του μαστού	137
ΚΕΦ. 7ο		151
7.1	Εισαγωγή	151
7.2	Συμπεράσματα από την περιγραφική ανάλυση των μεταβλητών	151

7.3	Συμπεράσματα από τις συσχετίσεις των μεταβλητών	153
7.4	Συμπεράσματα από την προσαρμογή των μοντέλων	154
	Βιβλιογραφία	156

Κατάλογος Πινάκων

ΚΕΦ. 2^ο

Πίνακας 2.1 : Περιγραφικά μέτρα για τη μεταβλητή “tumor size”.	47
Πίνακας 2.2 : Περιγραφικά μέτρα για τη μεταβλητή “Regional_Nodes_Examined”.	49
Πίνακας 2.3: Περιγραφικά μέτρα για τη μεταβλητή “Regional_Nodes_Positive”.	50
Πίνακας 2.4: Περιγραφικά μέτρα για τη μεταβλητή “Survival_Months”.	51
Πίνακας 2.5 : Συχνότητες ανά ηλικιακή ομάδα.	53
Πίνακας 2.6 : Συχνότητες ανά φυλή.	54
Πίνακας 2.7 : Συχνότητες ανά οικογενειακή κατάσταση.	55
Πίνακας 2.8 : Συχνότητες ανά μία μέτρηση του πόσο τα καρκινικά κύτταρα μοιάζουν με κανονικά.	56
Πίνακας 2.9 : Συχνότητες ανά την προέλευση πρωτογενούς.	57
Πίνακας 2.10 : Συχνότητες ανά την έκταση της νόσου.	58
Πίνακας 2.11 : Συχνότητες ανά τους υποδοχείς οιστρογόνου.	59
Πίνακας 2.12 : Συχνότητες ανά τους υποδοχείς προγεστερόνης.	60
Πίνακας 2.13 : Συχνότητες ανά του δευτερεύοντος τύπου μαστού.	61
Πίνακας 2.14 : Συχνότητες ανά του 6 ^ο σταδίου.	62
Πίνακας 2.15 : Συχνότητες ανά του T σταδίου.	64

Πίνακας 2.16 : Συχνότητες ανά του Ν σταδίου.	65
Πίνακας 2.17 : Συχνότητες ανά του αν κάποιος πέθανε από καρκίνο του μαστού.	66

ΚΕΦ. 3^ο

Πίνακας 3.1 : Προϋποθέσεις για τη χρήση του κριτηρίου.	67
Πίνακας 3.2 : Μέγεθος δείκτη συσχέτισης του Pearson.	69
Πίνακας 3.3 : Όταν τα δεδομένα είναι περίπου ελλειπτικά κατανομημένα και δεν υπάρχουν εμφανείς ακραίες τιμές, η συσχέτιση Spearman και Pearson δίνουν παρόμοιες τιμές.	70
Πίνακας 3.4 : Έλεγχος κανονικότητας των σφαλμάτων.	72
Πίνακας 3.5 : Έλεγχος συσχετίσεων με τη μέθοδο του Spearman.	73
Πίνακας 3.6 : Περίληψη επεξεργασίας υπόθεσης «Συσχέτιση θνησιμότητας από τον καρκίνο του μαστού με ηλικία».	75
Πίνακας 3.7 : «Διασταυρώσεις» των μεταβλητών Status και Age.	75
Πίνακας 3.8 : Έλεγχος χ^2 για τις μεταβλητές Status και Age.	75
Πίνακας 3.9 : Περίληψη επεξεργασίας υπόθεσης «Συσχέτιση θνησιμότητας από τον καρκίνο του μαστού με φυλή».	76
Πίνακας 3.10 : «Διασταυρώσεις» των μεταβλητών Status και Race.	77
Πίνακας 3.11 : Έλεγχος χ^2 για τις μεταβλητές Status και Race..	77

Πίνακας 3.12 : Περίληψη επεξεργασίας υπόθεσης «Συσχέτιση θνησιμότητας από τον καρκίνο του μαστού με οικογενειακή κατάσταση».	78
Πίνακας 3.13 : «Διασταυρώσεις» των μεταβλητών Status και Marital_Status.	79
Πίνακας 3.14 : Έλεγχος χ^2 για τις μεταβλητές Status και Marital_Status.	79
Πίνακας 3.15 : Περίληψη επεξεργασίας υπόθεσης «Συσχέτιση θνησιμότητας από τον καρκίνο του μαστού με την μέτρηση για το κατά πόσο τα καρκινικά κύτταρα μοιάζουν με τα κανονικά».	80
Πίνακας 3.16 : «Διασταυρώσεις» των μεταβλητών Status και Grade.	81
Πίνακας 3.17 : Έλεγχος χ^2 για τις μεταβλητές Status και Grade.	81
Πίνακας 3.18 : Περίληψη επεξεργασίας υπόθεσης «Συσχέτιση θνησιμότητας από τον καρκίνο του μαστού με την προέλευση πρωτογενούς».	82
Πίνακας 3.19 : «Διασταυρώσεις» των μεταβλητών Status και Laterality.	83
Πίνακας 3.20 : Έλεγχος χ^2 για τις μεταβλητές Status και Laterality.	83
Πίνακας 3.21 : Περίληψη επεξεργασίας υπόθεσης «Συσχέτιση θνησιμότητας από τον καρκίνο του μαστού με την έκταση της νόσου».	83
Πίνακας 3.22 : «Διασταυρώσεις» των μεταβλητών Status και StageA.	84
Πίνακας 3.23 : Έλεγχος χ^2 για τις μεταβλητές Status και StageA.	84
Πίνακας 3.24 : Περίληψη επεξεργασίας υπόθεσης «Συσχέτιση θνησιμότητας από τον καρκίνο του μαστού με τους υποδοχείς οιστρογόνου».	85

Πίνακας 3.25 : «Διασταυρώσεις» των μεταβλητών Status και ER_Status.	86
Πίνακας 3.26 : Έλεγχος χ^2 για τις μεταβλητές Status και ER_Status.	86
Πίνακας 3.27 : Περίληψη επεξεργασίας υπόθεσης «Συσχέτιση θνησιμότητας από τον καρκίνο του μαστού με τους υποδοχείς προγεστερόνης».	87
Πίνακας 3.28 : «Διασταυρώσεις» των μεταβλητών Status και PR_Status.	88
Πίνακας 3.29 : Έλεγχος χ^2 για τις μεταβλητές Status και PR_Status.	88
Πίνακας 3.30 : Περίληψη επεξεργασίας υπόθεσης «Συσχέτιση θνησιμότητας από τον καρκίνο του μαστού με τον δευτερεύον τύπο μαστού».	89
Πίνακας 3.31 : «Διασταυρώσεις» των μεταβλητών Status και BreastSubtype.	90
Πίνακας 3.32 : Έλεγχος χ^2 για τις μεταβλητές Status και BreastSubtype.	90
Πίνακας 3.33 : Περίληψη επεξεργασίας υπόθεσης «Συσχέτιση θνησιμότητας από τον καρκίνο του μαστού με το 6 ^ο στάδιο».	91
Πίνακας 3.34 : «Διασταυρώσεις» των μεταβλητών Status και 6 th Stage.	92
Πίνακας 3.35 : Έλεγχος χ^2 για τις μεταβλητές Status και 6 th Stage.	92
Πίνακας 3.36 : Περίληψη επεξεργασίας υπόθεσης «Συσχέτιση θνησιμότητας από τον καρκίνο του μαστού με το T στάδιο».	93
Πίνακας 3.37 : «Διασταυρώσεις» των μεταβλητών Status και T Stage.	94
Πίνακας 3.38 : Έλεγχος χ^2 για τις μεταβλητές Status και T Stage.	94

Πίνακας 3.39 : Περίληψη επεξεργασίας υπόθεσης «Συσχέτιση θνησιμότητας από τον καρκίνο του μαστού με το N στάδιο».	95
Πίνακας 3.40 : «Διασταυρώσεις» των μεταβλητών Status και N Stage.	96
Πίνακας 3.41 : Έλεγχος χ^2 για τις μεταβλητές Status και N Stage.	96
Πίνακας 3.42 : Έλεγχος συσχετίσεων με τη μέθοδο του Mann-Whitney.	98
Πίνακας 3.43 : Έλεγχος συσχετίσεων με τη μέθοδο του Kruskal - Wallis - μήνες επιβίωσης με την ηλικία.	101
Πίνακας 3.44 : levels της μεταβλητής Age που συσχετίζονται μεταξύ τους.	102
Πίνακας 3.45 : Έλεγχος συσχετίσεων με τη μέθοδο του Kruskal - Wallis - μήνες επιβίωσης με την φυλή.	106
Πίνακας 3.46 : Έλεγχος συσχετίσεων με τη μέθοδο του Kruskal - Wallis - μήνες επιβίωσης με την οικογενειακή κατάσταση.	107
Πίνακας 3.47 : levels της μεταβλητής Marital_Status που συσχετίζονται μεταξύ τους.	108
Πίνακας 3.48 : Έλεγχος συσχετίσεων με τη μέθοδο του Kruskal - Wallis - μήνες επιβίωσης με μία μέτρηση του κατά πόσο τα καρκινικά κύτταρα μοιάζουν με τα κανονικά.	109
Πίνακας 3.49 : levels της μεταβλητής Grade που συσχετίζονται μεταξύ τους.	110
Πίνακας 3.50 : Έλεγχος συσχετίσεων με τη μέθοδο του Mann Whitney - μήνες επιβίωσης με την προέλευση πρωτογενούς.	110
Πίνακας 3.51 : Έλεγχος συσχετίσεων με τη μέθοδο του Mann Whitney - μήνες επιβίωσης με το στάδιο A.	111

Πίνακας 3.52 : Έλεγχος συσχετίσεων με τη μέθοδο του Mann Whitney - μήνες επιβίωσης με τους υποδοχείς οιστρογόνου.	112
Πίνακας 3.53 : Έλεγχος συσχετίσεων με τη μέθοδο του Mann Whitney - μήνες επιβίωσης με τους υποδοχείς προγεστερόνης.	113
Πίνακας 3.54 : Έλεγχος συσχετίσεων με τη μέθοδο του Kruskal - Wallis - μήνες επιβίωσης με τον δευτερεύον τύπο μαστού.	114
Πίνακας 3.55 : levels της μεταβλητής BreastSubtype που συσχετίζονται μεταξύ τους.	115
Πίνακας 3.56 : Έλεγχος συσχετίσεων με τη μέθοδο του Kruskal - Wallis - μήνες επιβίωσης με το 6 ^ο στάδιο.	115
Πίνακας 3.57 : levels της μεταβλητής 6 th Stage που συσχετίζονται μεταξύ τους.	117
Πίνακας 3.58 : Έλεγχος συσχετίσεων με τη μέθοδο του Kruskal - Wallis - μήνες επιβίωσης με το T στάδιο.	118
Πίνακας 3.59 : levels της μεταβλητής Stage_T που συσχετίζονται μεταξύ τους.	119
Πίνακας 3.60 : Έλεγχος συσχετίσεων με τη μέθοδο του Kruskal - Wallis - μήνες επιβίωσης με το N στάδιο.	121
Πίνακας 3.61 : levels της μεταβλητής Stage_N που συσχετίζονται μεταξύ τους.	122
ΚΕΦ. 4 ^ο	
Πίνακας 4.1 : Η μέθοδος Stepwise για τον πληθυσμό ατόμων που πέθαναν από καρκίνο του μαστού - Status = 1.	125

Πίνακας 4.2 : Χαρακτηριστικά των μεταβλητών που μπήκαν στο μοντέλο με τη μέθοδο Stepwise. 127

Πίνακας 4.3 : Έλεγχος κανονικότητας των σφαλμάτων. 129

ΚΕΦ. 5^ο

Πίνακας 5.1 : Περίληψη επεξεργασίας του μοντέλου της λογιστικής παλινδρόμησης με εξαρτημένη μεταβλητή τη μεταβλητή “Status”. 132

Πίνακας 5.2 : Έλεγχος Hosmer and Lemeshow. 132

Πίνακας 5.3 : Classification table με εξαρτημένη μεταβλητή το “Status”. 133

Πίνακας 5.4 : Ερμηνεία μεταβλητών. 133

ΚΕΦ. 6^ο

Πίνακας 6.1 : Περίληψη επεξεργασίας του μοντέλου του Cox με εξαρτημένη μεταβλητή τη μεταβλητή “Survival_Months”. 138

Πίνακας 6.2 : Δοκιμές γενικών συντελεστών μοντέλου. 138

Πίνακας 6.3 : Ερμηνεία μεταβλητών. 139

Πίνακας 6.4 : Περιγραφικά μέτρα μεταβλητών. 148

Πίνακας 6.5 : Περιγραφικά μέτρα μεταβλητών. 149

Κατάλογος Διαγραμμάτων

ΚΕΦ. 2ο

Διάγραμμα 2.1 : Ιστόγραμμα για τη μεταβλητή “tumor size”.	48
Διάγραμμα 2.2 : Ιστόγραμμα για τη μεταβλητή “Regional_Nodes_Examined”.	49
Διάγραμμα 2.3: Ιστόγραμμα για τη μεταβλητή “Regional_Nodes_Positive”.	50
Διάγραμμα 2.4: Ιστόγραμμα για τη μεταβλητή “Survival_Months”.	52
Διάγραμμα 2.5 : Bar chart συχνοτήτων με βάση τις ηλικιακές ομάδες.	54
Διάγραμμα 2.6 : Bar chart συχνοτήτων με βάση την φυλή.	55
Διάγραμμα 2.7 : Bar chart συχνοτήτων με βάση την οικογενειακή κατάσταση.	56
Διάγραμμα 2.8 : Bar chart συχνοτήτων με βάση μία μέτρηση του πόσο τα καρκινικά κύτταρα μοιάζουν με κανονικά.	57
Διάγραμμα 2.9 : Bar chart συχνοτήτων με βάση την προέλευση πρωτογενούς.	58
Διάγραμμα 2.10 : Bar chart συχνοτήτων με βάση την έκταση της νόσου.	59
Διάγραμμα 2.11 : Bar chart συχνοτήτων με βάση τους υποδοχείς οιστρογόνου.	60
Διάγραμμα 2.12 : Bar chart συχνοτήτων με βάση τους υποδοχείς προγεστερόνης.	61
Διάγραμμα 2.13 : Bar chart συχνοτήτων με βάση τον δευτερεύον τύπο μαστού.	62

Διάγραμμα 2.14 : Bar chart συχνοτήτων με βάση το 6 ^ο στάδιο.	63
Διάγραμμα 2.15 : Bar chart συχνοτήτων με βάση το T στάδιο.	64
Διάγραμμα 2.16 : Bar chart συχνοτήτων με βάση το N στάδιο.	65
Διάγραμμα 2.17 : Bar chart συχνοτήτων με βάση το αν κάποιος πέθανε από καρκίνο του μαστού.	66

ΚΕΦ. 3^ο

Διάγραμμα 3.1 : Διάγραμμα διασποράς μεταξύ των μεταβλητών "Tumor_Size" και "Survival_Months".	73
Διάγραμμα 3.2 : Διάγραμμα διασποράς μεταξύ των μεταβλητών "Regional_Nodes_Examined" και "Survival_Months".	74
Διάγραμμα 3.3 : Ραβδόγραμμα μεταξύ των μεταβλητών "Age" και "Status".	76
Διάγραμμα 3.4 : Ραβδόγραμμα μεταξύ των μεταβλητών "Race" και "Status".	78
Διάγραμμα 3.5 : Ραβδόγραμμα μεταξύ των μεταβλητών "Marital_Status" και "Status".	80
Διάγραμμα 3.6 : Ραβδόγραμμα μεταξύ των μεταβλητών "Grade" και "Status".	82
Διάγραμμα 3.7 : Ραβδόγραμμα μεταξύ των μεταβλητών "StageA" και "Status".	85
Διάγραμμα 3.8 : Ραβδόγραμμα μεταξύ των μεταβλητών "ER_Status" και "Status".	87
Διάγραμμα 3.9 : Ραβδόγραμμα μεταξύ των μεταβλητών "PR_Status" και "Status".	89
Διάγραμμα 3.10 : Ραβδόγραμμα μεταξύ των μεταβλητών "BreastSubtype" και "Status".	91

Διάγραμμα 3.11 : Ραβδόγραμμα μεταξύ των μεταβλητών “6 th Stage ” και “Status”.	93
Διάγραμμα 3.12 : Ραβδόγραμμα μεταξύ των μεταβλητών “T Stage ” και “Status”.	95
Διάγραμμα 3.13 : Ραβδόγραμμα μεταξύ των μεταβλητών “N Stage ” και “Status”.	97
Διάγραμμα 3.14 : Θηκόγραμμα μεταξύ των μεταβλητών “Tumor_Size” και “Status”.	99
Διάγραμμα 3.15 : Θηκόγραμμα μεταξύ των μεταβλητών “Regional_Nodes_Examined” και “Status”.	99
Διάγραμμα 3.16 : Θηκόγραμμα μεταξύ των μεταβλητών “Regional_Nodes_Positive” και “Status”.	100
Διάγραμμα 3.17 : Θηκόγραμμα μεταξύ των μεταβλητών “Survival_Months” και “Status”.	100
Διάγραμμα 3.18 : Θηκόγραμμα μεταξύ των μεταβλητών “Survival_Months” και “Age”.	101
Διάγραμμα 3.19 : Θηκόγραμμα μεταξύ των μεταβλητών “Survival_Months” και “Marital_Status”.	107
Διάγραμμα 3.20 : Θηκόγραμμα μεταξύ των μεταβλητών “Survival_Months” και “Grade”.	109
Διάγραμμα 3.21 : Θηκόγραμμα μεταξύ των μεταβλητών “Survival_Months” και “StageA”.	111
Διάγραμμα 3.22 : Θηκόγραμμα μεταξύ των μεταβλητών “Survival_Months” και “ER_Status”.	112
Διάγραμμα 3.23 : Θηκόγραμμα μεταξύ των μεταβλητών “Survival_Months” και “PR_Status”.	113
Διάγραμμα 3.24 : Θηκόγραμμα μεταξύ των μεταβλητών “Survival_Months” και “BreastSubtype”.	114

Διάγραμμα 3.25 : Θηκόγραμμα μεταξύ των μεταβλητών "Survival_Months" και "6 th Stage".	116
Διάγραμμα 3.26 : Θηκόγραμμα μεταξύ των μεταβλητών "Survival_Months" και "T Stage".	118
Διάγραμμα 3.27 : Θηκόγραμμα μεταξύ των μεταβλητών "Survival_Months" και "N Stage".	121

ΚΕΦ. 6^ο

Διάγραμμα 6.1 : Συνάρτηση Επιβίωσης.	146
Διάγραμμα 6.2 : Συνάρτηση Κινδύνου.	147
Διάγραμμα 6.3 : Θηκόγραμμα Συνάρτηση Επιβίωσης.	148
Διάγραμμα 6.4 : Θηκόγραμμα Συνάρτηση Κινδύνου.	149

Κατάλογος Εικόνων

ΚΕΦ. 1^ο

Εικόνα 1.1: Μαστογραφία καρκίνου. Παρατηρήστε το ανώμαλο περίγραμμα που θυμίζει ακτινωτό άστρο. Πορογενής διηθητικός καρκίνος.	26
Εικόνα 1.2: Συχνότητα του καρκίνου του μαστού ανά τεταρτημόριο.	27
Εικόνα 1.3: Όσο αυξάνεται η ηλικία μέχρι τα 75 τόσο αυξάνεται και η πιθανότητα προσβολής από καρκίνο του μαστού.	28
Εικόνα 1.4: Ποσοστό κληρονομικότητας που οφείλεται στον καρκίνο του μαστού.	29
Εικόνα 1.5: Από τα αδενικά λόβια (1) που παράγουν το γάλα προέρχεται ο λοβιακός καρκίνος. Από τους πόρους που μεταφέρουν το γάλα προέρχεται ο πορογενής καρκίνος (2).	33
Εικόνα 1.6: In situ και διηθητικός καρκίνος.	34
Εικόνα 1.7: Νόσος Paget της θηλής.	36
Εικόνα 1.8: Φλεγμονώδης καρκίνος του μαστού.	37

ΚΕΦΑΛΑΙΟ 1

ΕΙΣΑΓΩΓΗ

1.1 Εισαγωγή

Ο όρος «καρκίνος του μαστού» αναφέρεται στην ανάπτυξη κακοήθους όγκου στην περιοχή του μαστού. Αποτελεί μία από τις συχνότερα εμφανιζόμενες μορφές καρκίνου παγκοσμίως και είναι η πρώτη σε αριθμό κρουσμάτων στο γυναικείο πληθυσμό. Προκαλείται από ανεξέλεγκτο πολλαπλασιασμό παθολογικών κυττάρων που ως αποτέλεσμα προκαλούν το σχηματισμό κακοήθους όγκου στην περιοχή του μαστού και ουσιαστικά αποτελεί κυτταρική νόσο. Τα παθολογικά αυτά κύτταρα έχουν τη δυνατότητα εξάπλωσης σε γειτονικούς ιστούς με δυσάρεστες συνέπειες για ολόκληρο τον οργανισμό. Η πιθανότητα εμφάνισης της νόσου σε άρρενες είναι υπαρκτή αλλά πολύ μικρή. Όσον αφορά στις γυναίκες όλες αντιμετωπίζουν τον κίνδυνο εμφάνισης της νόσου - όχι, όμως στον ίδιο βαθμό.

Σύμφωνα με τη Διεθνή Έκθεση για τις καρκινικές νόσους που παρουσιάστηκε στη Γενεύη της Ελβετίας τον Απρίλιο του 2003 από την IARC (International Agency for Research on Cancer), ο καρκίνος του μαστού αποτελεί την πιο συχνή μορφή καρκίνου μεταξύ των γυναικών, με περίπου 1.000.000 νέα κρούσματα παγκοσμίως. Ορισμένοι αριθμοί δίνουν την κοινωνική διάσταση του καρκίνου του μαστού. Στις ΗΠΑ το 2004 οι νέες περιπτώσεις καρκίνου του μαστού στις γυναίκες ήταν 217.000 περίπου. Στην Ελλάδα αναφέρονται 4.500 περίπου νέες περιπτώσεις το χρόνο, ενώ υπολογίζεται ότι 1 στις 8 γυναίκες παγκοσμίως θα παρουσιάσει καρκίνο μαστού σε κάποια φάση της ζωής της. Στην Ευρώπη, το 60% των κρουσμάτων καρκίνου του μαστού διαγιγνώσκεται σε πρώιμο στάδιο. Το αντίστοιχο ποσοστό στην Ελλάδα είναι μόλις 5%. Τα στοιχεία αυτά καταδεικνύουν πόσο ελλιπής είναι η σχετική ενημέρωση μεταξύ των Ελληνίδων.

Είναι ένα από τα είδη καρκίνου που επιφέρουν τους περισσότερους θανάτους ετησίως (519.000 θανάτους το 2004). Παρατηρείται εξαιρετικά σπάνια στους άνδρες, με περίπου 100 φορές μικρότερη συχνότητα συγκριτικά με τις γυναίκες και τα ίδια ποσοστά επιβίωσης. Ωστόσο, τα ποσοστά θανάτου από καρκίνο του μαστού χαρακτηρίζονται από πτωτική τάση από τις αρχές του 1990, με τις μεγαλύτερες μειώσεις να εντοπίζονται στις γυναίκες κάτω των 50. Οι ερευνητές αποδίδουν την πτώση αυτή στην έγκαιρη διάγνωση μέσω μαστογραφιών καθώς και στις βελτιώσεις που έχουν επέλθει στις σχετικές θεραπευτικές αγωγές. Ο αριθμός των ατόμων που έχουν αντιμετωπίσει με επιτυχία τον καρκίνο του μαστού αυξάνεται συνεχώς - από τον Ιανουάριο του 2006, υπήρξαν περίπου 2,5 εκατομμύρια γυναίκες στις ΗΠΑ που, βάσει της έκθεσης, έχουν ξεπεράσει με επιτυχία την περιπέτεια του καρκίνου του μαστού.

Οι διάφορες μελέτες που έχουν πραγματοποιηθεί με την πάροδο των ετών για τον καρκίνο του μαστού επικεντρώνονται μεταξύ άλλων και στους διάφορους παράγοντες

οι οποίοι είναι δυνατόν να επηρεάζουν τον οργανισμό και να οδηγούν στην εμφάνιση του. Σκοπός αυτής της εργασίας είναι να μελετήσει τόσο τους δημογραφικούς παράγοντες όσο και τους κλινικούς-διαγνωστικούς δείκτες που μπορεί να έχουν σχέση με τον καρκίνο του μαστού. Στην κατανόηση αυτών συνέβαλλαν πληροφορίες που αντλήθηκαν από σπουδαιούς οργανισμούς και γιατρούς, όπως την κλινική-κέντρο μαστού, το Άλμα ζωής Πανελλήνιος Σύλλογος Γυναικών με Καρκίνο του Μαστού, τον κ. Γιάννη Π. Φύσσα, Μαστολόγο Χειρουργό Μαστού, τον κ. Γεώργιο Ανθιμίδη, Γενικό Χειρουργό και τον κ. Γεώργιο Β. Καλόγηρο, Γυναικολόγο Μαιευτήρα.

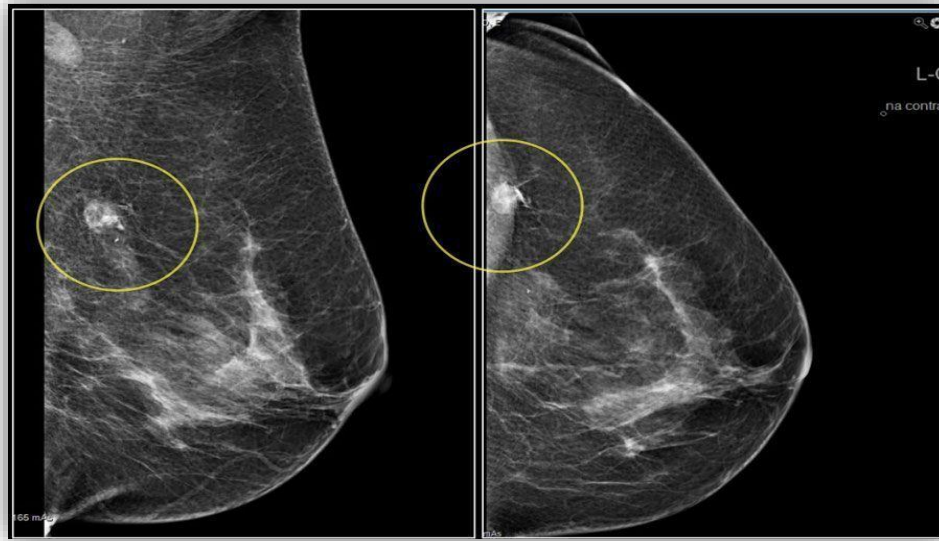
1.2 Ανατομία του μαστού - Καρκίνος του μαστού

Ο γυναικείος μαστός αποτελείται κυρίως από αδενικό, λιπώδη και συνδετικό ιστό. Διασχίζεται από αρτηρίες, φλέβες, λεμφικά αγγεία και νεύρα. Στην ανατομία της περιοχής περιλαμβάνονται επίσης, οι λεμφαδένες στις μασχαλιαίες κοιλότητες, πάνω και κάτω από την κλείδα, καθώς και μέσα στον ίδιο τον μαστό. Ο συνδετικός ιστός καθιστά το στήθος ελαστικό. Το πραγματικό σχήμα του μαστού καθορίζεται από την κατάσταση των κρεμαστήρων συνδέσμων, το βαθμό πτώσης, την περιεκτικότητα σε λιπώδη ιστό και φυσικά από το δέρμα του μαστού.

Προκειμένου να είναι δυνατό να προσδιοριστεί η ακριβής θέση μιας αλλαγής στον μαστό, χωρίζεται σχηματικά σε τέσσερα τεταρτημόρια. Αυτά σχηματίζονται από μία κάθετη και μία οριζόντια νοητή γραμμή που διαπερνά τη θηλή. Οι περισσότερες κακοήθεις αλλαγές συμβαίνουν στο εξωτερικό άνω τεταρτημόριο, δηλαδή αυτό που βρίσκεται πιο κοντά στη σύστοιχη μασχαλιαία κοιλότητα.

Ο μαζικός αδένας αντιδρά στις κυκλικές μεταβολές των γυναικείων ορμονών. Γι' αυτό και το γυναικείο στήθος παρουσιάζει μεταβολές κατά τη διάρκεια του εμμηνορρυσιακού κύκλου και αλλάζει μορφή ανάλογα με τη φάση αυτού. Η εγκυμοσύνη επίσης, αλλάζει το μέγεθος και τη σύσταση του μαστού, ανάλογα με το επίπεδο ορμονών και την φάση της κύησης. Η πλήρης μορφολογική ανάπτυξη και ωρίμανση του μαστού λαμβάνει χώρα μόνο μετά την εγκυμοσύνη και τον θηλασμό. Στην εμμηνόπαυση αυτές οι διαδικασίες μετασχηματισμού σταματούν. Ο ιστός παρόλα αυτά παραμένει ευαίσθητος στις ορμονικές επιδράσεις και μεταβάλλεται με την πάροδο του χρόνου, ακόμα και σε γυναίκες προχωρημένης ηλικίας.

Η ανάπτυξη των καρκινικών κυττάρων σχετίζεται στις περισσότερες περιπτώσεις με ορμονικές επιδράσεις. Οι ορμόνες είναι ουσίες που παράγονται φυσικά στο σώμα. Λειτουργούν ως «αγγελιοφόροι» που λένε σε όργανα ή ιστούς να εκτελούν διάφορες λειτουργίες. Τα κύτταρα του μαστού περιέχουν ειδικές πρωτεΐνες που ονομάζονται υποδοχείς ορμονών. Αυτοί οι υποδοχείς «λαμβάνουν» μηνύματα από ορμόνες στο σώμα και ανταποκρίνονται λέγοντας στα κύτταρα τι να κάνουν. Συνήθως, πρόκειται για ένα μονήρη, κακοήθη όγκο ή ογκίδιο. Στο 10-15% των περιπτώσεων μπορεί να βρεθούν και άλλες εστίες καρκίνου κοντά στον πρώτο όγκο και στο 2-4% και άλλες εστίες πιο μακριά, στον ίδιο μαστό. Στο 2% των γυναικών με καρκίνο μαστού, μπορεί να βρεθεί τη μέρα της διάγνωσης καρκίνος και στον άλλο μαστό και η πιθανότητα είναι 0.5-1% ετησίως, δηλαδή 15-30% σε διάρκεια 30 ετών. Στις μισές γυναίκες με πολυεστιακή ανάπτυξη καρκίνου, θα υπάρχει και διήθηση των μασχαλιαίων λεμφαδένων.



Εικόνα 1.1: Μαστογραφία καρκίνου. Παρατηρήστε το ανώμαλο περίγραμμα που θυμίζει ακτινωτό άστρο. Πορογενής διηθητικός καρκίνος.

1.2.1 Αιτίες εμφάνισης του καρκίνου του μαστού

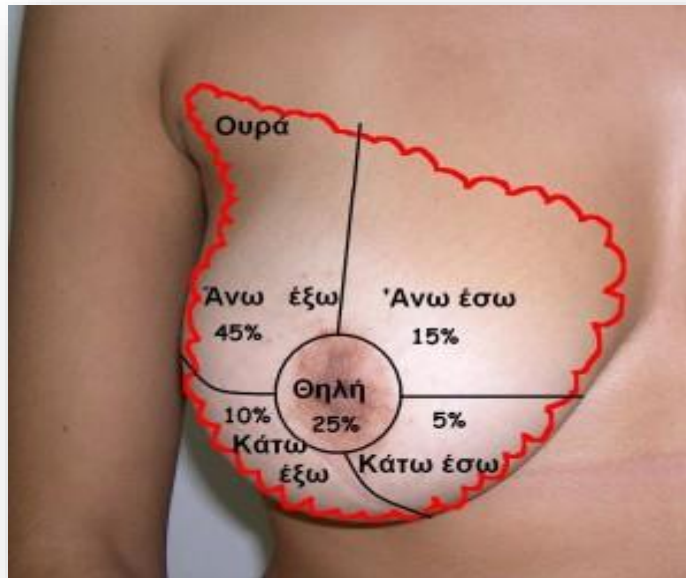
Ο καρκίνος του μαστού δεν είναι μια ασθένεια με σαφή αιτιολογία, όπως π.χ. η ηπατίτιδα, η φυματίωση, η ευλογιά, ο τέτανος κλπ. που μπορούμε με βεβαιότητα σήμερα να πούμε, ότι προκαλούνται από συγκεκριμένα γνωστά μικρόβια και ιούς. Ούτε υπάρχει ένα μόνο είδος καρκίνου του μαστού. Ξέρουμε ότι κάποιες γυναίκες έχουν περισσότερες πιθανότητες από κάποιες άλλες να πάθουν καρκίνο του μαστού και έχουμε βρει παράγοντες που αυξάνουν και άλλους που ελαττώνουν αυτόν τον κίνδυνο.

Αυτοί είναι δύο ειδών: **γενετικοί**, που τους μεταφέρουμε μέσα στα κύτταρά μας και **περιβαλλοντικοί** που μας επηρεάζουν μετά τη γέννησή μας. Η γνώση των παραγόντων αυτών μπορεί να βοηθήσει στην αξιολόγηση των γυναικών υψηλού κινδύνου ώστε αυτές να παρακολουθούνται σχολαστικά και να αντιμετωπίζονται έγκαιρα αν νοσήσουν, αλλά και να αλλάξουν ορισμένες συνήθειες της ζωής τους που θεωρούνται επικίνδυνες. Όλοι αυτοί οι παράγοντες πρέπει να περιλαμβάνονται στο ερωτηματολόγιο του ιστορικού του μαστού και οι ερωτήσεις για αυτούς, δεν πρέπει ποτέ να παραλείπονται. Παράγοντες αυξημένου κινδύνου είναι: Η **προχωρημένη ηλικία**, η βεβαρυμένη με καρκίνο ατομική και οικογενειακή **κληρονομικότητα**, η **μεγάλη διάρκεια της γόνιμης περιόδου**, η **καθυστερημένη εγκυμοσύνη**, η **παράλειψη του θηλασμού**, ο **τόπος καταγωγής**, η **παχυσαρκία**, η **κατάχρηση λιπαρών ουσιών**, η **φτωχή διαίτα σε βιταμίνες, χόρτα και φρούτα**, το **κάπνισμα**, η **κατάχρηση οινοπνεύματος**, οι **ακτινοβολίες**, η **λήψη θηλυκών ορμονών**, η **ύπαρξη κυτταρικής δυσπλασίας** και ο **καρκίνος του άλλου μαστού**.

Πρέπει να τονισθεί ότι επιβαρυντικοί παράγοντες δεν σημαίνουν ότι όποια γυναίκα τους έχει, θα πάθει υποχρεωτικά κάποτε καρκίνο του μαστού, αλλά και αντίστροφα, δεν σημαίνει πως όποια γυναίκα δεν τους έχει, δεν πρόκειται ποτέ να πάθει καρκίνο. Αναλυτικότερα:

❖ Σε ποιες περιοχές του μαστού της γυναίκας εμφανίζεται συχνότερα ο καρκίνος;

Στην παρακάτω φωτογραφία φαίνεται η συχνότητα του καρκίνου του μαστού ανά τεταρτημόριο. Παρατηρήστε ότι σχεδόν οι μισές περιπτώσεις (45%) εμφανίζονται στο άνω έξω τεταρτημόριο και την ουρά του μαστού και ότι το 25% εμφανίζονται στην περιοχή της θηλής. Μόνο το 20% των καρκίνων προσβάλλουν το έσω ημιμόριο (άνω και κάτω έσω τεταρτημόρια)



Εικόνα 1.2: Συχνότητα του καρκίνου του μαστού ανά τεταρτημόριο.

❖ Σε ποιες ηλικίες εμφανίζεται πιο συχνά ο καρκίνος του μαστού;

Είναι λανθασμένη η αντίληψη ότι ο καρκίνος του μαστού εμφανίζεται πιο συχνά στις νέες γυναίκες. Όσο αυξάνεται η ηλικία τόσο αυξάνεται και η πιθανότητα. Το 80% συμβαίνει μετά την ηλικία των 50 ετών. Δεν υπάρχει καρκίνος του μαστού σε ηλικίες μικρότερες των 20 ετών και είναι εξαιρετικά σπάνιος μεταξύ 20 και 30 ετών. Στην πραγματικότητα η συχνότητά του διπλασιάζεται κάθε δέκα χρόνια και μέχρι την ηλικία των 75 ετών. Μετά αρχίζει ελαφρά να μειώνεται. Δηλαδή σε γυναίκες 70 ετών έχει διπλάσια συχνότητα σε σχέση με αυτές των 60 ετών και τετραπλάσια σε σχέση με γυναίκες των 50. Στη δεκαετία 50-65 ετών εμφανίζονται οι μισοί καρκίνοι του μαστού.

Θυμηθείτε: Στα 70 σας, έχετε 10πλάσια πιθανότητα να εμφανίσετε καρκίνο μαστού απ' όταν ήρασαν 40 ετών!



Εικόνα 1.3: Όσο αυξάνεται η ηλικία μέχρι τα 75 τόσο αυξάνεται και η πιθανότητα προσβολής από καρκίνο του μαστού.

❖ Ποιο ρόλο παίζει η κληρονομικότητα στον όγκο στο στήθος;

Είναι λανθασμένη η κοινή αντίληψη πως μόνο αν είχε καρκίνο του μαστού η μητέρα ή η γιαγιά μπορεί να πάθουν και τα παιδιά. Μόνο στο 30% των καρκίνων του μαστού υπάρχει κάποιας μορφής κληρονομική επιβάρυνση με βάση το οικογενειακό ιστορικό. Άρα οι 70% είναι τυχαίοι και εμφανίζονται σε γυναίκες χωρίς γνωστό οικογενειακό ιστορικό καρκίνου. Είναι όπως λέμε σποραδικοί καρκίνοι μαστού. Γι' αυτό πρέπει όλες οι γυναίκες μεταξύ 40 και 70 ετών κι όχι μόνο όσες είχαν μητέρα με ιστορικό καρκίνου του μαστού, να υποβάλλονται σε προληπτική κλινική εξέταση από ειδικό μαστολόγο και σε μαστογραφία.

Το 70% των καρκίνων μαστού είναι τυχαίοι, χωρίς ιστορικό. Από τους υπόλοιπους 30%, λιγότεροι από το 1/3 οφείλονται στα γνωστά γονίδια BRCA-1 ή BRCA-2. Η κληρονομικότητα των υπολοίπων υπάρχει αλλά παραμένει αδιευκρίνιστη.

Υπάρχουν οικογένειες, ελάχιστες ευτυχώς, που η νόσος εμφανίστηκε και συνεχίζει ακόμα να εμφανίζεται σε όλες τις γενιές. Οι γυναίκες που προέρχονται από τέτοιες οικογένειες απαιτούν ειδική αντιμετώπιση. Ας μην ξεχνάμε επίσης, ότι δεν έχουμε πληροφορίες για τις παλιότερες γενιές.



Εικόνα 1.4: Ποσοστό κληρονομικότητας που οφείλεται στον καρκίνο του μαστού.

- ❖ **Τι είναι το «γόνιμο χρονικό διάστημα» της ζωής μιας γυναίκας και ποιο ρόλο παίζει;**

Όταν λέμε «γόνιμο χρονικό διάστημα» της ζωής μια γυναίκας εννοούμε τη χρονική περίοδο της γυναίκας με έμμηνο ρύση. Έχει βρεθεί ότι όσο περισσότερα χρόνια κράτησε η περίοδος μιας γυναίκας, όσο δηλαδή περισσότερες ωορρηξίες είχε στη ζωή της, τόσο μεγαλύτερος είναι ο κίνδυνος ανάπτυξης καρκίνου στο μαστό. Άρα μια γυναίκα που η περιόδός της άρχισε στα 10 και τελείωσε στα 57, έχει ελάχιστα μεγαλύτερη πιθανότητα να πάθει καρκίνο μαστού, από μια άλλη που η περιόδός της άρχισε στα 15 και τελείωσε στα 45, εφ' όσον βέβαια οι υπόλοιποι παράγοντες κινδύνου είναι ισοτιμιοί και για τις δύο.

- ❖ **Παίζει ρόλο αν μια γυναίκα έχει παιδιά και σε ποια ηλικία γέννησε;**

Έχει βρεθεί ότι οι γυναίκες που γέννησαν το πρώτο παιδί τους πριν από την ηλικία των 28 ετών, άγνωστο γιατί, έχουν ελάχιστα μειωμένη πιθανότητα να πάθουν καρκίνο του μαστού. Αυτό είναι κάτι που πρέπει να γνωρίζουν όλες οι γυναίκες, αλλά δεν είναι εύκολο να συστήσουμε βιασύνη για γάμο αν δε συντρέχουν όλες οι αναγκαίες προϋποθέσεις. Αν όμως μια γυναίκα είναι παντρεμένη από μικρή ηλικία και θέλει να κάνει παιδιά, τότε μάλιστα, μπορούμε να την παροτρύνουμε.

- ❖ **Ο θηλασμός ποια θέση έχει, προστατεύει ή όχι;**

Πολλά έχουν γραφτεί γύρω από αυτό το θέμα και δε συμφωνούν όλοι στην ίδια άποψη. Οι περισσότερες στατιστικές παρατηρήσεις δείχνουν ότι ο θηλασμός για πάνω από 4 μήνες, μειώνει λίγο την πιθανότητα ανάπτυξης καρκίνου του μαστού, μερικοί υποστηρίζουν ότι δεν έχει καμία επίδραση και κανείς ποτέ δεν υποστήριξε ότι βλάπτει. Συνεπώς, είναι κάτι που ανεπιφύλακτα πρέπει να συνηθίζουμε στις μητέρες δεδομένου

ότι με βεβαιότητα κάνει καλό και στα μωρά τους. Στην Ελλάδα μόνο μια στις τρεις γυναίκες θηλάζει το μωρό της, ενώ σε άλλες Ευρωπαϊκές χώρες, παρ' ότι οι δεσμοί της οικογένειας δεν είναι τόσο στενοί, το ποσοστό θηλασμού φθάνει και το 60%! Η αιτία δεν είναι ότι οι Ελληνίδες μαμάδες αγαπούν και φροντίζουν τα παιδιά τους λιγότερο από τις Ευρωπαίες. Είναι ότι σήμερα οι περισσότερες νέες μητέρες πρέπει να εργάζονται. Οι περισσότερες, ιδιαίτερα στο σκληρό ιδιωτικό τομέα, παίρνουν λιγότερη άδεια λοχείας απ' όση δικαιούνται, γιατί φοβούνται μην απολυθούν.

❖ **Παίζει κάποιο ρόλο η περιοχή καταγωγής;**

Ναι. Υπάρχουν περιοχές της γης όπου ο καρκίνος του μαστού είναι συχνός (ΗΠΑ, Δ. Ευρώπη) και περιοχές που είναι πιο σπάνιος. Στις ΗΠΑ οι γυναίκες έχουν τριπλάσιο αριθμό καρκίνων του μαστού από τις Γιαπωνέζες. (Εξαιρούνται οι περιοχές της Χιροσίμα και του Ναγκασάκι που εμφάνισαν καρκίνο από τη ραδιενέργεια). Θα μπορούσαμε να πούμε ότι υπάρχουν πιο πολλά επιβαρυντικά γονίδια στην Αμερική, ή πιο πολλά “προφυλακτικά” γονίδια στην Ιαπωνία και αφού οι άνθρωποι παντρεύονται μεταξύ τους αυτά τα γονίδια ανακυκλώνονται στις ίδιες γεωγραφικές περιοχές. Όμως αυτό δεν εξηγεί γιατί οι Γιαπωνέζες όταν μεταναστεύουν στην Αμερική χωρίς να έχουν ζευγαρώσει με Αμερικάνους, αλλά πάλι με Ιάπωνες, παρουσιάζουν τέτοια αύξηση στους καρκίνους του μαστού που στην επόμενη γενιά φθάνει στα αμερικάνικα νούμερα. Μοναδική εξήγηση, είναι ότι φθάνοντας οι Γιαπωνέζες στην Αμερική αποκτούν αμερικάνικες διατροφικές συνήθειες και τρόπο ζωής, πράγμα που σημαίνει ότι ο κίνδυνος βρίσκεται και σε κάποιους εξωτερικούς παράγοντες και όχι μόνο στα γονίδια τους. Στην Ελλάδα παρατηρείται μεγάλη αύξηση της συχνότητας του καρκίνου του μαστού τα μεταπολεμικά χρόνια, που φθάνει το διπλασιασμό ανά 20ετία. Τα γονίδιά μας δεν άλλαξαν, ο τρόπος ζωής μας όμως...

❖ **Η παχυσαρκία θεωρείται επιβαρυντικός παράγων για τον καρκίνο του μαστού;**

Έχει παρατηρηθεί στατιστικά λίγο αυξημένη πιθανότητα θανάτου από καρκίνο του μαστού σε παχύσαρκες γυναίκες. Αυτό εξηγείται γιατί το λίπος περιέχει ένα ένζυμο που λέγεται αρωματάση και μετατρέπει κάποιες ορμόνες σε οιστρογόνα. Πιο πολύ το λίπος, πιο πολλή η αρωματάση, συνεπώς και πιο πολλή ορμόνη μετατρέπεται σε οιστρογόνα. Ορισμένοι όμως υποστηρίζουν πως αυτή η αυξημένη επιθετικότητα του καρκίνου οφείλεται απλά και μόνο σε καθυστερημένη διάγνωση και έναρξη θεραπείας, γιατί ο όγκος κρύβεται πιο εύκολα μέσα στο μαστό λόγω παχυσαρκίας.

❖ **Το κάπνισμα;**

Κάποιες έρευνες της δεκαετίας του '80 υποστήριξαν ότι το κάπνισμα ασκεί μια άγνωστη προστατευτική δράση, μειώνοντας λίγο την πιθανότητα ανάπτυξης καρκίνου του μαστού, στηριζόμενες στο ότι το κάπνισμα ελαττώνει τα οιστρογόνα και επιταχύνει την εμμηνόπαυση. Όλες οι νεότερες έρευνες όχι μόνο δεν επιβεβαιώνουν αυτήν την άποψη, αντιθέτως δείχνουν λίγο αυξημένη πιθανότητα καρκίνου του μαστού σε βαριές καπνίστριες για πολλά χρόνια. Ιδιαίτερα μάλιστα όταν πρόκειται για νέες γυναίκες πριν την εμμηνόπαυση. Εξ άλλου το τσιγάρο ευθύνεται επίσης και

για τόσα άλλα κακά, τόσους άλλους καρκίνους και βλάβες της καρδιάς και των αγγείων που αναμφίβολα πρέπει να αποθαρρύνεται η χρήση του.

❖ **Το οινόπνευμα;**

Τα τελευταία χρόνια κάποιες μελέτες αναφέρουν ότι οι γυναίκες που για χρόνια καταναλώνουν πάνω από 40 γρ. οινόπνεύματος την ημέρα έχουν αυξημένη πιθανότητα ανάπτυξης καρκίνου στο μαστό. Αυτό μάλλον οφείλεται στο ότι το οινόπνευμα, όπως και τα οιστρογόνα αποσυντίθενται στο ήπαρ (συκώτι). Το οινόπνευμα προκαλεί βλάβη στα ηπατικά κύτταρα και αυτά δυσκολεύονται να αποσυνθέσουν και να αποβάλλουν τα οιστρογόνα. Έτσι οι γυναίκες αυτές έχουν αυξημένα επίπεδα οιστρογόνων στο αίμα.

Πριν όμως απαγορεύσουμε τελείως το ποτό θα πρέπει να λάβουμε υπ' όψη μας το καλό που μπορεί να κάνει λίγο ποτό σε κάποιες γυναίκες από ψυχολογική άποψη, καθώς επίσης και τη γνώση μας πως ένα-δύο ποτήρια κόκκινου κρασιού κάνουν καλό στην καρδιά. Μια λογική συμβουλή θα ήταν να συστήσουμε στις γυναίκες που έχουν και άλλους επιβαρυντικούς παράγοντες να φυλάξουν τη διασκέδαση της οινοποσίας μόνο για ειδικές περιπτώσεις και απλά να αποφεύγουν την κατάχρηση του αλκοόλ.

❖ **Οι μαστογραφίες και οι ακτινογραφίες μπορεί να προκαλέσουν καρκίνο;**

Οι μοντέρνες ψηφιακές μαστογραφίες έχουν τόσο μικρή δόση ακτινοβολίας που δεν προκαλούν καρκίνο ακόμα και αν μια γυναίκα κάνει μια μαστογραφία το χρόνο για 30 χρόνια. Η μαστογραφία έχει λιγότερη ακτινοβολία από μια απλή ακτινογραφία θώρακος. Συνεπώς, ενώ το κέρδος από μια μαστογραφία είναι τεράστιο, γιατί η έγκαιρη διάγνωση και θεραπεία μπορεί ακόμα και μια ζωή να γλιτώσει, η δόση ακτινοβολίας που δίνει είναι μικρή.

Όλες οι εξετάσεις που ακτινοβολούν τον άνθρωπο με ακτίνες Χ (αξονικές τομογραφίες, ακτινογραφίες, πνευμογραφίες, αγγειογραφίες, στεφανιογραφίες κλπ), πρέπει να χρησιμοποιούνται όταν πραγματικά χρειάζονται, να συστήνονται πάντα με περίσκεψη από τους γιατρούς και να μη γίνεται κατάχρησή τους ιδιαίτερα σε νέα άτομα.

❖ **Υπάρχουν κάποιες ουσίες που προστατεύουν από τον καρκίνο του μαστού ;**

Πολλές ερευνητικές εργασίες θεωρούν προστατευτικές ουσίες τις βιταμίνες D,C και E. Με την έννοια όμως ότι μειώνουν κάπως την πιθανότητα προσβολής από τη νόσο, χωρίς να εξασφαλίζουν την ανοσία. Δεν υπάρχει φάρμακο που να το παίρνουν υγιείς γυναίκες και να μην παθαίνουν ποτέ καρκίνο του μαστού. Αυτό θα ήταν το όνειρο όλων. Υπάρχει όμως η ταμοξιφαίνη. Αυτό το φάρμακο χρησιμοποιείται κυρίως σαν επικουρική θεραπεία για καρκίνο του μαστού μετά τη χειρουργική εκτομή και έχει βρεθεί ότι ελαττώνει τις μεταστάσεις και αυξάνει την επιβίωση σε μετεμμηνοπαυσιακές γυναίκες κατά 30%. Έχει χρησιμοποιηθεί προληπτικά σε υγιείς γυναίκες υψηλού κινδύνου με κληρονομικό καρκίνο έχοντας μειώσει την εμφάνιση καρκίνου του μαστού κατά 50%. Δεν χορηγείται προληπτικά στον πληθυσμό γιατί προκαλεί εμμηνόπαυση, μπορεί να έχει παρενέργειες και τελικά να κάνει κακό αντί για καλό.

- ❖ **Αν παρατηρήθηκε κακοήθης όγκος στον έναν μαστό, τι κίνδυνος υπάρχει να εντοπιστεί και στον άλλο;**

Οι γυναίκες που εμφάνισαν καρκίνο του ενός μαστού, έχουν λίγο αυξημένη πιθανότητα να εμφανίσουν και στον άλλο. Η πιθανότητα είναι 0,5-1% ανά έτος, δηλαδή 15-30% στα 30 χρόνια. Στο 2% των γυναικών, τη μέρα της διάγνωσης του καρκίνου στον ένα μαστό, θα βρεθεί ταυτόχρονα καρκίνος και στον άλλο.

1.2.2 Είδη καρκίνου του μαστού

Υπάρχουν πολλά είδη καρκίνων του μαστού. Εξαρτάται πώς τους κατατάσσουμε. Ανάλογα με το πόσο βαθιά έχει προχωρήσει ένας καρκίνος μέσα στο μαστό, χωρίζεται σε *in situ* και σε διηθητικό. Ανάλογα με τα κύτταρα από τα οποία προέρχεται, χωρίζεται σε πορογενή και σε λοβιακό. Ανάλογα με το πόσο κακοήθης είναι τα κύτταρά του, χωρίζεται σε βαθμό κακοήθειας 1 (χαμηλό), 2 (μέτριο) ή 3 (υψηλό). Ανάλογα με το πόσο διαφοροποιημένα είναι τα κύτταρά του χωρίζεται σε χαμηλής, μέσης και υψηλής διαφοροποίησης. Επίσης, χωρίζονται σε δύο ομάδες, ανάλογα με το αν τα κύτταρα έχουν θετικούς ορμονικούς υποδοχείς οιστρογόνων και προγεστερόνης, αν έχουν αρνητικούς ορμονικούς υποδοχείς. Χωρίζονται επίσης, σε δύο ομάδες, ανάλογα με το αν έχουν υποδοχείς HER2 ή όχι, δηλαδή HER-2θετικοί και HER-2 αρνητικοί.

Στην παρούσα ενότητα παρουσιάζουμε γενικές πληροφορίες σχετικά με τα εξής είδη καρκίνου του μαστού:

- Διηθητικός πορογενής & λοβιακός καρκίνος
- *In situ* πορογενής & λοβιακός καρκίνος
- Σάρκωμα
- Νόσος Paget της θηλής
- Φλεγμονώδης καρκίνος μαστού

Διηθητικός καρκίνος, πορογενής και λοβιακός

Τι είναι ο πορογενής διηθητικός καρκίνος του μαστού;

Είναι ο καρκίνος που αρχίζει από κύτταρα των γαλακτοφόρων πόρων, αυτών δηλαδή που μεταφέρουν το γάλα από τα λόβια προς τη θηλή. Αποτελεί το 80% του συνόλου των καρκινωμάτων του μαστού, πρόκειται δηλαδή για τη συχνότερη μορφή και είναι συνήθως μονήρης. Είναι πιθανό να υπάρχει και δεύτερη εστία στον ίδιο ή και στον άλλο μαστό, γι' αυτό και απαιτείται πάντοτε πολύ καλός έλεγχος πριν το χειρουργείο, εάν πρόκειται να γίνει ογκεκτομή και όχι μαστεκτομή. Μπορεί να εκδηλωθεί με όλα τα πιθανά συμπτώματα ανάλογα με την περίπτωση: ψηλαφητό όγκο, έκκριση από τη θηλή, εισολκή του δέρματος, ψηλαφητούς λεμφαδένες μασχάλης, πόνο κλπ. Είναι συνήθως ορατός στη μαστογραφία, εκτός εάν πρόκειται για πολύ πυκνούς μαστούς, οπότε η πιθανότητα να χαθεί και να μείνει αδιάγνωστος είναι μεγάλη. Μπορεί να δώσει μεταστάσεις σε άλλα όργανα. Η επιθετικότητα και ο βαθμός κακοήθειάς του εμφανίζουν πολύ μεγάλη ποικιλία ανάλογα με την περίπτωση.

Τι είναι ο λοβιακός διηθητικός καρκίνος του μαστού;

Είναι ο καρκίνος που αρχίζει από τους λοβιακούς αδένες που παράγουν το γάλα. Αποτελεί το 10% του συνόλου των καρκινωμάτων του μαστού. Στο 20% των περιπτώσεων, μπορεί να βρεθεί και στον άλλο μαστό (αμφοτερόπλευρος καρκίνος μαστού). Εκδηλώνεται συνήθως με ψηλαφητό όγκο στο μαστό. Συνήθως είναι ορατός στη μαστογραφία. Ενίοτε επιφυλάσσει εκπλήξεις και μπορεί να μεγαλώνει εντελώς αόρατος στη μαστογραφία, στους υπερήχους και στην ψηλάφηση και να φανεί όταν διηθήσει το δέρμα! Δεν προκαλεί έκκριση υγρού από τη θηλή.



Εικόνα 1.5: Από τα αδενικά λόβια (1) που παράγουν το γάλα προέρχεται ο λοβιακός καρκίνος. Από τους πόρους που μεταφέρουν το γάλα προέρχεται ο πορογενής καρκίνος (2).

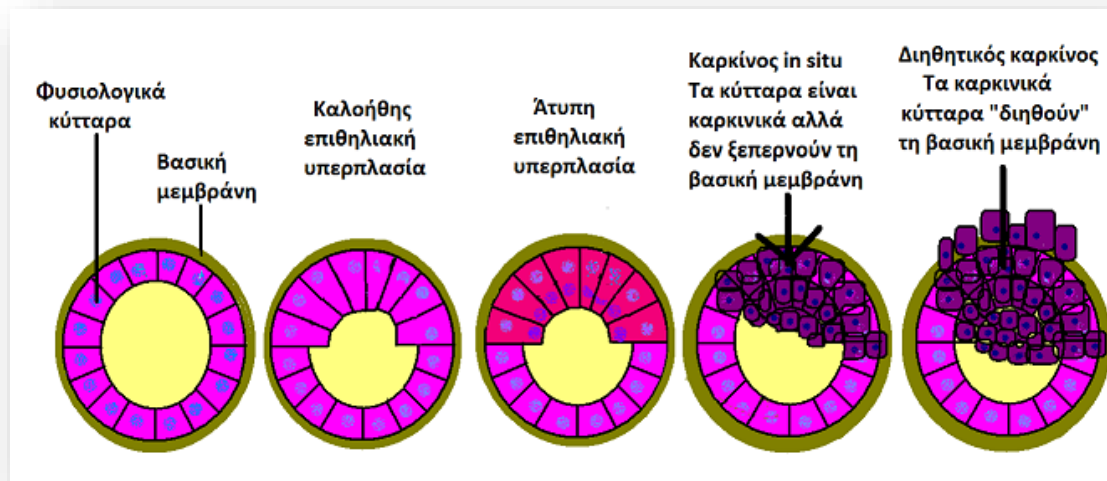
In situ καρκίνος, λοβιακός, πορογενής

In situ και διηθητικός καρκίνος

Συνήθως προηγείται του καρκίνου μια υπερπλασία κάποιων κυττάρων. Αργότερα τα κύτταρα αυτά εμφανίζουν ατυπία (προκαρκινικά) και πιο μετά γίνονται καρκινικά (κακοήθη). Αυτά μεγαλώνουν κλεισμένα μέσα σε μια μεμβράνη που λέγεται βασική μεμβράνη και τα απομονώνει από τα υπόλοιπα κύτταρα του μαστού. Στη φάση αυτή ο καρκίνος θεωρείται εντοπισμένη νόσος, δεν κάνει μεταστάσεις σε άλλα όργανα και άρα είναι ιάσιμος με απλή χειρουργική αφαίρεση. Αυτή η μορφή του ονομάζεται καρκίνος in situ και είναι αρχομένη νόσος.

Μεγαλώνοντας σιγά σιγά, ο καρκίνος διασπά τη βασική αυτή μεμβράνη και διεισδύει (διηθεί όπως λέμε) στα γειτονικά κύτταρα, τα αιμοφόρα και λεμφοφόρα

αγγεία. Τότε ονομάζεται διηθητικός. (Invasive ή Infiltrative) Ο διηθητικός καρκίνος είναι πιο προχωρημένη νόσος με την έννοια ότι τα κύτταρά του μπορεί να μπουν στην κυκλοφορία του αίματος και άρα μπορεί να γεννηθούν μεταστάσεις σε άλλα όργανα, χωρίς όμως αυτό να είναι υποχρεωτικό. Δηλαδή, διαβάζοντας σε ένα χαρτί ιστολογικής εξέτασης κάποιου νοσοκομείου τη λέξη «διηθητικό καρκίνωμα», αυτό δε σημαίνει αυτομάτως και ότι έχουν γίνει μεταστάσεις αλλού. Σημαίνει μόνο ότι κάτι τέτοιο είναι δυνατό να γίνει, σε αντίθεση με το in situ όπου κάτι τέτοιο είναι αδύνατο.



Εικόνα 1.6: In situ και διηθητικός καρκίνος.

Τι είναι ο λοβιακός in situ καρκίνος;

Γράφεται διεθνώς ως LCIS. Είναι η λιγότερο επιθετική μορφή καρκίνου του μαστού, γιατί έχει πιθανότητα 20% στα 15 χρόνια να μετατραπεί σε διηθητικό καρκίνο μαστού. Είναι αόρατος στη μαστογραφία και συνήθως βρίσκεται σε τμήμα μαστού που αφαιρέθηκε για άλλο λόγο (βιοψία, αφαίρεση κάποιου άλλου ογκιδίου, πλαστική μείωσης μαστών κλπ). Στο 80% των περιπτώσεων είναι πολυεστιακός, δηλαδή βρίσκεται και σε άλλα σημεία του ίδιου μαστού. Στο 25% βρίσκεται και στον άλλο μαστό. Πρόκειται για πάθηση που για την αντιμετώπισή της δεν υπάρχει ομοφωνία. Πολλές λύσεις έχουν προταθεί ανάλογα με την ιδιοσυγκρασία της κάθε ασθενούς αλλά και της ιατρικής ομάδας που θα την αντιμετωπίσει.

Με την πάθηση αυτή συμβαίνει το εξής παράδοξο: Ενώ στον επιθετικό διηθητικό καρκίνο του μαστού προσπαθούμε σήμερα να διασώσουμε το μαστό και 2 στις 3 φορές το πετυχαίνουμε, με απλή ογκεκτομή, στο «καλοηθέστερο κακόηθες» ογκίδιο του μαστού, το LCIS, κάποιος γιατρός μπορεί να συστήσουν ακόμη και υποδόρια μαστεκτομή λόγω της αυξημένης πολυεστιακότητας! Σήμερα, υπάρχει μια τάση, να καταταγεί το λοβιακό in situ καρκίνωμα στους παράγοντες αυξημένου κινδύνου για ανάπτυξη καρκίνου μαστού και όχι στα καρκινώματα.

Τι είναι ο πορογενής in situ καρκίνος;

Γράφεται διεθνώς ως DCIS. Όπως αναφέραμε στην αρχή του κεφαλαίου, “in situ” σημαίνει πρόδρομη μορφή καρκίνου που μεγαλώνει τοπικά. Δε δίνει μεταστάσεις, εκτός κι αν εξαλλαγεί σε διηθητικό καρκίνο, πράγμα που έχει πιθανότητα 30% να συμβεί στα 10 χρόνια. Φαίνεται στη μαστογραφία επειδή συνήθως περιέχει μικροαποτιτανώσεις, οι οποίες όμως δεν δείχνουν σαφώς το μέγεθός του, γιατί ο όγκος μπορεί να εκτείνεται και πιο πέρα από αυτές. Στο 30% των περιπτώσεων βρίσκεται και σε άλλα σημεία του ίδιου μαστού και στο 5% μπορεί να βρεθεί και στον άλλο μαστό.

Σάρκωμα μαστού, ένας σπάνιος, επιθετικός όγκος

Τι είναι το σάρκωμα;

Πρόκειται για μια σπάνια ξεχωριστή κατηγορία κακοήθους μορφής, που προέρχεται από τα κύτταρα των ινών, των αγγείων ή του λίπους του μαστού. Αποτελεί λιγότερο από το 1% των κακοήθων παθήσεων του μαστού. Δεν μεθίσταται στους λεμφαδένες της μασχάλης, γι’ αυτό και αντιμετωπίζεται με ευρεία ογκεκτομή ή απλή μαστεκτομή, δηλαδή χωρίς αφαίρεση των λεμφαδένων της μασχάλης. Το σάρκωμα προκαλεί συνήθως, αν δεν το προλάβουμε, αιματογενείς μεταστάσεις. Σε σάρκωμα (φυλλοειδές κυστεοσάρκωμα) μπορεί να μεταπέσει ο φυλλοειδής όγκος, που αρχικά μοιάζει με μεγάλο ινοαδένωμα χωρίς να είναι.

Νόσος Paget της θηλής

Τι είναι η νόσος Paget της θηλής;

Πρόκειται για σπάνια μορφή πορογενούς διηθητικού καρκινώματος των τελικών γαλακτοφόρων πόρων της θηλής. Διηθείται η θηλή του μαστού γιατί ο όγκος βρίσκεται σχεδόν από κάτω της. Προκαλείται έτσι μια κοκκινίλα της θηλής και της θηλαίας άλω, που αρχικά μπορεί να μοιάζει με έκζεμα, γι’ αυτό και πολλές γυναίκες πηγαίνουν πρώτα στο δερματολόγο. Με την πρόοδο της νόσου η θηλή αρχίζει να αιμορραγεί. Αποτελεί λιγότερο από το 1% των κακοηθειών του μαστού.



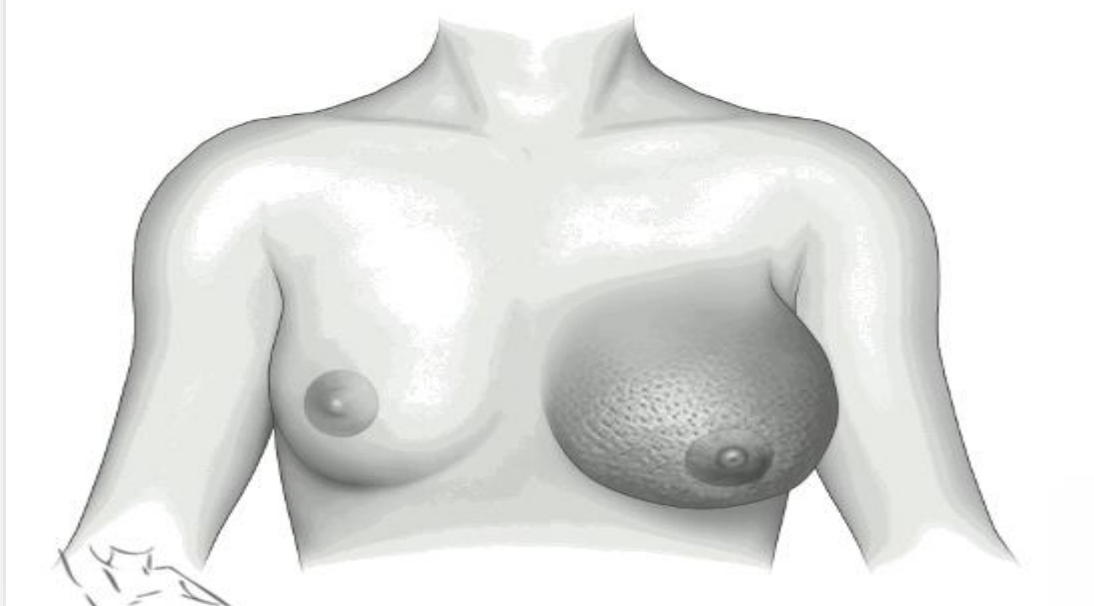
Εικόνα 1.7: Νόσος Paget της θηλής.

Φλεγμονώδης καρκίνος του μαστού

Τι είναι ο φλεγμονώδης καρκίνος μαστού;

Ο φλεγμονώδης καρκίνος μαστού είναι μια ειδική κατηγορία καρκίνου. Συνοδεύεται από μικροβιακή ή άσηπτη φλεγμονή των λεμφαγγείων μιας περιοχής του μαστού. Επειδή το δέρμα της πάσχουσας περιοχής εμφανίζει σαφή ερυθρότητα, μπορεί εύκολα να θεωρήσει κανείς πως πρόκειται για απλή φλεγμονή και να χορηγήσει αντιβίωση. Συνήθως, η ερυθρότητα στο φλεγμονώδη καρκίνο είναι διάχυτη και καταλαμβάνει μεγάλο μέρος της επιφάνειας του δέρματος, ακόμα και το 1/3. Αντίθετα, στη φλεγμονή και το απόστημα του μαστού η ερυθρότητα συνήθως εστιάζεται ιδιαίτερα σε ένα σημείο του δέρματος που είναι το κέντρο της φλεγμονής και όσο απομακρυνόμαστε από εκεί αραιώνει. Συνήθως, με την εμπειρία μπορεί κανείς να τον ξεχωρίσει από μια απλή φλεγμονή. Αντιμετωπίζουμε όμως περιπτώσεις όπου δε μπορούμε εξ αρχής να ξεχωρίσουμε αν κάτω από τη φλεγμονή υπάρχει και όγκος. Ακόμα και με μαγνητική μαστογραφία, πάλι μπορεί να έχουμε αμφιβολίες. Τότε, χορηγούμε αντιβιοτικά μήπως υποχωρήσει η φλεγμονή και επανεξετάζουμε σε λίγες μέρες. Ο φλεγμονώδης καρκίνος κατατάσσεται εξ αρχής στο 3ο στάδιο καρκίνου μαστού.

Φλεγμονώδης Καρκίνος Μαστού



Εικόνα 1.8: Φλεγμονώδης καρκίνος του μαστού.

1.2.3 Συμπτώματα

Στα πιο χαρακτηριστικά συμπτώματα του καρκίνου του μαστού που καλείται να αντιμετωπίσει ή να διαγνώσει συνήθως ένας εξειδικευμένος μαστολόγος συγκαταλέγονται τα εξής:

- **Ένας αισθητός όγκος στο μαστό ή στη μασχάλη:** Αν ψηλαφίσετε το στήθος σας και νιώθετε ένα εξόγκωμα σαν μικρό μπιζέλι ή λίγο μεγαλύτερο, να βρίσκεται στα πλάγια του μαστού κοντά στον αδένα και παραμένει ακόμα και μετά το πέρας της περιόδου, είναι καλό να επισκεφθείτε τον γιατρό σας.
- **Πρήξιμο και διόγκωση του μαστού ή του ιστού:** Το στήθος της γυναίκας, συνήθως πρήζεται φυσιολογικά κατά την περίοδο και διογκώνεται προκαλώντας ενόχληση και πόνο ειδικά κοντά στους αδένες, ωστόσο αν δείτε ότι αυτή η κατάσταση συνεχίζεται και είναι εξίσου ενοχλητική και επώδυνη, ακόμα κι όταν η περίοδος σας έχει τελειώσει, τότε το σύμπτωμα αυτό θεωρείται μάλλον ανησυχητικό.
- **Αισθητή αλλαγή στο σχήμα, το περίγραμμα ή την υφή του στήθους:** Αν έχετε ξεπεράσει την ηλικία ανάπτυξης του στήθους σας και παρ' όλα αυτά παρατηρήσετε αισθητή αλλαγή είτε στο μέγεθος και το σχήμα μην το αφήσετε χωρίς εξέταση.
- **Διαφανές ή κόκκινο υγρό από τη θηλή:** Η έκκριση υγρού από τη θηλή είναι ανησυχητική. Προσοχή όμως, αναφερόμαστε σε ροή που γίνεται από μόνη της χωρίς να ασκήσουμε εμείς πίεση στο στήθος.

- **Ερυθρότητα στην περιοχή του μαστού ή της θηλής:** Γενικά, αν παρατηρήσετε μια περιοχή στο μαστό σας που φαίνεται σαφώς διαφορετική από το συνηθισμένο και παρουσιάζει έντονη ερυθρότητα και ενόχληση, φολιδωτό δέρμα ή ξηροδερμία, θα πρέπει να το δει ένας εξειδικευμένος μαστολόγος.
- **Μία σκληρή περιοχή σαν μάρμαρο κάτω από το δέρμα:** Τέτοιου είδους εξογκώματα είναι αρκετά ανησυχητικά και συχνά ενδεικτικά για την ύπαρξη δυνητικά επικίνδυνου όγκου στον μαστό.

1.2.4 Στάδια της νόσου

Η σωστή σταδιοποίηση του καρκίνου του μαστού είναι εξαιρετικής σημασίας. Όχι μόνο επιτρέπει την ακριβή πρόγνωση αλλά επίσης, σε πολλές επιπτώσεις, η απόφαση για τη θεραπεία βασίζεται κατά πολύ στην ταξινόμηση κατά TNM. Έτσι, τα στάδια του καρκίνου του μαστού καθορίζονται σύμφωνα με την ταξινόμηση κατά TNM (American Joint Committee, 1992), κατά την οποία ο καρκίνος του μαστού σταδιοποιείται ανάλογα με το **μέγεθος του πρωτοπαθούς όγκου (T)**, την **προσβολή των σύστοιχων μασχαλιαίων λεμφαδένων (N)** και την **παρουσία ή όχι μεταστάσεων (M)**.

Υπάρχουν λοιπόν, τέσσερα στάδια.

Το Στάδιο 0 είναι ο καρκίνος *in situ* που εντοπίζεται στο σημείο προέλευσής του και δεν έχει αρχίσει να διηθεί τους γύρω ιστούς ή να διασπείρεται και περιλαμβάνει τα *in situ* λοβώδη καρκινώματα και τα *in situ* καρκινώματα των εκφορητικών πόρων του μαστού.

- ✚ **Στο Στάδιο I (T1N0M0)** ο όγκος είναι μεγέθους 2cm ή μικρότερος και δεν υπάρχει ένδειξη διασποράς στους λεμφαδένες ή σε απομακρυσμένα σημεία.
- ✚ **Το Στάδιο II Α** περιλαμβάνει τις κατηγορίες T0N1M0, όπου δεν υπάρχει ένδειξη πρωτοπαθούς όγκου αλλά διακρίνεται μετάσταση στους μασχαλιαίους λεμφαδένες, T2N0M0, με όγκο μεγαλύτερο από 2cm αλλά όχι μεγαλύτερο από 5cm και χωρίς διασπορά στους λεμφαδένες.
Στο Στάδιο II Β περιλαμβάνονται οι κατηγορίες T2N1M0, με όγκο 2-5cm και με θετικούς λεμφαδένες και T3N0M0, με όγκο μεγαλύτερο από 5cm χωρίς λεμφαδενικές μεταστάσεις.
- ✚ **Το Στάδιο III Α**, στη συνέχεια, αποτελείται από τις κατηγορίες T0N2M0, T1N2M0, T2N2M0, T3N1M0 και T3N2M0, κατά τις οποίες ο όγκος είναι μεγαλύτερος από 5cm και διακρίνεται προσβολή των σύστοιχων μασχαλιαίων λεμφαδένων που συμφύονται μεταξύ τους ή με άλλους ιστούς.
Στο Στάδιο III Β περιλαμβάνονται οι κατηγορίες T4 anyN M0 και anyT N3 M0, όπου υπάρχουν προσβεβλημένοι θωρακικοί λεμφαδένες και ο όγκος επεκτείνεται στο θωρακικό τοίχωμα ή προσβάλλει και εξελκώνει το δέρμα.
- ✚ Τέλος, **το Στάδιο IV** περιλαμβάνει κατηγορίες με οποιοδήποτε μέγεθος όγκου, με διασπορά στους λεμφαδένες, αλλά και με απομακρυσμένες μεταστάσεις ή προσβολή του δέρματος και του θωρακικού τοιχώματος πέρα από την περιοχή του μαστού.

1.2.5 Διαγνωστική διερεύνηση της νόσου

ΜΑΣΤΟΓΡΑΦΙΑ

Η **μαστογραφία** αποτελεί έναν εξαιρετικά πολύτιμο σύμμαχο ενάντια στον καρκίνο του μαστού. Την πιο σημαντική σήμερα εξέταση για την εκτίμηση ενός προβλήματος στο μαστό καθώς και την κύρια εξέταση του προληπτικού ελέγχου. Η μαστογραφία μπορεί να διαγνώσει ένα ογκίδιο μαστού πολύ πριν αυτό γίνει ψηλαφητό. Όταν μία γυναίκα ψηλαφά ένα ογκίδιο στο μαστό της αυτό είναι κατά μέσο όρο περίπου 2 εκατοστά ενώ η μαστογραφία μπορεί να διαγνώσει ογκίδια μόλις λίγων χιλιοστών. Η ποιότητα της μαστογραφίας έχει εξαιρετική σημασία στη σωστή αποτύπωση και διάγνωση ενός προβλήματος στο μαστό. Θα πρέπει να γίνεται σωστά και να ερμηνεύεται από ειδικό ακτινολόγο μαστού με εμπειρία!

Η γυναίκα που δεν έχει βεβαρυμένο ιστορικό θα πρέπει να κάνει την πρώτη της μαστογραφία (μαστογραφία αναφοράς), σε ηλικία 35 - 40 ετών ώστε να υπάρχει σύγκριση με τις μαστογραφίες που θα διενεργούνται στο μέλλον και πρέπει να γίνεται κάθε χρόνο μετά την ηλικία των 40. Σήμερα, η σύσταση είναι μια γυναίκα να συνεχίζει να κάνει ετήσιες μαστογραφίες και μετά τα εβδομήντα χρόνια, εφόσον είναι σε καλή γενική κατάσταση και μπορεί να επισκεφθεί το κέντρο της επιλογής της για να κάνει τη μαστογραφία της.

Εάν μία γυναίκα έχει βεβαρημένο οικογενειακό ιστορικό θα πρέπει να ξεκινήσει τον τακτικό, κλινικό, μαστογραφικό έλεγχο τουλάχιστον δέκα χρόνια νωρίτερα, στην ηλικία των 30 ετών και θα πρέπει να υποβάλλεται και σε μαστογραφία και σε υπερηχογράφημα μαστών καθώς και σε Μαγνητική Μαστών κατά περίπτωση.

Με τους σύγχρονους ψηφιακούς μαστογράφους οι γυναίκες δέχονται 50% με 70% λιγότερη ακτινοβολία συγκριτικά με παλαιότερης τεχνολογίας μαστογράφους. Η ψηφιακή μαστογραφία ενδείκνυται ιδιαίτερα σε προεμμηνοπαυσιακές αλλά και μεγαλύτερες γυναίκες με πυκνούς μαστούς. Έχει μεγαλύτερη διαγνωστική ακρίβεια και αναδεικνύει καλύτερα τις πιθανές αποτιτανώσεις στο μαστό.

Για τη μαστογραφία θα πρέπει να γνωρίζουμε ότι ένα ποσοστό 8% - 12% περίπου των καρκίνων του μαστού, παρά τις άριστες συνθήκες τεχνικής και μηχανήματος, δεν απεικονίζονται. Γι' αυτό και θα πρέπει να συμπληρώνεται με κλινική εξέταση από εξειδικευμένο γιατρό σε συνδυασμό με υπερηχογράφημα μαστών, ιδιαίτερα αν οι γυναίκες έχουν πυκνούς μαστούς. Με το τρίπτυχο -κλινική εξέταση, μαστογραφία, υπερηχογράφημα μαστών -επιτυγχάνεται η σωστή διάγνωση κατά 98% με 99%.

Ένα ύποπτο εύρημα στην ψηφιακή μαστογραφία θα πρέπει να επανελέγχεται επιτόπου με περαιτέρω διαγνωστική μαστογραφία, η οποία περιλαμβάνει περισσότερες εικόνες της περιοχής προς διερεύνηση. Στην περίπτωση που διαφανεί πως ενδέχεται να είναι καρκίνος, τότε θα χρειαστεί βιοψία.

ΥΠΕΡΗΧΟΓΡΑΦΗΜΑ ΜΑΣΤΩΝ

Πρόκειται για ένα ανώδυνο τεστ που διενεργείται με τη χρήση ηχητικών κυμάτων που σκιαγραφούν ένα μέρος του σώματος. Το **υπερηχογράφημα μαστών** είναι ένα εξαιρετικά χρήσιμο διαγνωστικό εργαλείο σε συνδυασμό με τη μαστογραφία, δεν προκαλεί πόνο και δεν εκθέτει το άτομο σε ακτινοβολία. Συνήθως, το υπερηχογράφημα μαστών χρησιμοποιείται επικουρικά για να διερευνηθούν πυκνοί μαστοί για τυχόν ευρήματα αλλά και γυναίκες με οικογενειακό ιστορικό καρκίνου, ενώ βοηθά ώστε να διαχωριστούν οι κύστες από τα συμπαγή ογκίδια και ορισμένες φορές για να διαφοροποιηθεί ένας καλοήθης όγκος από έναν κακοήγη. Σε άτομα με όγκο στο μαστό, το υπερηχογράφημα μπορεί να αποτυπώσει τυχόν διόγκωση των μασχαλιαίων λεμφαδένων. Ωστόσο, σε καμία περίπτωση δεν μπορεί το υπερηχογράφημα μαστών να υποκαταστήσει τη μαστογραφία.

ΜΑΓΝΗΤΙΚΗ ΤΟΜΟΓΡΑΦΙΑ

Η **μαγνητική τομογραφία** μπορεί να χρησιμοποιηθεί σε συνδυασμό με τη μαστογραφία για την απεικόνιση των μαστών των γυναικών που εμφανίζουν υψηλό κίνδυνο ανάπτυξης καρκίνου του μαστού ή μπορεί να βοηθήσει στην περαιτέρω διερεύνηση ύποπτων περιοχών στο μαστό για ύπαρξη ευρημάτων. Επίσης, η ψηφιακή μαστογραφία επιλέγεται σε γυναίκες με διαγνωσμένο καρκίνο του μαστού για να διασαφηνίσει το ακριβές μέγεθος του καρκίνου και για να εντοπίσει τυχόν άλλους καρκίνους στο μαστό ή για να εξακριβώσει εάν υπάρχουν όγκοι στον άλλο μαστό.

ΝΕΕΣ ΔΙΑΓΝΩΣΤΙΚΕΣ ΔΙΑΔΕΡΜΙΚΕΣ ΒΙΟΨΙΕΣ (ΜΙΚΡΟ-ΕΠΕΜΒΑΤΙΚΕΣ)

Σύγχρονες διαγνωστικές τεχνικές όπως η παρακέντηση δια **λεπτής βελόνης** (FNA) και η **διαδερμική βιοψία κόπτουσας βελόνης** (Core biopsy), μας δίνουν τη δυνατότητα λήψης ιστού από την βλάβη είτε με την μορφή κυττάρων ή με την μορφή ιστού. Στις περιπτώσεις που η βλάβη είναι μικρή και δεν είναι ψηλαφητή, η λήψη γίνεται με υπερηχογραφική καθοδήγηση. Επίσης, τεχνικές στερεοτακτικής βιοψίας με τη βοήθεια του μαστογράφου μας επιτρέπουν να παίρνουμε ιστό από περιοχές ύποπτες, μη ψηλαφητές καθώς και από περιοχές με αποιτανώσεις στο μαστό για ιστολογική διάγνωση. Σήμερα, διαγνωστικές μέθοδοι όπως η βιοψία με τη χρήση του mammotome, μας επιτρέπουν να έχουμε ιστολογική επιβεβαίωση μιας καλοήθους βλάβης αποφεύγοντας σε αρκετές περιπτώσεις μη αναγκαίες χειρουργικές επεμβάσεις, αλλά μας επιτρέπουν και σε άλλες περιπτώσεις να έχουμε προεγχειρητική διάγνωση ώστε να προγραμματίσουμε σωστά ένα ογκολογικό χειρουργείο.

1.2.6 Θεραπευτική αγωγή του καρκίνου του μαστού

Τα τελευταία χρόνια έχουν αναπτυχθεί πολλοί και αποτελεσματικοί τρόποι θεραπείας για τον καρκίνο του μαστού. Αυτό συμβαίνει ώστε να αντιμετωπίζεται με επιτυχία κάθε ιδιαίτερη μορφή καρκίνου. Η επιλογή της κατάλληλης θεραπείας για

κάθε περίπτωση είναι πολύπλοκη καθώς υπάρχουν πολλοί παράγοντες που πρέπει να ληφθούν υπόψη. Παρακάτω αναφέρονται οι 5 τύποι θεραπείας που έχουν καθιερωθεί σήμερα:

Χειρουργική θεραπεία.

Στις περισσότερες περιπτώσεις η πρώτη θεραπεία που εφαρμόζεται σε περίπτωση καρκίνου του μαστού είναι η χειρουργική επέμβαση για την αφαίρεση του όγκου. Η απόφαση για το είδος της επέμβασης εξαρτάται από πολλούς παράγοντες όπως: Το στάδιο του καρκίνου, το είδος του καρκίνου, το ορμονικό και γενετικό «προφίλ» του, τις επιθυμίες της γυναίκας.

Οι τύποι των χειρουργικών επεμβάσεων που εφαρμόζονται είναι οι ακόλουθοι:

- **Ουγκεκτομή ή μερική μαστεκτομή.** Είναι η πιο συντηρητική μορφή επέμβασης στην οποία αφαιρείται ο όγκος και υγιές τμήμα του ιστού γύρω του αλλά όχι ολόκληρος ο μαστός.
- **Μαστεκτομή.** Σε αυτήν την περίπτωση αφαιρείται ολόκληρος ο μαστός στον οποίο βρίσκεται ο καρκίνος. Επίσης, είναι δυνατό να αφαιρεθούν και μερικοί μασχαλιαίοι λεμφαδένες ώστε να ελεγχθούν για τυχόν επέκταση του καρκίνου.
- **Ριζική μαστεκτομή.** Στην επέμβαση αυτή αφαιρείται όλος ο μαστός που έχει τον καρκίνο, πολλοί μασχαλιαίοι λεμφαδένες, τμήμα του θωρακικού τοιχώματος και μερικές φορές τμήμα των θωρακικών μυών.
- **Λεμφαδενεκτομή.** Είναι η αφαίρεση επιπλέον λεμφαδένων σε γειτονικές με τον όγκο περιοχές εάν υπάρχει υποψία επέκτασης του καρκίνου.
- **Βιοψία φρουρού λεμφαδένα.** Είναι η αφαίρεση του πρώτου λεμφαδένα που δέχεται την λέμφο από τον όγκο του μαστού. Είναι ο πρώτος λεμφαδένας στον οποίο είναι πιθανό να κάνει μετάσταση ο καρκίνος. Στην περιοχή του όγκου χορηγείται με ένεση με ένα ραδιοϊσότοπο ή μία χρωστική ουσία. Μέσα από τα λεμφαγγεία μεταφέρονται στους λεμφαδένες. Ο πρώτος λεμφαδένας που δέχεται το ραδιοϊσότοπο ή την χρωστική αφαιρείται και εξετάζεται στο μικροσκόπιο. Εάν δεν βρεθούν καρκινικά κύτταρα ίσως να μην χρειαστεί να αφαιρεθούν περισσότεροι λεμφαδένες. Αντίθετα εάν βρεθούν καρκινικά κύτταρα ίσως πρέπει να γίνει πιο εκτεταμένη λεμφαδενεκτομή.

Προεγχειρητική χημειοθεραπεία. Μερικές φορές είναι προτιμότερο πριν την επέμβαση να προηγηθεί χημειοθεραπεία. Σκοπός είναι να συρρικνωθεί ο όγκος και να αφαιρεθεί μικρότερη ποσότητα ιστού κατά την επακόλουθη επέμβαση.

Συμπληρωματική θεραπεία. Ακόμη και εάν στην επέμβαση αφαιρεθεί όλος ο όγκος σε μερικές περιπτώσεις συμπληρωματική θεραπεία είναι απαραίτητη. Περιλαμβάνει την ακτινοθεραπεία, την χημειοθεραπεία και την ορμονοθεραπεία. Στόχος είναι να θανατωθούν όσα καρκινικά κύτταρα πιθανό να έχουν απομείνει ώστε να μειωθούν οι πιθανότητες υποτροπής του καρκίνου.

Πλαστική αποκατάσταση μαστού μπορεί να γίνει μετά από μαστεκτομή. Μπορεί να γίνει ταυτόχρονα με την μαστεκτομή ή αργότερα με την χρήση ειδικών ενθεμάτων.

Ακτινοθεραπεία.

Στην ακτινοθεραπεία χρησιμοποιείται υψηλής ενέργειας ιονίζουσα ακτινοβολία. Σε περίπτωση καρκίνου του μαστού εφαρμόζεται συνήθως μετά την χειρουργική επέμβαση, με σκοπό να σκοτώσει ή να εμποδίσει να αναπτυχθούν καρκινικά κύτταρα που τυχόν έχουν απομείνει στην περιοχή. Υπεύθυνοι για την ακτινοθεραπεία είναι ειδικοί ιατροί, οι ακτινοθεραπευτές.

Χημειοθεραπεία.

Η χημειοθεραπεία χρησιμοποιεί ειδικά φάρμακα, που σταματούν την εξέλιξη του καρκίνου, είτε σκοτώνοντας τα καρκινικά κύτταρα είτε εμποδίζοντας τα να πολλαπλασιαστούν. Τα φάρμακα παίρνονται από το στόμα σαν χάπια ή δίνονται με μορφή ένεσης (ενδοφλέβια ή ενδομυϊκή). Υπάρχουν διάφορες κατηγορίες χημειοθεραπευτικών φαρμάκων. Συχνά στον καρκίνο του μαστού χρησιμοποιούνται περισσότερα από ένα φάρμακα.

Η χημειοθεραπεία χρησιμοποιείται:

- Σε περίπτωση **αρχόμενου διηθητικού καρκίνου** για να εξοντωθούν όσα καρκινικά κύτταρα έχουν τυχόν παραμείνει μετά την χειρουργική επέμβαση. Επίσης, για να ελαττωθεί ο κίνδυνος υποτροπής του καρκίνου.
- Σε περίπτωση **προχωρημένου καρκίνου με μεταστάσεις** για να καταστρέψει όσο το δυνατό περισσότερα καρκινικά κύτταρα.
- Σε κάποιες περιπτώσεις δίνεται πριν την χειρουργική επέμβαση (**προεγχειρητική χημειοθεραπεία**).

Η χημειοθεραπεία χορηγείται:

- Ενδοφλέβια με την μορφή ορού.
- Ενδομυϊκά με την μορφή ένεσης.
- Από το στόμα σαν χάπι ή κάψουλα.

Ορμονοθεραπεία.

Η ορμονοθεραπεία είναι χρήσιμη σε περίπτωση ορμονο-εξαρτώμενου καρκίνου του μαστού και έχει σαν στόχο:

- Να ελαττώσει τα επίπεδα των ορμονών (κυρίως της οιστραδιόλης)
- Να αναστέλλει την δράση τους πάνω στα καρκινικά κύτταρα.

Υπάρχουν μορφές καρκίνου του μαστού που έχουν ορμονικούς υποδοχείς οιστρογόνων ή/και προγεστερόνης. Η παρουσία αυτών των ορμονών προάγει τον πολλαπλασιασμό των κυττάρων και την ανάπτυξη του καρκίνου. Οι παραπάνω ορμόνες παράγονται κυρίως από τις ωοθήκες. Στόχος της θεραπείας είναι να αναστείλει την λειτουργία των ωοθηκών και την παραγωγή των ορμονών. Η θεραπεία αυτή δεν έχει ένδειξη σε μη-ορμονοεξαρτώμενους καρκίνους.

Στοχευμένη θεραπεία.

Η στοχευμένη θεραπεία χρησιμοποιεί φάρμακα ή ουσίες που αναγνωρίζουν και επιτίθενται μόνο στα καρκινικά κύτταρα χωρίς να επηρεάζονται τα φυσιολογικά κύτταρα. Σε κάποιες από τις θεραπείες αυτές χρησιμοποιούνται αντισώματα που μοιάζουν με εκείνα που φυσιολογικά παράγονται από το ανοσολογικό σύστημα του οργανισμού. Ονομάζονται **ανοσοθεραπείες** και έχουν σαν στόχο να ενεργοποιήσουν και να κατευθύνουν το ανοσοποιητικό σύστημα του οργανισμού εναντίον της ανάπτυξης των καρκινικών κυττάρων.

ΚΕΦΑΛΑΙΟ 2

ΠΕΡΙΓΡΑΦΙΚΗ ΑΝΑΛΥΣΗ

2.1 Εισαγωγή

Στο παρόν κεφάλαιο θα πραγματοποιηθεί μια παρουσίαση και σύντομη ανάλυση των περιγραφικών δεικτών που επιλέχθηκαν, με στόχο την καλύτερη μελέτη και κατανόηση του δείγματος, που χρησιμοποιείται για τη μελέτη του καρκίνου του μαστού.

Τα δεδομένα που θα χρησιμοποιηθούν στην παρακάτω ανάλυση τα αντλήσαμε από μία βάση δεδομένων του λογισμικού SEER*Stat.

Το SEER είναι μία έγκυρη πηγή για τις στατιστικές του καρκίνου στις Ηνωμένες Πολιτείες. Το Πρόγραμμα Επιτήρησης, Επιδημιολογίας και Τελικών Αποτελεσμάτων (SEER) παρέχει πληροφορίες σχετικά με τις στατιστικές για τον καρκίνο σε μια προσπάθεια να μειωθεί το βάρος του καρκίνου στον πληθυσμό των ΗΠΑ. Το SEER υποστηρίζεται από το Πρόγραμμα Ερευνών Επιτήρησης (SRP) στο Τμήμα Ελέγχου του Καρκίνου και Επιστημών Πληθυσμού (DCCPS) του NCI.

Τα δεδομένα της έρευνας SEER περιλαμβάνουν τη συχνότητα εμφάνισης και δεδομένα πληθυσμού που σχετίζονται με την ηλικία, το φύλο, τη φυλή, το έτος διάγνωσης και τις γεωγραφικές περιοχές (συμπεριλαμβανομένου του μητρώου SEER και της κομητείας). Το SEER δημοσιεύει νέα ερευνητικά δεδομένα κάθε άνοιξη με βάση την υποβολή δεδομένων του προηγούμενου Νοεμβρίου.

2.2 Περιγραφική Ανάλυση

Τα δεδομένα που δουλεύουμε αποτελούνται 5552 παρατηρήσεις και 20 στήλες. Κάθε στήλη αντιπροσωπεύει μία μεταβλητή, οι τιμές των οποίων παρουσιάζονται παρακάτω :

- 1) **Year of diagnosis** (2010 - 2015)
- 2) **Age** (20-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-59, 60-64, 65-69, 70-74, 75-79, 80-84, 85+)
- 3) **Sex** (female)
- 4) **Race** (unknown, white, black, other)
- 5) **State** (Washington)
- 6) **Marital Status** (single, married, separated, divorced, widowed, unmarried or domestic partner)
- 7) **Grade** [μία μέτρηση του πόσο τα καρκινικά κύτταρα μοιάζουν με κανονικά κύτταρα] .

Grade I = τα καρκινικά κύτταρα βαθμού I μοιάζουν **λίγο διαφορετικά από τα φυσιολογικά κύτταρα** και αναπτύσσονται σε αργά, καλά οργανωμένα πρότυπα.

Grade II = τα καρκινικά κύτταρα βαθμού II **δε μοιάζουν με τα φυσιολογικά / κανονικά κύτταρα** και αναπτύσσονται και διαιρούνται λίγο πιο γρήγορα από το κανονικό.

Grade III = τα κύτταρα βαθμού III **φαίνονται πολύ διαφορετικά από τα κανονικά κύτταρα**, αναπτύσσονται γρήγορα σε μη οργανωμένα, ακανόνιστα σχήματα, με πολλά διαιρούμενα να δημιουργούν νέα καρκινικά κύτταρα.

- 8) **Laterality** [προέλευση πρωτογενούς] - (Right origin of primary , Left origin of primary)
- 9) **Stage A** [έκταση της νόσου] - (Regional=τοπική , Distant=μακρινή)
- 10) **Tumor Size** (1 - 150 mm)
- 11) **ER Status** = υποδοχέας οιστρογόνου (Positive , Negative)
- 12) **PR Status** = υποδοχέας προγεστερόνης (Positive , Negative)
- 13) **Regional Nodes Examined** = περιφερειακοί όγκοι που εξετάστηκαν (0-99)
- 14) **Regional Nodes Positive** = περιφερειακοί όγκοι θετικοί (0-99)
- 15) **Breast Subtype** = δευτερεύον τύπος μαστού

Luminal A : θετικός στους υποδοχείς ορμονών (υποδοχείς οιστρογόνου ή/ και θετικός στον υποδοχέα προγεστερόνης) και αρνητικός στο Her2

Her2- / HR+

(χαμηλά επίπεδα πρωτεΐνης Ki-67 η οποία βοηθά στον έλεγχο της ταχύτητας ανάπτυξης καρκινικών κυττάρων, καρκίνοι χαμηλού βαθμού, τείνουν να αναπτύσσονται αργά και έχουν την καλύτερη πρόγνωση)

Luminal B : θετικός στους υποδοχείς ορμονών (υποδοχείς οιστρογόνου ή/ και θετικός στον υποδοχέα προγεστερόνης) και είτε θετικός είτε αρνητικός στο Her2

Her2- / HR+

Her2+ / HR+

(υψηλά επίπεδα πρωτεΐνης Ki-67 η οποία βοηθά στον έλεγχο της ταχύτητας ανάπτυξης καρκινικών κυττάρων, αναπτύσσονται ελαφρώς ταχύτερα από τους καρκίνους Luminal A και η πρόγνωσή τους είναι ελαφρώς χειρότερη)

Τριπλός Αρνητικός : αρνητικός στους υποδοχείς ορμονών (υποδοχείς οιστρογόνου ή/ και αρνητικός στον υποδοχέα προγεστερόνης) και αρνητικός στο Her2.

Triple Negative -> Her2- / HR-

HER2-enriched : ο εμπλουτισμένος με HER2 καρκίνος του μαστού είναι αρνητικός στους υποδοχείς ορμονών (υποδοχείς οιστρογόνου ή/ και αρνητικός στον υποδοχέα προγεστερόνης) και θετικός στο HER2.

Her2+ / HR-

(τείνουν να αναπτύσσονται γρηγορότερα από τους καρκίνους Luminal και μπορεί να έχουν χειρότερη πρόγνωση, αλλά συχνά αντιμετωπίζονται επιτυχώς με στοχευμένες θεραπείες που στοχεύουν στην πρωτεΐνη HER2)

- 16) **6th Stage** [είτε κλινικό είτε παθολογικό στάδιο]

Στάδιο 1 : Αυτό το στάδιο είναι συνήθως ένας μικρός καρκίνος ή όγκος που δεν έχει αναπτυχθεί βαθιά σε κοντινούς ιστούς. Επίσης, δεν έχει εξαπλωθεί στους λεμφαδένες ή σε άλλα μέρη του σώματος. Ονομάζεται συχνά καρκίνος πρώιμου σταδίου.

Στάδιο 2 και 3 : Γενικά, αυτά τα 2 στάδια υποδεικνύουν μεγαλύτερους καρκίνους ή όγκους που έχουν αναπτυχθεί βαθύτερα σε κοντινούς ιστούς. Μπορεί επίσης, να έχουν εξαπλωθεί σε λεμφαδένες αλλά όχι σε άλλα μέρη του σώματος.

Στάδιο 4 : Αυτό το στάδιο σημαίνει ότι ο καρκίνος έχει εξαπλωθεί σε άλλα όργανα ή μέρη του σώματος. Μπορεί επίσης να ονομαστεί προχωρημένος ή μεταστατικός καρκίνος.

17) **Stage T**

Το γράμμα «T» συν έναν αριθμό από 0 έως 4 περιγράφει το μέγεθος και τη θέση του όγκου, συμπεριλαμβανομένου του κατά πόσο ο όγκος έχει αναπτυχθεί σε κοντινούς ιστούς. Το μέγεθος του όγκου μετριέται σε εκατοστά. Ένας μεγαλύτερος όγκος ή ένας όγκος που έχει αναπτυχθεί βαθύτερα στον περιβάλλοντα ιστό λαμβάνει μεγαλύτερο αριθμό.

18) **Stage N**

Το γράμμα «N» συν έναν αριθμό από 0 έως 3 σημαίνει λεμφαδένες. Τις περισσότερες φορές, όσο είναι περισσότεροι οι λεμφαδένες με καρκίνο, τόσο μεγαλύτερος είναι ο αριθμός που αποδίδεται.

19) **Survival Months** (0 - 515)

20) **Status** (Alive , Dead)

Η τιμή Alive σημαίνει ότι το άτομο επέζησε από τον καρκίνο του μαστού αλλά μπορεί να πέθανε από κάποια άλλη αιτία. Η τιμή Dead σημαίνει ότι το άτομο απεβίωσε από τον καρκίνο του μαστού.

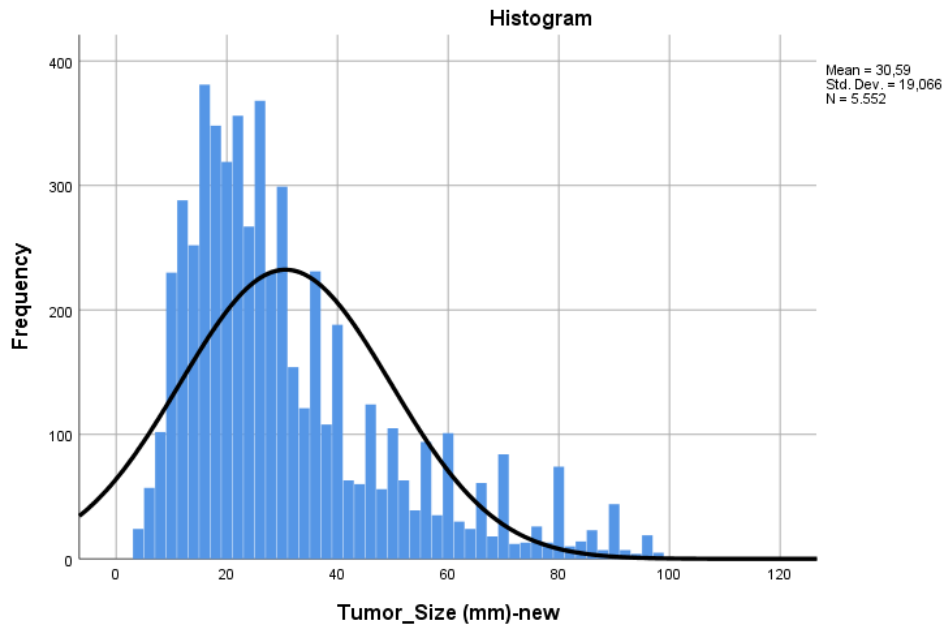
2.2.1 Ποσοτικές Μεταβλητές

1. Μέγεθος όγκου (Tumor_Size - μετριέται σε mm):

Παρακάτω παρουσιάζονται κάποια περιγραφικά μέτρα για τη μεταβλητή “tumor size”.

Statistics		
Tumor_Size (mm)-new		
N	Valid	5552
	Missing	0
Mean		30,59
Std. Error of Mean		,256
Median		25,00
Mode		25
Std. Deviation		19,066
Variance		363,502
Skewness		1,261
Std. Error of Skewness		,033
Kurtosis		1,178
Std. Error of Kurtosis		,066
Range		95
Minimum		4
Maximum		99
Sum		169808

Πίνακας 2.1 : Περιγραφικά μέτρα για τη μεταβλητή “tumor size”.



Διάγραμμα 2.1 : Ιστόγραμμα για τη μεταβλητή “tumor size”.

Σύμφωνα με τα παραπάνω, παρατηρούμε πως το μέσο μέγεθος του όγκου είναι 30,59 mm. Επίσης, παρατηρούμε πως η κατανομή αυτής της μεταβλητής είναι θετικά ασύμμετρη σύμφωνα με το συντελεστή ασυμμετρίας που είναι θετικός (= 1,178) και επαληθεύεται από το ιστόγραμμα.

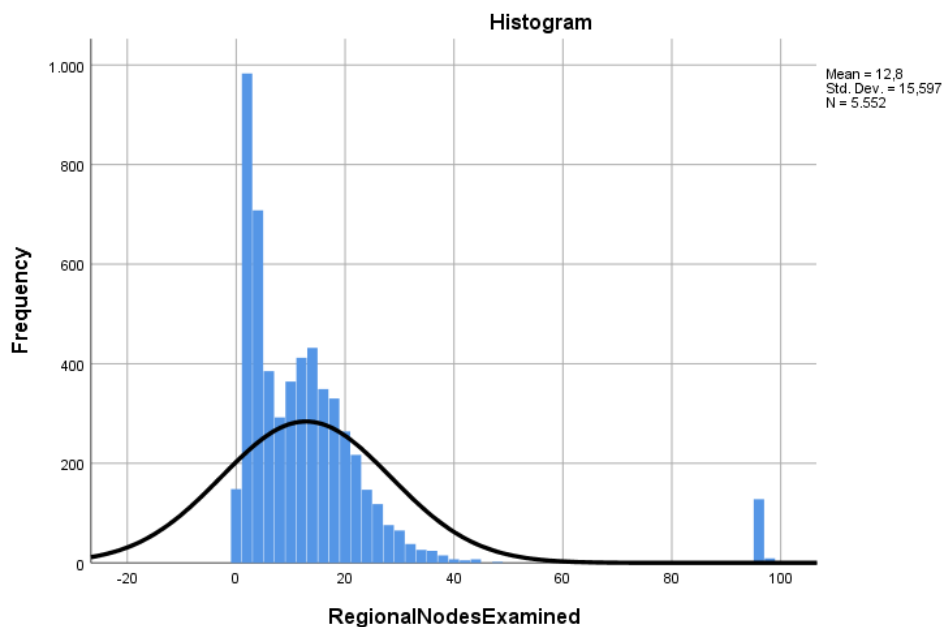
2. Περιφερειακοί Όγκοι που εξετάστηκαν (Regional Nodes Examined)

Παρακάτω παρουσιάζονται κάποια περιγραφικά μέτρα για τη μεταβλητή “Regional Nodes Examined”.

Statistics

RegionalNodesExamined		
N	Valid	5552
	Missing	0
Mean		12,80
Std. Error of Mean		,209
Median		10,00
Mode		2
Std. Deviation		15,597
Variance		243,262
Skewness		3,654
Std. Error of Skewness		,033
Kurtosis		16,486
Std. Error of Kurtosis		,066
Range		98
Minimum		0
Maximum		98
Sum		71074

Πίνακας 2.2 : Περιγραφικά μέτρα για τη μεταβλητή “Regional_Nodes_Examined”.



Διάγραμμα 2.2 : Ιστόγραμμα για τη μεταβλητή “Regional_Nodes_Examined”.

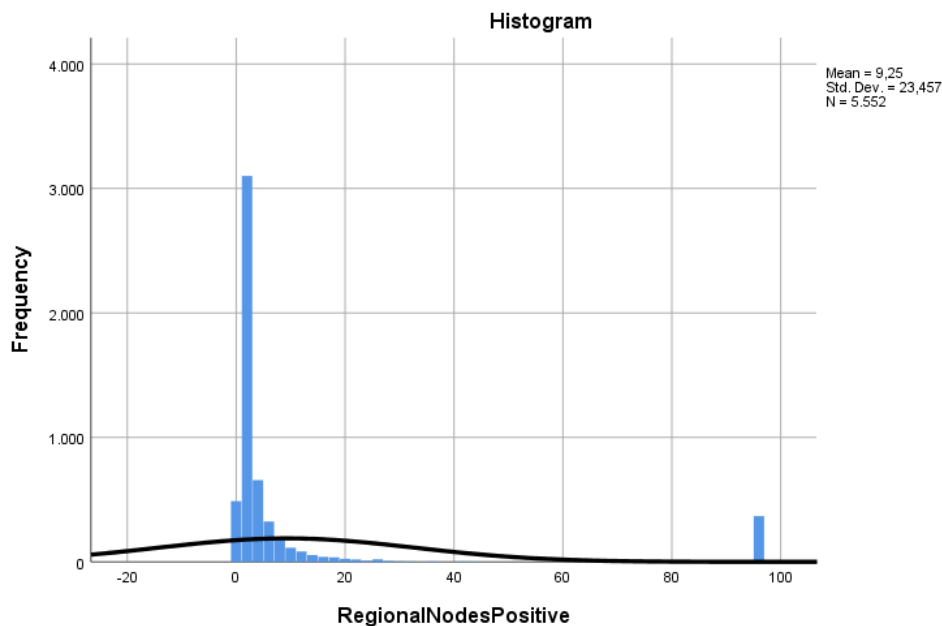
Σύμφωνα με τα παραπάνω, παρατηρούμε πως το μέσο πλήθος των περιφερειακών όγκων που εξετάστηκαν είναι 12,80. Επίσης, παρατηρούμε πως η κατανομή αυτής της μεταβλητής είναι θετικά ασύμμετρη σύμφωνα με το συντελεστή ασυμμετρίας που είναι θετικός (= 16,486) και επαληθεύεται από το ιστόγραμμα.

3. Περιφερειακοί Όγκοι που εξετάστηκαν και βγήκαν θετικοί (Regional Nodes Positive)

Παρακάτω παρουσιάζονται κάποια περιγραφικά μέτρα για τη μεταβλητή “Regional Nodes Positive”.

Statistics		
RegionalNodesPositive		
N	Valid	5552
	Missing	0
Mean		9,25
Std. Error of Mean		,315
Median		2,00
Mode		1
Std. Deviation		23,457
Variance		550,221
Skewness		3,276
Std. Error of Skewness		,033
Kurtosis		9,087
Std. Error of Kurtosis		,066
Range		97
Minimum		0
Maximum		97
Sum		51331

Πίνακας 2.3: Περιγραφικά μέτρα για τη μεταβλητή “Regional_Nodes_Positive”.



Διάγραμμα 2.3: Ιστόγραμμα για τη μεταβλητή “Regional_Nodes_Positive”.

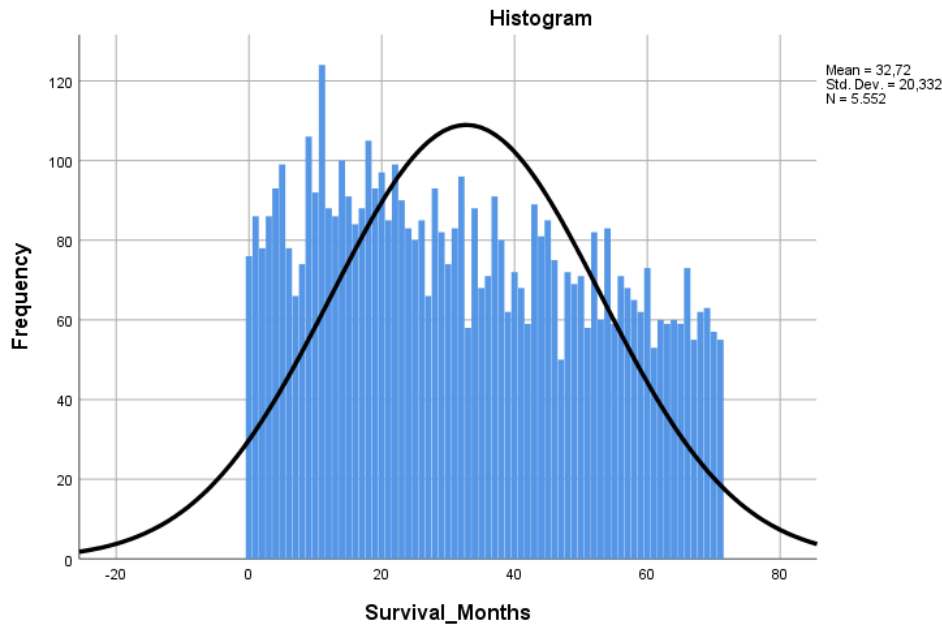
Σύμφωνα με τα παραπάνω, παρατηρούμε πως το μέσο πλήθος των περιφερειακών όγκων που εξετάστηκαν και βγήκαν θετικοί είναι 9,25. Επίσης, παρατηρούμε πως η κατανομή αυτής της μεταβλητής είναι θετικά ασύμμετρη σύμφωνα με το συντελεστή ασυμμετρίας που είναι θετικός ($= 9,087$) και επαληθεύεται από το ιστόγραμμα.

4. Μήνες Επιβίωσης (Survival Months)

Παρακάτω παρουσιάζονται κάποια περιγραφικά μέτρα για τη μεταβλητή “Survival Months”.

Statistics		
Survival_Months		
N	Valid	5552
	Missing	0
Mean		32,72
Std. Error of Mean		,273
Median		31,00
Mode		11
Std. Deviation		20,332
Variance		413,390
Skewness		,186
Std. Error of Skewness		,033
Kurtosis		-1,138
Std. Error of Kurtosis		,066
Range		71
Minimum		0
Maximum		71
Sum		181678

Πίνακας 2.4: Περιγραφικά μέτρα για τη μεταβλητή “Survival_Months”.



Διάγραμμα 2.4: Ιστόγραμμα για τη μεταβλητή “Survival_Months”.

Σύμφωνα με τα παραπάνω, παρατηρούμε πως ο μέσος των μηνών επιβίωσης είναι 32,72. Επίσης, παρατηρούμε πως η κατανομή αυτής της μεταβλητής είναι αρνητικά ασύμμετρη σύμφωνα με το συντελεστή ασυμμετρίας που είναι αρνητικός ($= -1,138$) και επαληθεύεται από το ιστόγραμμα.

2.2.2 Ποιοτικές Μεταβλητές

1. Ηλικία

Μία πολύ σημαντική μεταβλητή στα πλαίσια της παρούσας έρευνας είναι αυτή της ηλικίας. Ύστερα από την διεκπεραίωση διαφόρων μελετών, έχει διαπιστωθεί ότι η ηλικία μιας γυναίκας μπορεί να αποτελέσει επιβαρυντικό παράγοντα ως προς την ανάπτυξη καρκίνου του μαστού. Αυτό αποτελεί σοβαρή αιτία για την εκτενέστερη μελέτη της μεταβλητής αυτής.

Έχει γίνει ομαδοποίηση των ηλικιών στις παρακάτω κατηγορίες:

- 1^η ηλικιακή ομάδα: 20-24 έτη
- 2^η ηλικιακή ομάδα: 25-29 έτη
- 3^η ηλικιακή ομάδα: 30-34 έτη
- 4^η ηλικιακή ομάδα: 35-39 έτη
- 5^η ηλικιακή ομάδα: 40-44 έτη
- 6^η ηλικιακή ομάδα: 45-49 έτη
- 7^η ηλικιακή ομάδα: 50-54 έτη
- 8^η ηλικιακή ομάδα: 55-59 έτη
- 9^η ηλικιακή ομάδα: 60-64 έτη
- 10^η ηλικιακή ομάδα: 65-69 έτη
- 11^η ηλικιακή ομάδα: 70-74 έτη
- 12^η ηλικιακή ομάδα: 75-80 έτη

13^η ηλικιακή ομάδα: 80-84 έτη

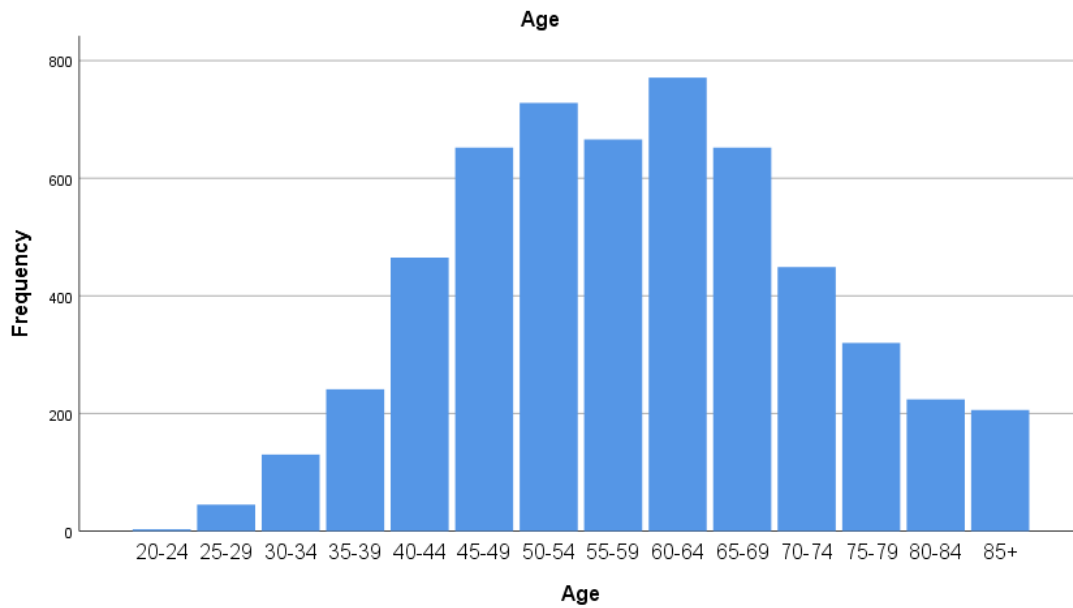
14^η ηλικιακή ομάδα: 85+ έτη

		Age			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	20-24	3	,1	,1	,1
	25-29	45	,8	,8	,9
	30-34	130	2,3	2,3	3,2
	35-39	241	4,3	4,3	7,5
	40-44	465	8,4	8,4	15,9
	45-49	652	11,7	11,7	27,7
	50-54	728	13,1	13,1	40,8
	55-59	666	12,0	12,0	52,8
	60-64	771	13,9	13,9	66,7
	65-69	652	11,7	11,7	78,4
	70-74	449	8,1	8,1	86,5
	75-79	320	5,8	5,8	92,3
	80-84	224	4,0	4,0	96,3
	85+	206	3,7	3,7	100,0
	Total	5552	100,0	100,0	

Πίνακας 2.5 : Συχνότητες ανά ηλικιακή ομάδα.

Μπορούμε να διαπιστώσουμε ότι το 0,1%, δηλαδή 3 από τις 5552 γυναίκες του δείγματος με καρκίνο του μαστού ανήκουν στην ηλικιακή ομάδα 20 με 24, το 0,8%, δηλαδή 45 από τις 5552 γυναίκες με καρκίνο του μαστού ανήκουν στην ηλικιακή ομάδα 25-29, το 2,3%, δηλαδή 130 από τις 5552 γυναίκες με καρκίνο του μαστού ανήκουν στην ηλικιακή ομάδα 30-34, το 4,3%, δηλαδή 241 από τις 5552 γυναίκες με καρκίνο του μαστού ανήκουν στην ηλικιακή ομάδα 35-39, το 8,4%, δηλαδή 465 από τις 5552 γυναίκες με καρκίνο του μαστού ανήκουν στην ηλικιακή ομάδα 40-44, το 11,7%, δηλαδή 652 από τις 5552 γυναίκες με καρκίνο του μαστού ανήκουν στην ηλικιακή ομάδα 45-49, το 13,1%, δηλαδή 728 από τις 5552 γυναίκες με καρκίνο του μαστού ανήκουν στην ηλικιακή ομάδα 50-54, το 12%, δηλαδή 666 από τις 5552 γυναίκες με καρκίνο του μαστού ανήκουν στην ηλικιακή ομάδα 55-59, το 13,9%, δηλαδή 771 από τις 5552 γυναίκες με καρκίνο του μαστού ανήκουν στην ηλικιακή ομάδα 60-64, το 11,7%, δηλαδή 652 από τις 5552 γυναίκες με καρκίνο του μαστού ανήκουν στην ηλικιακή ομάδα 65-69, το 8,1%, δηλαδή 449 από τις 5552 γυναίκες με καρκίνο του μαστού ανήκουν στην ηλικιακή ομάδα 70-74, το 5,8%, δηλαδή 320 από τις 5550 γυναίκες με καρκίνο του μαστού ανήκουν στην ηλικιακή ομάδα 75-79, το 4%, δηλαδή 224 από τις 5550 γυναίκες με καρκίνο του μαστού ανήκουν στην ηλικιακή ομάδα 80-84 και 3,7%, δηλαδή 206 από τις 5550 γυναίκες με καρκίνο του μαστού ανήκουν στην ηλικιακή ομάδα 85+. Έτσι λοιπόν, στο δείγμα μας, οι περισσότερες γυναίκες με καρκίνο του μαστού είναι ηλικίας 60-64 ετών.

Στο συμπέρασμα αυτό μας οδηγεί και το παρακάτω διάγραμμα συχνοτήτων με βάση τις ηλικιακές ομάδες.

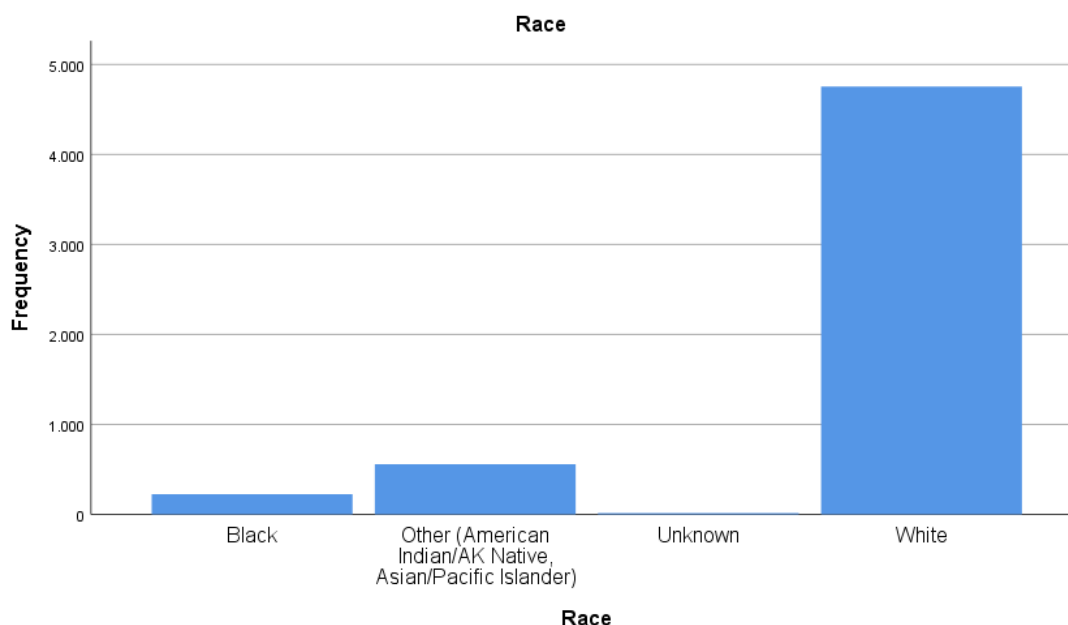


Διάγραμμα 2.5 : Bar chart συχνοτήτων με βάση τις ηλικιακές ομάδες.

2. Φυλή (Race)

		Race			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Black	225	4,1	4,1	4,1
	Other (American Indian/AK Native, Asian/Pacific Islander)	558	10,1	10,1	14,1
	Unknown	14	,3	,3	14,4
	White	4755	85,6	85,6	100,0
	Total	5552	100,0	100,0	

Πίνακας 2.6 : Συχνότητες ανά φυλή.



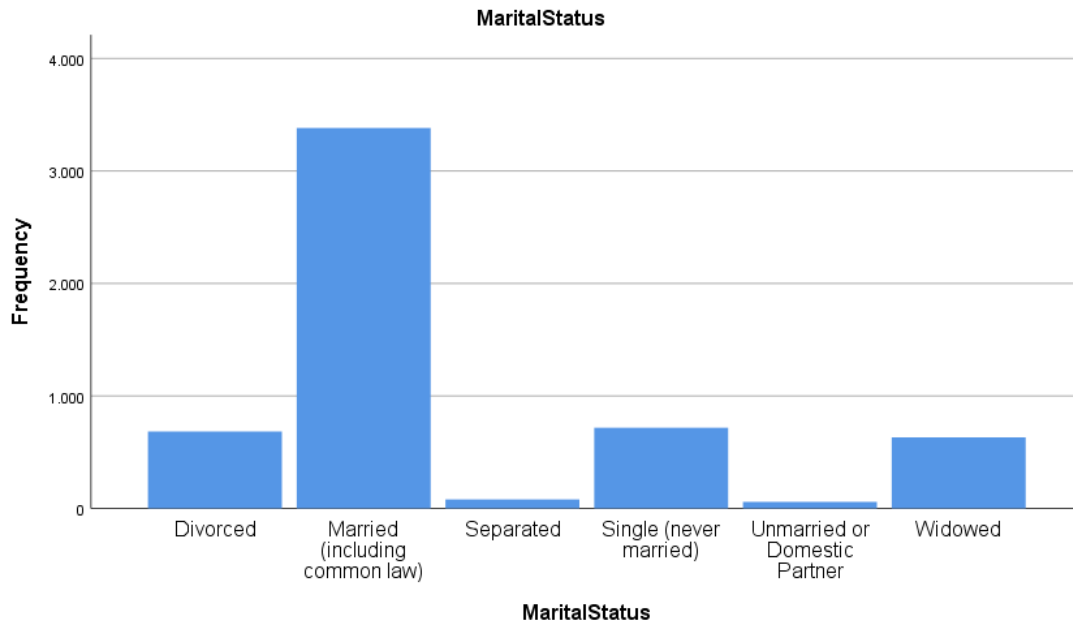
Διάγραμμα 2.6 : Bar chart συχνοτήτων με βάση την φυλή.

Σύμφωνα με τα παραπάνω, συμπεραίνουμε ότι το 4,1%, δηλαδή 225 από τις 5552 γυναίκες του δείγματος ήταν αфроαμερικάνες, το 85,6%, δηλαδή 4755 από τις 5552 γυναίκες του δείγματος ήταν λευκές, το 0,3%, δηλαδή 14 από τις 5552 γυναίκες του δείγματος ήταν άγνωστης φυλής και το 10,1%, δηλαδή 558 από τις 5552 γυναίκες του δείγματος ανήκαν σε άλλη φυλή (American Indian/AK Native, Asian/Pacific Islander).

3. Marital Status

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid Divorced	684	12,3	12,3	12,3
Married (including common law)	3382	60,9	60,9	73,2
Separated	80	1,4	1,4	74,7
Single (never married)	717	12,9	12,9	87,6
Unmarried or Domestic Partner	58	1,0	1,0	88,6
Widowed	631	11,4	11,4	100,0
Total	5552	100,0	100,0	

Πίνακας 2.7 : Συχνότητες ανά οικογενειακή κατάσταση.



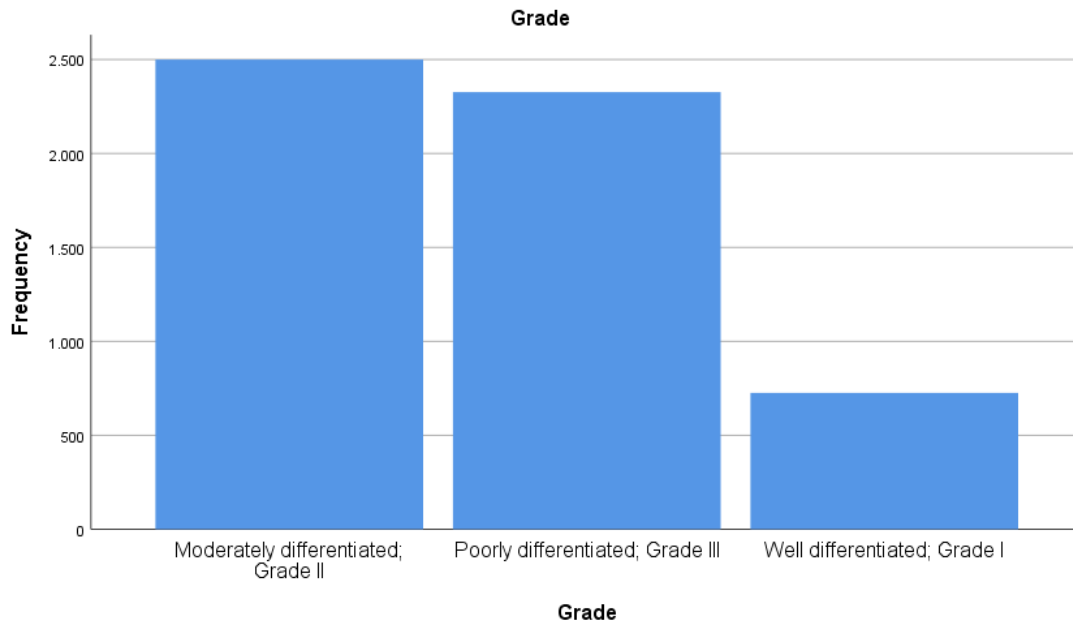
Διάγραμμα 2.7 : Bar chart συχνοτήτων με βάση την οικογενειακή κατάσταση.

Σύμφωνα με τα παραπάνω, συμπεραίνουμε ότι το 12,3%, δηλαδή 684 από τις 5552 γυναίκες του δείγματος ήταν διαζευγμένες, το 60,9%, δηλαδή 3382 από τις 5552 γυναίκες του δείγματος ήταν παντρεμένες, το 1,4%, δηλαδή 80 από τις 5552 γυναίκες του δείγματος ήταν σε διάσταση, το 12,9%, δηλαδή 717 από τις 5552 γυναίκες του δείγματος ήταν ανύπαντρες, το 1%, δηλαδή 58 από τις 5552 γυναίκες του δείγματος ήταν με συγκατοικο αλλά ανύπαντρες και το 11,4%, δηλαδή 631 από τις 5552 γυναίκες του δείγματος ήταν χήρες.

4. Grade

		Grade			Cumulative
		Frequency	Percent	Valid Percent	Percent
Valid	Moderately differentiated; Grade II	2500	45,0	45,0	45,0
	Poorly differentiated; Grade III	2326	41,9	41,9	86,9
	Well differentiated; Grade I	726	13,1	13,1	100,0
	Total	5552	100,0	100,0	

Πίνακας 2.8 : Συχνότητες ανά μία μέτρηση του πόσο τα καρκινικά κύτταρα μοιάζουν με κανονικά.



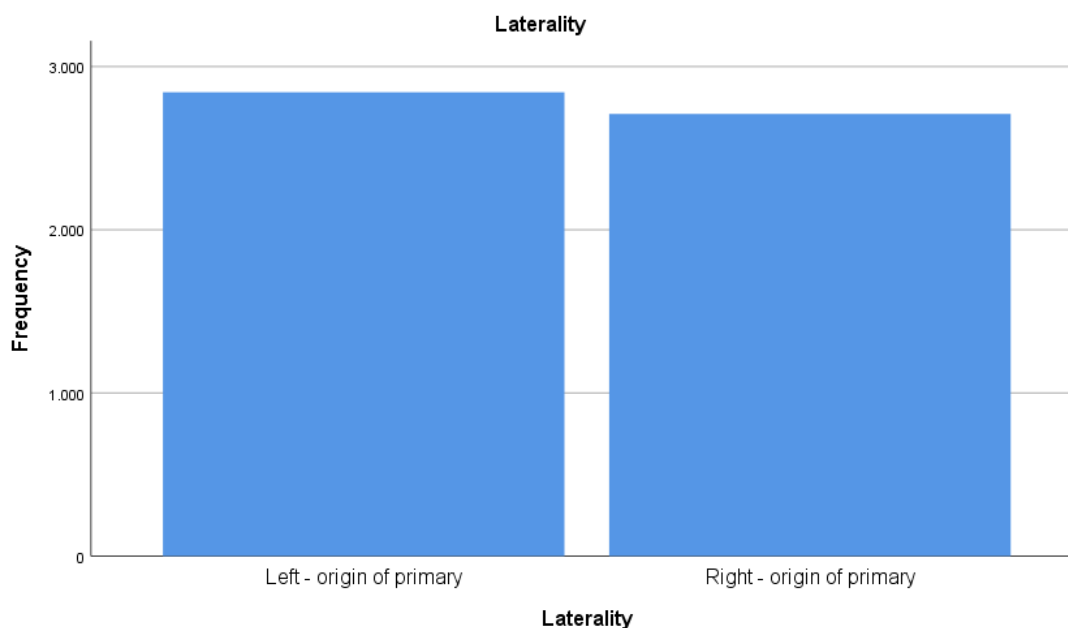
Διάγραμμα 2.8 : Bar chart συχνοτήτων με βάση μία μέτρηση του πόσο τα καρκινικά κύτταρα μοιάζουν με κανονικά.

Σύμφωνα με τα παραπάνω, συμπεραίνουμε ότι το 13,1%, δηλαδή 726 από τις 5552 γυναίκες του δείγματος ήταν είχαν καρκινικά κύτταρα βαθμού I, δηλαδή κύτταρα που μοιάζουν λίγο διαφορετικά από τα φυσιολογικά κύτταρα και αναπτύσσονται αργά, το 45%, δηλαδή 2500 από τις 5552 γυναίκες του δείγματος είχαν καρκινικά κύτταρα βαθμού II, δηλαδή κύτταρα που δεν μοιάζουν με τα φυσιολογικά/κανονικά κύτταρα και αναπτύσσονται λίγο πιο γρήγορα από τα κανονικά και το 41,9%, δηλαδή 1326 από τις 5552 γυναίκες του δείγματος είχαν καρκινικά κύτταρα βαθμού III, δηλαδή κύτταρα που φαίνονται πολύ διαφορετικά από τα φυσιολογικά/κανονικά κύτταρα και αναπτύσσονται γρήγορα σε μη οργανωμένα σχήματα, με πολλά διαιρούμενα να δημιουργούν νέα καρκινικά κύτταρα.

5. Laterality

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid Left - origin of primary	2842	51,2	51,2	51,2
Valid Right - origin of primary	2710	48,8	48,8	100,0
Total	5552	100,0	100,0	

Πίνακας 2.9 : Συχνότητες ανά την προέλευση πρωτογενούς.



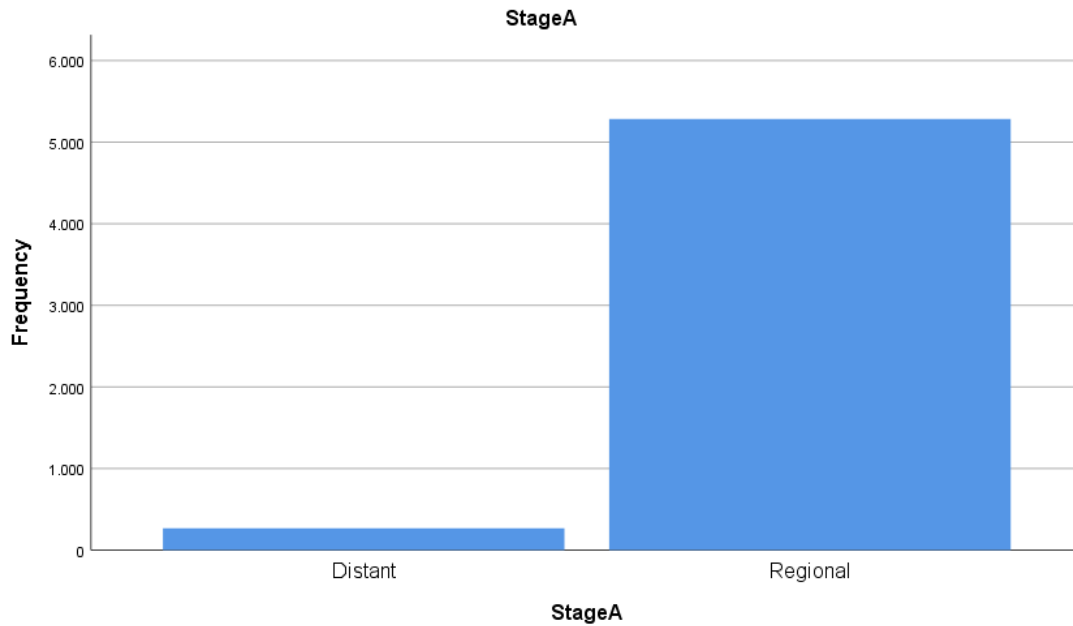
Διάγραμμα 2.9 : Bar chart συχνοτήτων με βάση την προέλευση πρωτογενούς.

Σύμφωνα με τα παραπάνω, συμπεραίνουμε ότι το 51,2%, δηλαδή 2842 από τις 5552 γυναίκες του δείγματος είχαν προέλευση πρωτογενούς όγκου στο αριστερό στήθος ενώ το 48,8% δηλαδή 2710 από τις 5552 γυναίκες του δείγματος είχαν προέλευση πρωτογενούς όγκου στο δεξιό στήθος.

6. Stage A

		StageA			Cumulative
		Frequency	Percent	Valid Percent	Percent
Valid	Distant	268	4,8	4,8	4,8
	Regional	5284	95,2	95,2	100,0
	Total	5552	100,0	100,0	

Πίνακας 2.10 : Συχνότητες ανά την έκταση της νόσου.



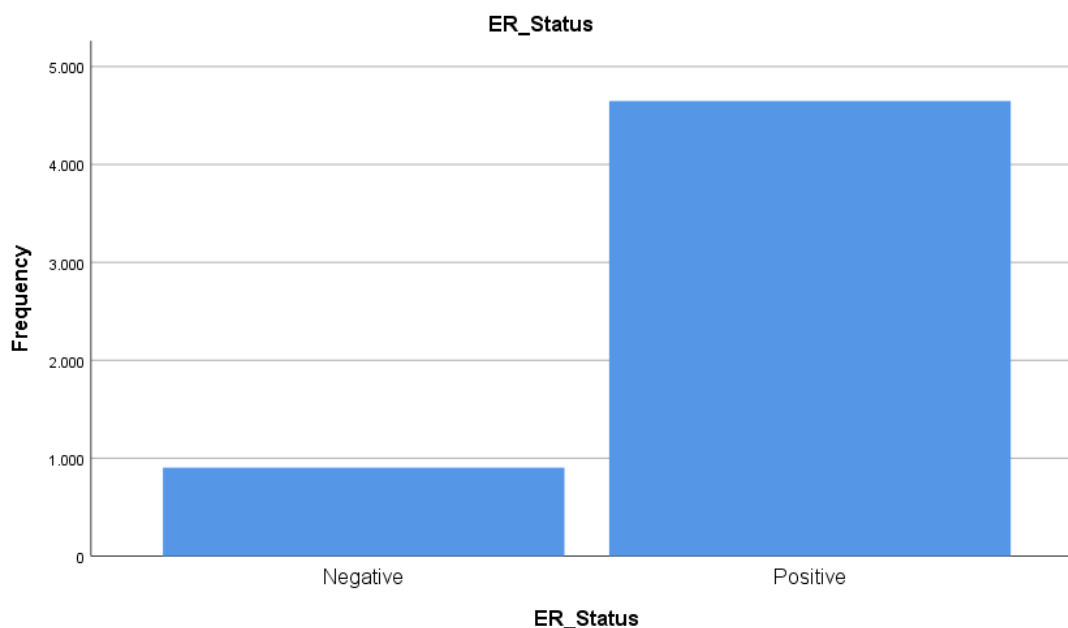
Διάγραμμα 2.10 : Bar chart συχνοτήτων με βάση την έκταση της νόσου.

Σύμφωνα με τα παραπάνω, συμπεραίνουμε ότι το 4,8%, δηλαδή 268 από τις 5552 γυναίκες του δείγματος είχαν έκταση της νόσου μακρινή ενώ το 95,2% δηλαδή 5284 από τις 5552 είχαν έκταση της νόσου τοπική.

7. ER Status

		ER_Status			Cumulative
		Frequency	Percent	Valid Percent	Percent
Valid	Negative	904	16,3	16,3	16,3
	Positive	4648	83,7	83,7	100,0
	Total	5552	100,0	100,0	

Πίνακας 2.11 : Συχνότητες ανά τους υποδοχείς οιστρογόνου.



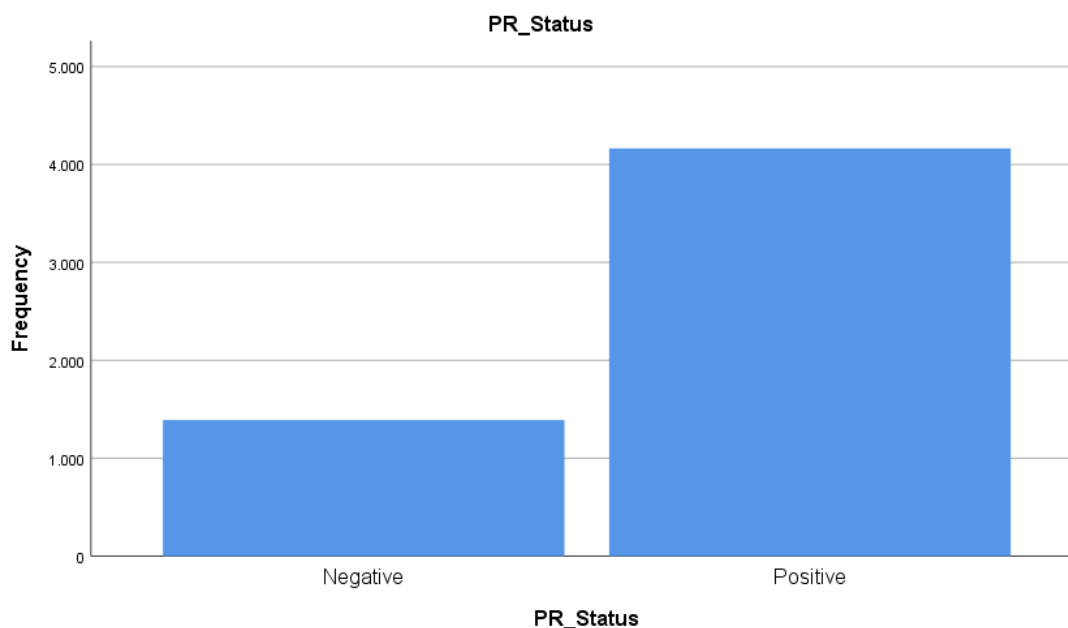
Διάγραμμα 2.11 : Bar chart συχνότητων με βάση τους υποδοχείς οιστρογόνου.

Σύμφωνα με τα παραπάνω, συμπεραίνουμε ότι το 16,3%, δηλαδή 904 από τις 5552 γυναίκες του δείγματος είχαν υποδοχείς οιστρογόνου αρνητικούς ενώ το 83,7% δηλαδή 4648 από τις 5552 είχαν υποδοχείς οιστρογόνου θετικούς.

8. PR Status

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Negative	1389	25,0	25,0	25,0
	Positive	4163	75,0	75,0	100,0
	Total	5552	100,0	100,0	

Πίνακας 2.12 : Συχνότητες ανά τους υποδοχείς προγεστερόνης.



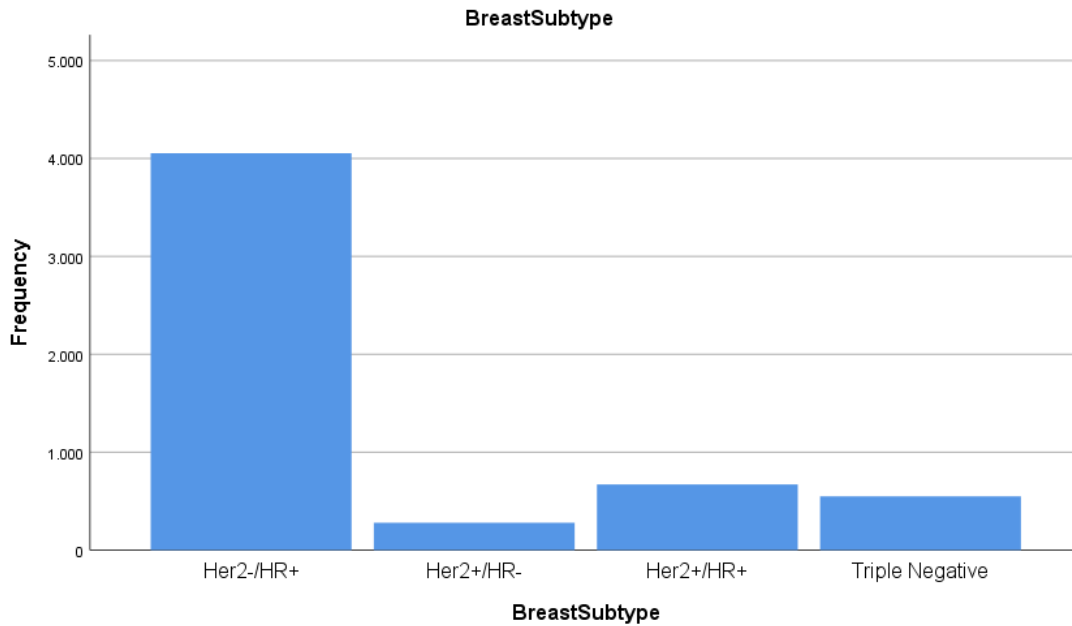
Διάγραμμα 2.12 : Bar chart συχνοτήτων με βάση τους υποδοχείς προγεστερόνης.

Σύμφωνα με τα παραπάνω, συμπεραίνουμε ότι το 25%, δηλαδή 1389 από τις 5552 γυναίκες του δείγματος είχαν υποδοχείς προγεστερόνης αρνητικούς ενώ το 75% δηλαδή 4163 από τις 5552 είχαν υποδοχείς προγεστερόνης θετικούς.

9. BreastSubtype

		BreastSubtype			Cumulative
		Frequency	Percent	Valid Percent	Percent
Valid	Her2-/HR+	4054	73,0	73,0	73,0
	Her2+/HR-	279	5,0	5,0	78,0
	Her2+/HR+	670	12,1	12,1	90,1
	Triple Negative	549	9,9	9,9	100,0
Total		5552	100,0	100,0	

Πίνακας 2.13 : Συχνότητες ανά του δευτερεύοντος τύπου μαστού.



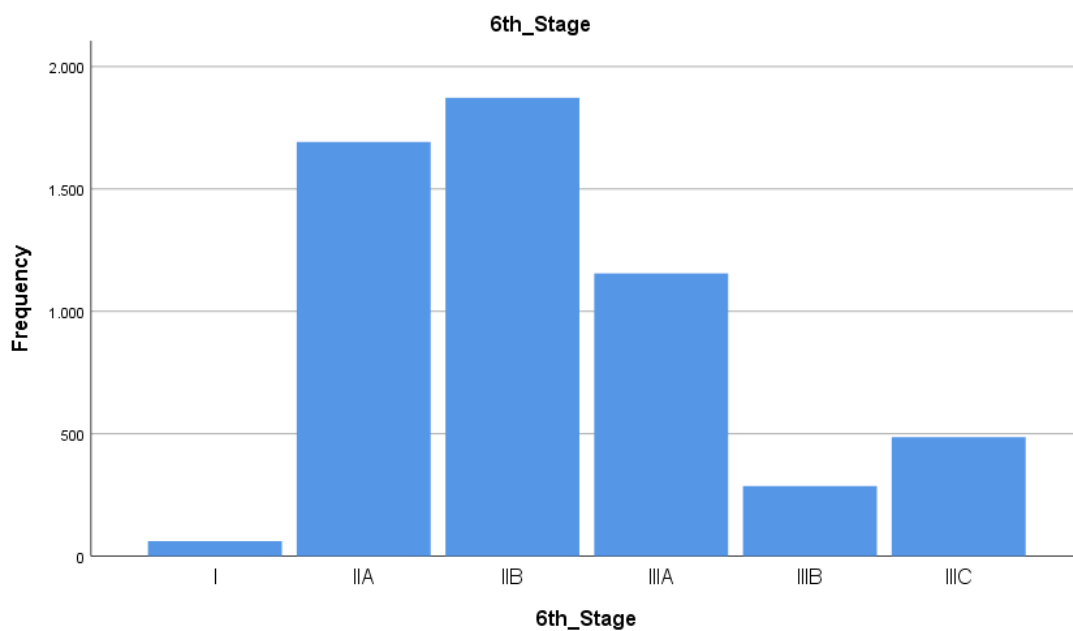
Διάγραμμα 2.13 : Bar chart συχνοτήτων με βάση τον δευτερεύον τύπο μαστού.

Σύμφωνα με τα παραπάνω, συμπεραίνουμε ότι το 73%, δηλαδή 4054 από τις 5552 γυναίκες του δείγματος είχαν δευτερεύον τύπου μαστό θετικό στους υποδοχείς ορμονών (οιστρογόνου ή προγεστερόνης) και αρνητικό στο Her2 (χρησιμοποιείται για την πρόβλεψη της ανταπόκρισης στην ορμονοθεραπεία και τη χημειοθεραπεία – παρακολούθηση της θεραπείας), το 5%, δηλαδή 279 από τις 5552 γυναίκες του δείγματος είχαν δευτερεύον τύπου μαστό αρνητικό στους υποδοχείς ορμονών και θετικό στο Her2, το 12,1%, δηλαδή 670 από τις 5552 γυναίκες του δείγματος είχαν δευτερεύον τύπου μαστό θετικό στους υποδοχείς ορμονών και θετικό στο Her2 και το 9,9%, δηλαδή 549 από τις 5552 γυναίκες του δείγματος είχαν δευτερεύον τύπου μαστό αρνητικό στους υποδοχείς ορμονών και στο Her2.

10. 6th Stage

		6th_Stage			Cumulative
		Frequency	Percent	Valid Percent	Percent
Valid	I	61	1,1	1,1	1,1
	IIA	1692	30,5	30,5	31,6
	IIB	1872	33,7	33,7	65,3
	IIIA	1155	20,8	20,8	86,1
	IIIB	286	5,2	5,2	91,2
	IIIC	486	8,8	8,8	100,0
	Total	5552	100,0	100,0	

Πίνακας 2.14 : Συχνότητες ανά του 6^ο σταδίου.



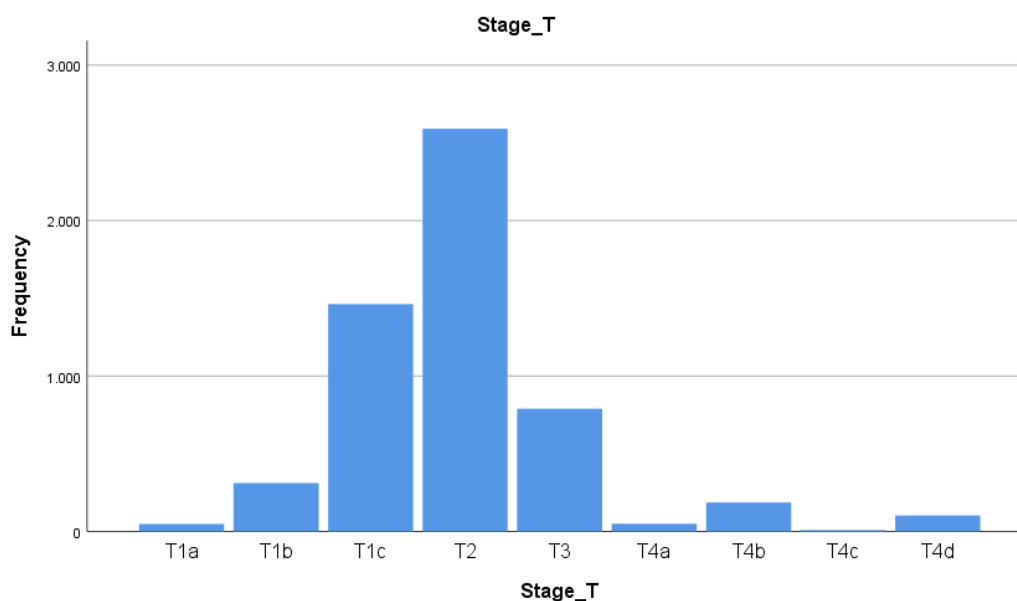
Διάγραμμα 2.14 : Bar chart συχνοτήτων με βάση το 6^ο στάδιο.

Σύμφωνα με τα παραπάνω, συμπεραίνουμε ότι το 1,1%, δηλαδή 61 από τις 5552 γυναίκες του δείγματος είχαν καρκίνο σε στάδιο 1 (αυτό το στάδιο είναι συνήθως ένας μικρός καρκίνος ή όγκος που δεν έχει αναπτυχθεί βαθιά σε κοντινούς ιστούς κι επίσης δεν έχει εξαπλωθεί στους λεμφαδένες ή σε άλλα μέρη του σώματος - ονομάζεται συχνά καρκίνος πρώιμου σταδίου), το 30,5%, δηλαδή 1692 από τις 5552 γυναίκες του δείγματος είχαν καρκίνο σε στάδιο 2A, το 33,7%, δηλαδή 1872 από τις 5552 γυναίκες του δείγματος είχαν καρκίνο σε στάδιο 2B, το 20,8%, δηλαδή 1155 από τις 5552 γυναίκες του δείγματος είχαν καρκίνο σε στάδιο 3A, το 5,2%, δηλαδή 286 από τις 5552 γυναίκες του δείγματος είχαν καρκίνο σε στάδιο 3B και το 8,8%, δηλαδή 486 από τις 5552 γυναίκες του δείγματος είχαν καρκίνο σε στάδιο 3C. (Γενικά, αυτά τα 2 στάδια - 2^ο και 3^ο- υποδεικνύουν μεγαλύτερους καρκίνους ή όγκους που έχουν αναπτυχθεί βαθύτερα σε κοντινούς ιστούς. Μπορεί επίσης, να έχουν εξαπλωθεί σε λεμφαδένες αλλά όχι σε άλλα μέρη του σώματος.)

11. Stage T

		Stage_T			Cumulative
		Frequency	Percent	Valid Percent	Percent
Valid	T1a	48	,9	,9	,9
	T1b	311	5,6	5,6	6,5
	T1c	1463	26,4	26,4	32,8
	T2	2591	46,7	46,7	79,5
	T3	790	14,2	14,2	93,7
	T4a	50	,9	,9	94,6
	T4b	187	3,4	3,4	98,0
	T4c	9	,2	,2	98,1
	T4d	103	1,9	1,9	100,0
	Total	5552	100,0	100,0	

Πίνακας 2.15 : Συχνότητες ανά του T σταδίου.



Διάγραμμα 2.15 : Bar chart συχνοτήτων με βάση το T στάδιο.

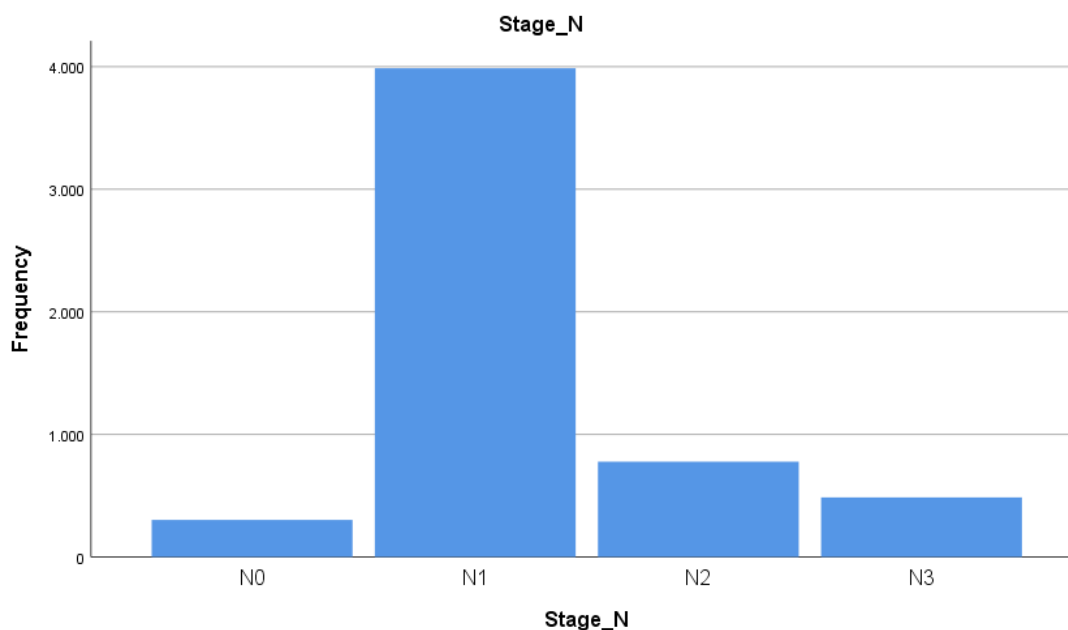
Σύμφωνα με τα παραπάνω, συμπεραίνουμε ότι το 0,9%, δηλαδή 48 από τις 5552 γυναίκες του δείγματος είχαν καρκίνο σε στάδιο T1a, το 5,6%, δηλαδή 311 από τις 5552 γυναίκες του δείγματος είχαν καρκίνο σε στάδιο T1b, το 26,4%, δηλαδή 1463 από τις 5552 γυναίκες του δείγματος είχαν καρκίνο σε στάδιο T1c, το 46,7%, δηλαδή 2591 από τις 5552 γυναίκες του δείγματος είχαν καρκίνο σε στάδιο T2, το 14,2%, δηλαδή 790 από τις 5552 γυναίκες του δείγματος είχαν καρκίνο σε στάδιο T3 και το 0,9%, δηλαδή 50 από τις 5552 γυναίκες του δείγματος είχαν καρκίνο σε στάδιο T4a. (Το γράμμα «T» συν ένας αριθμός από 0 έως 4 περιγράφει το μέγεθος και τη θέση του όγκου,

συμπεριλαμβανομένου του κατά πόσο ο όγκος έχει αναπτυχθεί σε κοντινούς ιστούς. Το μέγεθος του όγκου μετριέται σε εκατοστά. Ένας μεγαλύτερος όγκος ή ένας όγκος που έχει αναπτυχθεί βαθύτερα στον περιβάλλοντα ιστό λαμβάνει μεγαλύτερο αριθμό.)

12. Stage N

		Stage_N			Cumulative
		Frequency	Percent	Valid Percent	Percent
Valid	N0	302	5,4	5,4	5,4
	N1	3987	71,8	71,8	77,3
	N2	777	14,0	14,0	91,2
	N3	486	8,8	8,8	100,0
	Total	5552	100,0	100,0	

Πίνακας 2.16 : Συχνότητες ανά του N σταδίου.



Διάγραμμα 2.16 : Bar chart συχνοτήτων με βάση το N στάδιο.

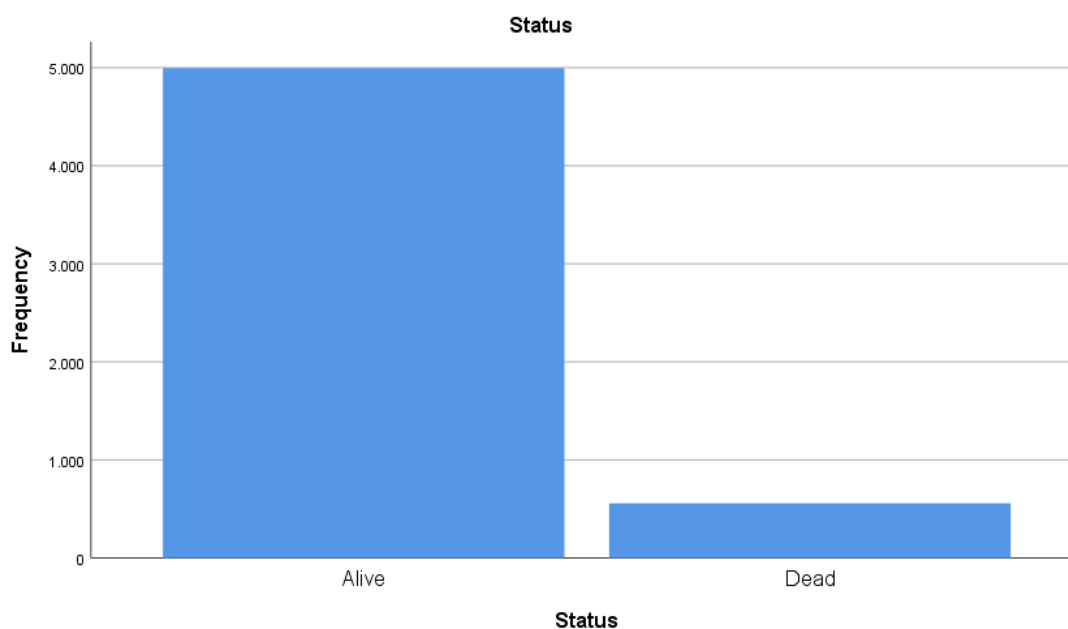
Σύμφωνα με τα παραπάνω, συμπεραίνουμε ότι το 5,4%, δηλαδή 302 από τις 5552 γυναίκες του δείγματος είχαν καρκίνο σε στάδιο N0, το 71,8%, δηλαδή 3987 από τις 5552 γυναίκες του δείγματος είχαν καρκίνο σε στάδιο N1, το 14%, δηλαδή 777 από τις 5552 γυναίκες του δείγματος είχαν καρκίνο σε στάδιο N2, το 8,8%, δηλαδή 486 από τις 5552 γυναίκες του δείγματος είχαν καρκίνο σε στάδιο N3. (Το γράμμα «N» συν ένας αριθμός από 0 έως 3 σημαίνει λεμφαδένες. Τις περισσότερες φορές, όσο είναι

περισσότεροι οι λεμφαδένες με καρκίνο, τόσο μεγαλύτερος είναι ο αριθμός που αποδίδεται.)

13. Status

		Status			Cumulative
		Frequency	Percent	Valid Percent	Percent
Valid	Alive	4995	90,0	90,0	90,0
	Dead	557	10,0	10,0	100,0
	Total	5552	100,0	100,0	

Πίνακας 2.17 : Συχνότητες ανά του αν κάποιος πέθανε από καρκίνο του μαστού.



Διάγραμμα 2.17 : Bar chart συχνοτήτων με βάση το αν κάποιος πέθανε από καρκίνο του μαστού.

Σύμφωνα με τα παραπάνω, συμπεραίνουμε ότι το 90%, δηλαδή 4995 από τις 5552 γυναίκες του δείγματος δεν πέθαναν από καρκίνο του μαστού ενώ το 10% δηλαδή 557 από τις 5552 πέθαναν από αυτή την αιτία.

ΚΕΦΑΛΑΙΟ 3

ΣΥΣΧΕΤΙΣΕΙΣ

3.1 Εισαγωγή

Σε αυτό το κεφάλαιο θα μελετήσουμε τις συσχετίσεις μεταξύ των μεταβλητών.

3.1.1 Έλεγχος χ^2

Διαφορές ή συσχέτιση :	Διαφορές
Κλίμακα μέτρησης :	Κατηγορική
Σχεδιασμός :	Ανεξάρτητα δείγματα
Σημειώσεις :	Τα δεδομένα πρέπει να έχουν την μορφή συχνοτήτων. Παρότι ενδιαφερόμαστε για τις διαφορές στην επίδραση της ανεξάρτητης μεταβλητής, το τεστ ουσιαστικά εξετάζει τη σχέση μεταξύ των κατηγοριών στις στήλες και τις γραμμές ενός πίνακα.

Πίνακας 3.1 : Προϋποθέσεις για τη χρήση του κριτηρίου.

Το στατιστικό κριτήριο χ^2 (χ -τετράγωνο - chi-square) είναι πιθανότατα η δοκιμασία που χρησιμοποιείται συχνότερα για τον έλεγχο των υποθέσεων των ερευνών που πραγματοποιούνται από τους επιστήμονες. Σε ένα βαθμό, το χ^2 είναι πολύ δημοφιλές λόγω της σχετικής ευκολίας με την οποία ικανοποιούνται οι προϋποθέσεις για τη χρήση του. Πρόκειται για ένα μη παραμετρικό κριτήριο που δεν απαιτεί καμία υπόθεση για την ακριβή μορφή της κατανομής του πληθυσμού.

Το χ^2 είναι το κατάλληλο κριτήριο για την περίπτωση κατά την οποία τα δεδομένα της έρευνάς μας είναι ποιοτικά (η κλίμακα μέτρησης που χρησιμοποιήθηκε ήταν κατηγορική).

Η δημοτικότητα του χ^2 σχετίζεται, επίσης, με τη χρησιμότητά του σε ένα ευρύτερο φάσμα ερευνητικών καταστάσεων. Έτσι, αν και το χ^2 μπορεί να χρησιμοποιηθεί στην απλή περίπτωση όπου συγκρίνονται οι συχνότητες δύο συνθηκών μιας ποιοτικής μεταβλητής, μπορεί, επίσης, να εφαρμοστεί και σε πιο περίπλοκες ερευνητικές συνθήκες με δύο ποιοτικές μεταβλητές.

Ένα σημείο που πρέπει να λαμβάνεται υπόψη όταν χρησιμοποιείται το χ^2 , είναι το εξής: Σε όλους τους υπόλοιπους πειραματικούς σχεδιασμούς που έχουν αναφερθεί μέχρι τώρα ο ερευνητής αποφασίζει πόσους συμμετέχοντες θα έχει σε κάθε ομάδα, οπότε ο αριθμός των μετρήσεων κάθε ομάδας είναι προκαθορισμένος. Ωστόσο, το χ^2 ελέγχει μία εναλλακτική υπόθεση η οποία προβλέπει πόσοι συμμετέχοντες κάθε

ομάδας θα βρεθούν σε συγκεκριμένες κατηγορίες. Επομένως, αυτό δεν μπορεί να αποφασιστεί εκ των προτέρων. Η συγκεκριμένη ιδιαιτερότητα επιβάλλει στον ερευνητή να χρησιμοποιήσει πολλούς συμμετέχοντες προκειμένου να εξασφαλίσει ότι ένας ικανός αριθμός από αυτούς θα βρεθεί στην κάθε κατηγορία. Έτσι, για να υπολογίσουμε τον ελάχιστο αριθμό συμμετεχόντων, μπορούμε να ακολουθήσουμε τον εξής απλό κανόνα: Για κάθε κατηγορία θα πρέπει να έχουμε περίπου 20 συμμετέχοντες. Αυτός είναι ο καλύτερος τρόπος για να εξασφαλίσουμε πως θα πληρείται μία από τις πιο σημαντικές προϋποθέσεις για την εφαρμογή του κριτηρίου αυτού.

Ο έλεγχος ανεξαρτησίας χ^2 (Chi-Square) ελέγχει την υπόθεση :

H_0 : Οι μεταβλητές είναι ανεξάρτητες
έναντι της εναλλακτικής

H_1 : Οι μεταβλητές δεν είναι ανεξάρτητες

Ο παραπάνω έλεγχος υποθέσεων αξιολογεί τη μηδενική υπόθεση, συγκρίνοντας τον αριθμό των υποκείμενων που αναμένεται σε κάθε ονομαστική κατηγορία με τον αριθμό των υποκείμενων που παρατηρείται ότι ανήκει σε κάθε ονομαστική κατηγορία, με βάση τις μετρήσεις του δείγματος. Ο στατιστικός δείκτης ελέγχου που χρησιμοποιείται για αυτή την αξιολόγηση είναι το χ^2 . Οι κρίσιμες τιμές για το στατιστικό δείκτη ελέγχου δίνονται από την κατανομή χ^2 , ενώ οι μαθηματικές σχέσεις υπολογισμού του είναι οι εξής:

$$\chi^2 = \sum \frac{(f_o - f_e)^2}{f_e},$$

όπου f_e και f_o η αναμενόμενη και η παρατηρούμενη συχνότητα,

$df = C - 1$, όπου C ο αριθμός των ονομαστικών κατηγοριών.

Αν και, όπως αναφέραμε, τα μη-παραμετρικά τεστ δε χρειάζεται να ικανοποιούν συγκεκριμένες προϋποθέσεις, τα συμπεράσματα που προκύπτουν από την εφαρμογή του στατιστικού τεστ χ^2 είναι πιο αξιόπιστα όταν χρησιμοποιείται αντιπροσωπευτικό δείγμα, και η αναμενόμενη συχνότητα όλων των κατηγοριών είναι μεγαλύτερη από πέντε.

Σε έναν πίνακα συνάφειας αν ένα κελί εμφανίσει αναμενόμενη συχνότητα μικρότερη του 5 (επομένως το 25% των αναμενόμενων συχνοτήτων), τότε σε αυτήν την περίπτωση αντί του ελέγχου χ^2 μπορούμε εναλλακτικά να χρησιμοποιήσουμε την ακριβή δοκιμασία του Fisher (Fisher's exact test).

3.1.2 Συντελεστής συσχέτισης του Pearson

Ο συντελεστής συσχέτισης Pearson r είναι ο κατάλληλος (παραμετρικός) στατιστικός δείκτης προκειμένου να αξιολογηθεί αν υπάρχει γραμμική σχέση μεταξύ δύο

μεταβλητών. Με τον όρο παραμετρικός αναφερόμαστε σε μια σειρά στατιστικών δεικτών που ικανοποιούν συγκεκριμένες προϋποθέσεις. Αν αυτές οι προϋποθέσεις παραβιάζονται, τότε γίνεται χρήση μη-παραμετρικών στατιστικών δεικτών, οι οποίοι δεν επηρεάζονται από την ισχύ ή όχι των προϋποθέσεων. Για τον συντελεστή συσχέτισης Pearson r οι προϋποθέσεις αυτές είναι:

(α) οι μεταβλητές πρέπει να είναι συνεχείς σε κλίμακα ίσων διαστημάτων και να κατανέμονται κανονικά &

$$r = \frac{\sum(X-\mu_X)(Y-\mu_Y)}{N\sigma_X\sigma_Y}$$

Οι τιμές X και Y αναφέρονται στις τιμές κάθε συμμετέχοντα, ενώ οι τιμές μ_X και μ_Y στους μέσους όρους των μεταβλητών X και Y . Επομένως η τιμή κάθε συμμετέχοντα στην μεταβλητή X αφαιρείται από το μέσο όρο της μεταβλητής και όλες οι τιμές προστίθενται. Το ίδιο γίνεται και για τη μεταβλητή Y . Το γεγονός ότι αυτά τα ζευγάρια τιμών πολλαπλασιάζονται μεταξύ τους εκφράζει τη “συνδιακόμευση” των δύο μεταβλητών.

Πρόκειται για έναν αριθμό μεταξύ του -1 και 1.

$r < 0.3$	Καμία ή πολύ ασθενής συσχέτιση
$0.3 < r < 0.5$	Ασθενής συσχέτιση
$0.5 < r < 0.7$	Μέτρια συσχέτιση
$0.7 < r < 1$	Ισχυρή συσχέτιση
$r = 1$	Απόλυτη συσχέτιση

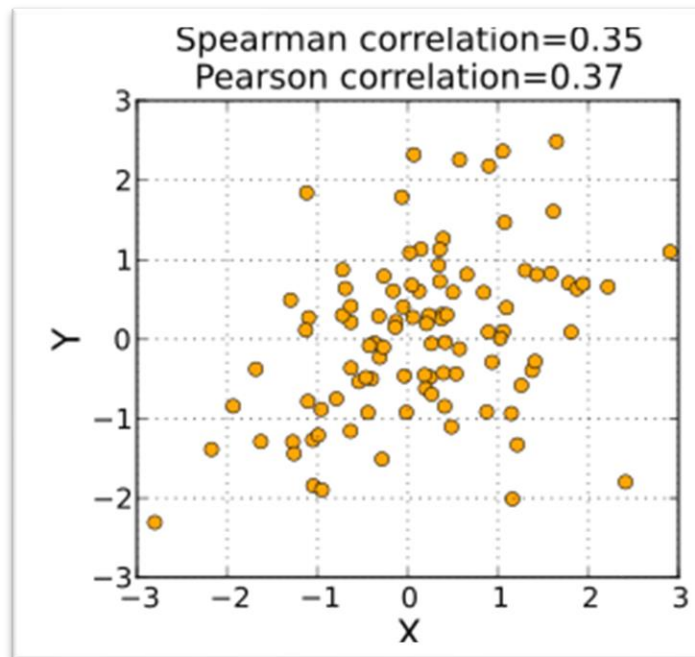
Πίνακας 3.2 : Μέγεθος δείκτη συσχέτισης του Pearson.

3.1.3 Ο συντελεστής συσχέτισης του Spearman

Στη στατιστική, ο συντελεστής συσχέτισης Spearman, είναι ένα μη-παραμετρικό μέτρο της στατιστικής εξάρτησης μεταξύ δύο μεταβλητών. Αξιολογεί το πόσο καλά μπορεί να περιγραφεί η σχέση μεταξύ των δύο μεταβλητών χρησιμοποιώντας μια μονότονη συνάρτηση. Εάν δεν υπάρχουν επαναλαμβανόμενες τιμές των δεδομένων, μια τέλεια συσχέτιση Spearman κατά +1 ή -1 συμβαίνει όταν κάθε μία από τις μεταβλητές είναι μια τέλεια μονότονη συνάρτηση της άλλης. Ο συντελεστής Spearman, όπως κάθε συντελεστής συσχέτισης, είναι κατάλληλος και για συνεχείς και για διακριτές μεταβλητές, συμπεριλαμβανομένων των τακτικών διακριτών μεταβλητών.

$$r_s = \frac{\sum_{i=1}^n (R_i - \bar{R})(S_i - \bar{S})}{\sqrt{\sum_{i=1}^n (R_i - \bar{R})^2} \sqrt{\sum_{i=1}^n (S_i - \bar{S})^2}}$$

όπου $\bar{R} = \frac{\sum_{i=1}^n R_i}{n}$ και $\bar{S} = \frac{\sum_{i=1}^n S_i}{n}$



Πίνακας 3.3 : Όταν τα δεδομένα είναι περίπου ελλειπτικά κατανομημένα και δεν υπάρχουν εμφανείς ακραίες τιμές, η συσχέτιση Spearman και Pearson δίνουν παρόμοιες τιμές.

Ο συντελεστής συσχέτισης του Spearman είναι επίσης, πιο «ευέλικτος» και στην ύπαρξη ακραίων τιμών, μπορεί να χρησιμοποιηθεί και για διατάξιμες μεταβλητές και δεν απαιτεί κάποια υπόθεση για τους πληθυσμούς.

3.1.4 Έλεγχος Mann-Whitney U

Έλεγχος του αν δύο ανεξάρτητα δείγματα διαφέρουν ή όχι (αν προέρχονται από τον ίδιο πληθυσμό). Η μηδενική υπόθεση H_0 είναι ότι οι διάμεσοι (medians) της μεταβλητής στα δύο δείγματα είναι ίσες. Πρόκειται για μη παραμετρική δοκιμασία, αντιστοιχη του t-test δύο ανεξάρτητων δειγμάτων.

1. Διατάσσονται κατά αύξουσα σειρά οι τιμές της μεταβλητής και των δύο δειγμάτων μαζί.
2. Υπολογίζεται το άθροισμα των θέσεων διάταξης των παρατηρήσεων που ανήκουν σε κάθε δείγμα. Αν δύο ή περισσότερες παρατηρήσεις έχουν ίδια τιμή, ως «διορθωμένη θέση διάταξης» επιλέγεται η μέση τιμή.

3. Υπολογίζονται οι τιμές του στατιστικού δείκτη U :

$$U_1 = R_1 - \frac{n_1(n_1+1)}{2} \text{ και } U_2 = R_2 - \frac{n_2(n_2+1)}{2}$$

4. Η μικρότερη τιμή (U_{\min}) των U_1 και U_2 , με τη βοήθεια πινάκων καθορίζει τη στατιστική σημαντικότητα της διαφοράς των δύο δειγμάτων. Δεδομένης της H_0 , η κατανομή του U είναι ομοιόμορφη με αντικατάσταση.

3.1.5 Kruskal Wallis

Ο έλεγχος των Kruskal- Wallis χρησιμοποιείται στην περίπτωση που θέλουμε να ελέγξουμε εάν τα δείγματα προέρχονται από κοινό πληθυσμό. Δηλαδή θέλουμε να ελέγξουμε αν οι πληθυσμοί ανήκουν στην ίδια κατανομή. Ο έλεγχος των Kruskal-Wallis υποθέτει ότι οι k πληθυσμοί έχουν κοινή διακύμανση.

Έστω k δείγματα μεγέθους n_i τα οποία προέρχονται από k πληθυσμούς με αθροιστικές συναρτήσεις κατανομών F_1, \dots, F_k . Μας ενδιαφέρει ο έλεγχος της υπόθεσης:

$$H_0: m_1 = \dots = m_k$$

έναντι της εναλλακτικής

$$H_1: \text{υπάρχει τουλάχιστον ένα ζεύγος } (i,j) \text{ } i \neq j \text{ τέτοιο ώστε } m_i \neq m_j$$

Καταρχήν για τον υπολογισμό της στατιστικής συνάρτησης τα δείγματα ενώνονται, το νέο δείγμα διατάσσεται κατά αύξουσα σειρά και υπολογίζονται οι τάξεις $R(X_{ij})$ των παρατηρήσεων. Στη συνέχεια υπολογίζονται τα αθροίσματα των τάξεων R_i για το κάθε δείγμα.

Μη ύπαρξη δεσμών :

Στην περίπτωση που δεν υπάρχουν δεσμοί η στατιστική συνάρτηση του ελέγχου είναι:

$$H = \frac{12}{(n+1)n} \sum_{i=1}^k R_i^2 - 3(n+1)$$

Η για δείγματα με $n_i \geq 5$ προσεγγίζει την χ^2 κατανομή με $df=k-1$ βαθμούς ελευθερίας. Επομένως, η κρίσιμη περιοχή του ελέγχου σε επίπεδο σημαντικότητας α είναι :

$$H \geq \chi^2_{k-1, 1-\alpha}$$

Υπαρξη Δεσμών :

Στην περίπτωση που υπάρχουν δεσμοί η στατιστική συνάρτηση που χρησιμοποιείται είναι:

$$H_d = \frac{H}{1 - \frac{\sum_{i=1}^t d_i(d_i^2-1)}{n(n^2-2)}}$$

Όπου t το πλήθος των διαφορετικών δεσμών και d_i το πλήθος των παρατηρήσεων του i -δεσμού.

3.2 Έλεγχος συσχέτισης ποσοτικών μεταβλητών

Εφαρμόζουμε έλεγχο κανονικότητας για να αποφασίσουμε ποιον συντελεστή συσχέτισης θα εφαρμόσουμε (συντελεστή συσχέτισης του Pearson αν τα δεδομένα μου ακολουθούν την κανονική κατανομή, αν όχι του Spearman).

Έλεγχος Κανονικότητας :

H_0 : Η κατανομή είναι κανονική κατανομή.
έναντι της εναλλακτικής

H_1 : Η κατανομή δεν είναι κανονική κατανομή.

One-Sample Kolmogorov-Smirnov Test

		YearOfDiagnosis	Tumor_Size (mm)-new	RegionalNodesE xamined	RegionalNodesP ositive	Survival_Months
N		5552	5552	5552	5552	5552
Normal Parameters ^{a,b}	Mean	2012,53	30,59	12,80	9,25	32,72
	Std. Deviation	1,701	19,066	15,597	23,457	20,332
Most Extreme Differences	Absolute	,149	,146	,206	,379	,075
	Positive	,143	,146	,165	,379	,075
	Negative	-,149	-,096	-,206	-,347	-,059
Test Statistic		,149	,146	,206	,379	,075
Asymp. Sig. (2-tailed)		,000 ^c	,000 ^c	,000 ^c	,000 ^c	,000 ^c

a. Test distribution is Normal.

b. Calculated from data.

c. Lilliefors Significance Correction.

Πίνακας 3.4 :Έλεγχος κανονικότητας των σφαλμάτων.

Παρατηρούμε πως το p -value=0,000 σε όλες τις μεταβλητές, συνεπώς σε επίπεδο σημαντικότητας 5% απορρίπτουμε την μηδενική μας υπόθεση. Αυτό σημαίνει πως οι ποσοτικές μεταβλητές δεν ακολουθούν την κανονική κατανομή.

Σύμφωνα με τα παραπάνω καταλήγουμε στο συμπέρασμα ότι πρέπει να χρησιμοποιήσουμε τον συντελεστή συσχέτισης του Spearman για να βρούμε ποιες μεταβλητές συνδέονται μεταξύ τους και πόσο ισχυρά.

Συντελεστής συσχέτισης Spearman μεταξύ του χρόνου επιβίωσης από την διάγνωση και των υπόλοιπων μεταβλητών :

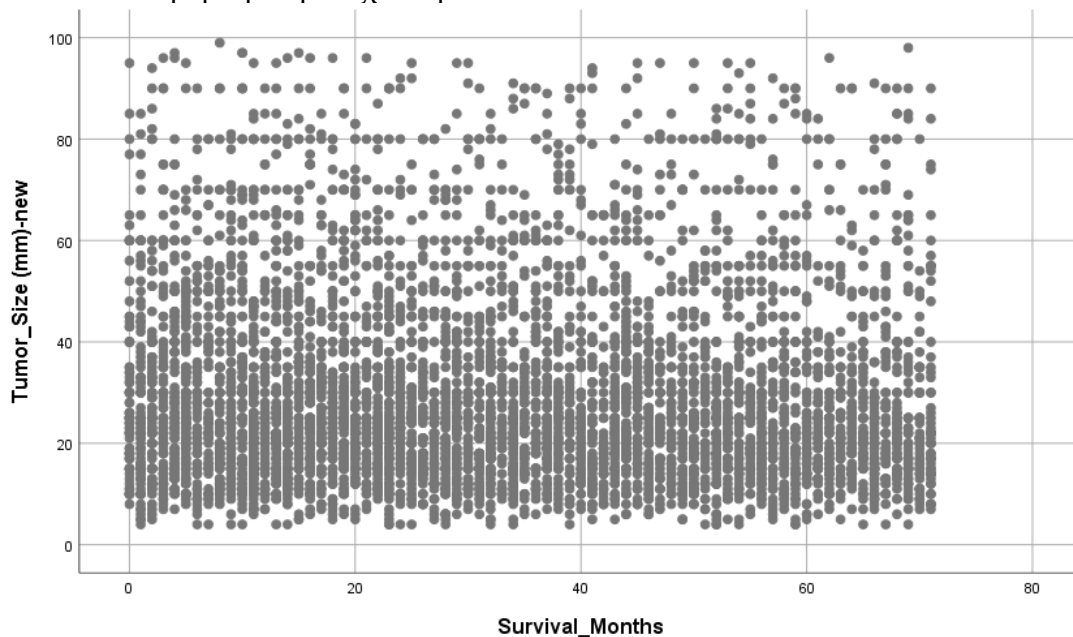
H₀: Δεν υπάρχει στατιστικά σημαντική σχέση μεταξύ των μεταβλητών.
έναντι της εναλλακτικής
H₁: Υπάρχει στατιστικά σημαντική σχέση μεταξύ των μεταβλητών.

			Correlations				
			YearOfDiagnosis	Tumor_Size (mm)-new	RegionalNodesExamined	RegionalNodesPositive	Survival_Months
Spearman's rho	YearOfDiagnosis	Correlation Coefficient	1,000	,035**	-,127**	-,020	-,891**
		Sig. (2-tailed)	.	,010	,000	,143	,000
		N	5552	5552	5552	5552	5552
	Tumor_Size (mm)-new	Correlation Coefficient	,035**	1,000	,198**	,206**	-,060**
		Sig. (2-tailed)	,010	.	,000	,000	,000
		N	5552	5552	5552	5552	5552
	RegionalNodesExamined	Correlation Coefficient	-,127**	,198**	1,000	,567**	,115**
		Sig. (2-tailed)	,000	,000	.	,000	,000
		N	5552	5552	5552	5552	5552
	RegionalNodesPositive	Correlation Coefficient	-,020	,206**	,567**	1,000	,004
		Sig. (2-tailed)	,143	,000	,000	.	,762
		N	5552	5552	5552	5552	5552
	Survival_Months	Correlation Coefficient	-,891**	-,060**	,115**	,004	1,000
		Sig. (2-tailed)	,000	,000	,000	,762	.
		N	5552	5552	5552	5552	5552

** . Correlation is significant at the 0.01 level (2-tailed).

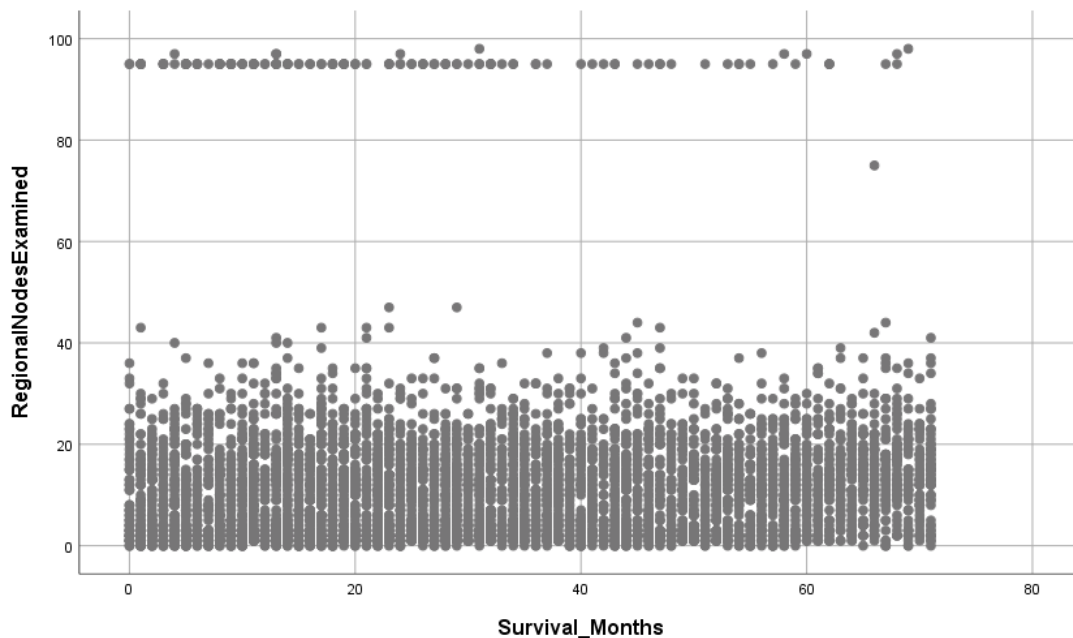
Πίνακας 3.5 : Έλεγχος συσχετίσεων με τη μέθοδο του Spearman.

- Σύμφωνα με τον 3.5 πίνακα το $p\text{-value}=0,000 < 0,05$, συνεπώς απορρίπτουμε τη μηδενική υπόθεση, δηλαδή υπάρχει στατιστικά σημαντική σχέση μεταξύ των μεταβλητών Survival_Months και Tumor_Size. Ο συντελεστής συσχέτισης του Spearman για το μέγεθος του όγκου και τους μήνες επιβίωσης είναι $-0,060$. Αυτό σημαίνει ότι οι μεταβλητές Tumor_Size και Survival_Months έχουν πολύ ασθενή αρνητική συσχέτιση.



Διάγραμμα 3.1 : Διάγραμμα διασποράς μεταξύ των μεταβλητών "Tumor_Size" και "Survival_Months".

- Σύμφωνα με τον 3.5 πίνακα το $p\text{-value}=0,000 < 0,05$, συνεπώς απορρίπτουμε τη μηδενική υπόθεση, δηλαδή υπάρχει στατιστικά σημαντική σχέση μεταξύ των μεταβλητών *Survival_Months* και *Regional_Nodes_Examined*. Ο συντελεστής συσχέτισης του Spearman για το πλήθος των περιφερειακών όγκων που εξετάστηκαν και τους μήνες επιβίωσης είναι 0,115. Αυτό σημαίνει ότι οι μεταβλητές *Regional_Nodes_Examined* και *Survival_Months* έχουν πολύ ασθενή θετική συσχέτιση.



Διάγραμμα 3.2 : Διάγραμμα διασποράς μεταξύ των μεταβλητών “ *Regional_Nodes_Examined*” και “*Survival_Months*”.

- Σύμφωνα με τον 3.5 πίνακα $p\text{-value}=0,762 > 0,05$, συνεπώς δεχόμαστε τη μηδενική υπόθεση, δηλαδή δεν υπάρχει στατιστικά σημαντική σχέση μεταξύ των μεταβλητών *Survival_Months* και *Regional_Nodes_Positive*.

3.3 Έλεγχος συσχέτισης ποιοτικών μεταβλητών

Σε αυτή την παράγραφο θα εφαρμόσουμε έλεγχο συσχέτισης για ποιοτικές μεταβλητές. Έλεγχος υποθέσεων:

H_0 : Οι μεταβλητές δεν είναι συσχετισμένες
έναντι της εναλλακτικής

H_1 : Οι μεταβλητές είναι συσχετισμένες

Συσχέτιση θνησιμότητας από τον καρκίνο του μαστού με ηλικία :

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Status * Age	5552	100,0%	0	0,0%	5552	100,0%

Πίνακας 3.6 : Περίληψη επεξεργασίας υπόθεσης «Συσχέτιση θνησιμότητας από τον καρκίνο του μαστού με ηλικία».

			Age													Total	
			20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	65-69	70-74	75-79	80-84		85+
Status	Alive	Count	2	40	118	230	441	609	676	621	702	605	404	255	174	118	4995
		Expected Count	2,7	40,5	117,0	216,8	418,3	586,6	655,0	599,2	693,7	586,6	404,0	287,9	201,5	185,3	4995,0
Dead	Count	1	5	12	11	24	43	52	45	69	47	45	65	50	88	557	
	Expected Count	,3	4,5	13,0	24,2	46,7	65,4	73,0	66,8	77,3	65,4	45,0	32,1	22,5	20,7	557,0	
Total	Count	3	45	130	241	465	652	728	666	771	652	449	320	224	206	5552	
	Expected Count	3,0	45,0	130,0	241,0	465,0	652,0	728,0	666,0	771,0	652,0	449,0	320,0	224,0	206,0	5552,0	

Πίνακας 3.7 : «Διασταυρώσεις» των μεταβλητών Status και Age.

Chi-Square Tests

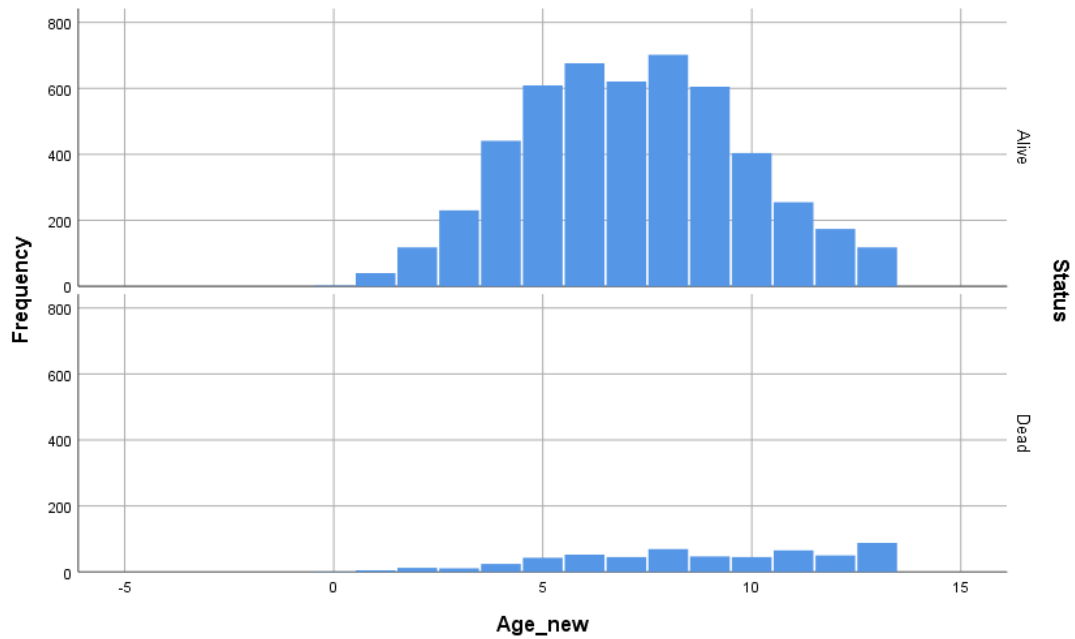
	Value	df	Asymptotic Significance (2- sided)	Exact Sig. (2- sided)
Pearson Chi-Square	370,895 ^a	13	,000	. ^b
Likelihood Ratio	266,103	13	,000	. ^b
Fisher's Exact Test	. ^b			. ^b
N of Valid Cases	5552			

a. 3 cells (10,7%) have expected count less than 5. The minimum expected count is ,30.

b. Cannot be computed because there is insufficient memory.

Πίνακας 3.8 : Έλεγχος χ^2 για τις μεταβλητές Status και Age.

Ο έλεγχος ανεξαρτησίας $\chi^2=370,895$ του Pearson έχει $p\text{-value}=0,000 < 0,05$, συνεπώς απορρίπτουμε τη μηδενική υπόθεση του ελέγχου για το αν η ηλικία συσχετίζεται με την θνησιμότητα από τον καρκίνο του μαστού. Άρα η ηλικία σχετίζεται με την θνησιμότητα από καρκίνο του μαστού.



Διάγραμμα 3.3 : Ραβδόγραμμα μεταξύ των μεταβλητών “Age” και “Status”.

Όπου:

45-49	5
35-39	3
65-69	9
40-44	4
55-59	7
80-84	12
25-29	1
20-24	0
50-54	6
70-74	10
75-79	11
60-64	8
85+	13
30-34	2

Συσχέτιση θνησιμότητας από τον καρκίνο του μαστού με φυλή :

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Status * Race	5552	100,0%	0	0,0%	5552	100,0%

Πίνακας 3.9 : Περίληψη επεξεργασίας υπόθεσης «Συσχέτιση θνησιμότητας από τον καρκίνο του μαστού με φυλή».

Status * Race Crosstabulation

Status			Race				Total
			Black	Other (American Indian/AK Native, Asian/Pacific Islander)	Unknown	White	
Alive	Count		198	514	14	4269	4995
	Expected Count		202,4	502,0	12,6	4278,0	4995,0
Dead	Count		27	44	0	486	557
	Expected Count		22,6	56,0	1,4	477,0	557,0
Total	Count		225	558	14	4755	5552
	Expected Count		225,0	558,0	14,0	4755,0	5552,0

Πίνακας 3.10 : «Διασταυρώσεις» των μεταβλητών Status και Race.

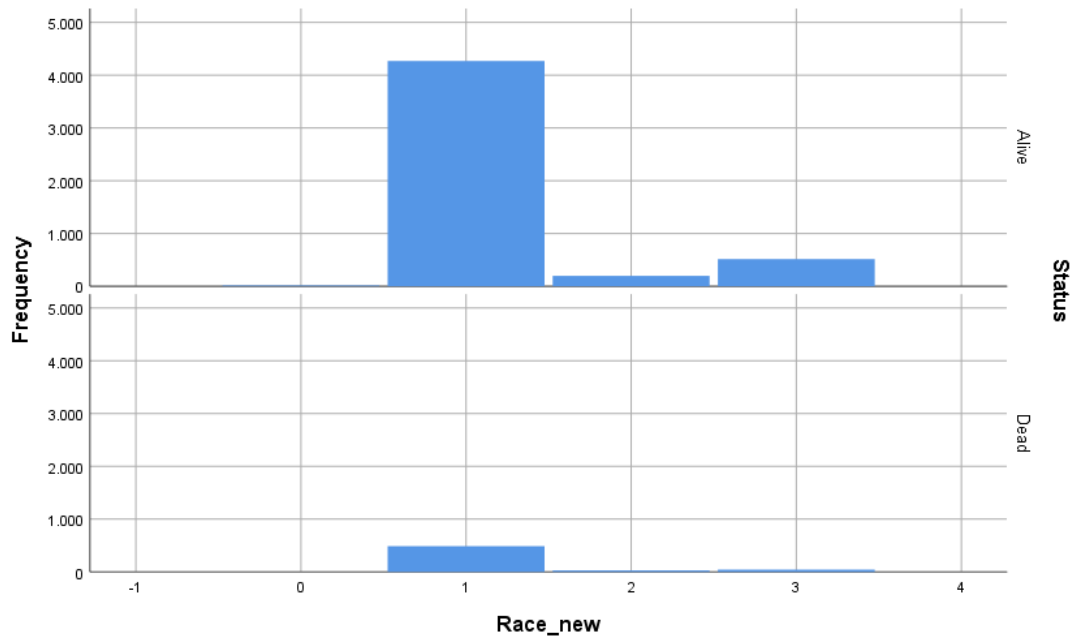
Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)
Pearson Chi-Square	5,563 ^a	3	,135	,132
Likelihood Ratio	7,113	3	,068	,071
Fisher's Exact Test	5,008			,154
N of Valid Cases	5552			

a. 1 cells (12,5%) have expected count less than 5. The minimum expected count is 1,40.

Πίνακας 3.11 : Έλεγχος χ^2 για τις μεταβλητές Status και Race..

Ο έλεγχος ανεξαρτησίας $\chi^2 = 5,563$ του Pearson έχει $p\text{-value} = 0,135 > 0,05$, συνεπώς δεχόμαστε τη μηδενική υπόθεση του ελέγχου για το αν η φυλή ενός ανθρώπου συσχετίζεται με την θνησιμότητα από τον καρκίνο του μαστού. Άρα η φυλή δεν συσχετίζεται με τη θνησιμότητα από καρκίνο του μαστού.



Διάγραμμα 3.4 : Ραβδόγραμμα μεταξύ των μεταβλητών "Race" και "Status".

Όπου :

unknown	0
white	1
black	2
other	3

Συσχέτιση θνησιμότητας από τον καρκίνο του μαστού με οικογενειακή κατάσταση :

	Case Processing Summary					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Status * MaritalStatus	5552	100,0%	0	0,0%	5552	100,0%

Πίνακας 3.12 : Περίληψη επεξεργασίας υπόθεσης «Συσχέτιση θνησιμότητας από τον καρκίνο του μαστού με οικογενειακή κατάσταση».

Status * MaritalStatus Crosstabulation

		MaritalStatus						Total	
		Divorced	Married (including common law)	Separated	Single (never married)	Unmarried or Domestic Partner	Widowed		
Status	Alive	Count	624	3127	72	629	54	489	4995
		Expected Count	615,4	3042,7	72,0	645,1	52,2	567,7	4995,0
	Dead	Count	60	255	8	88	4	142	557
		Expected Count	68,6	339,3	8,0	71,9	5,8	63,3	557,0
Total		Count	684	3382	80	717	58	631	5552
		Expected Count	684,0	3382,0	80,0	717,0	58,0	631,0	5552,0

Πίνακας 3.13 : «Διασταυρώσεις» των μεταβλητών Status και Marital_Status.

Chi-Square Tests

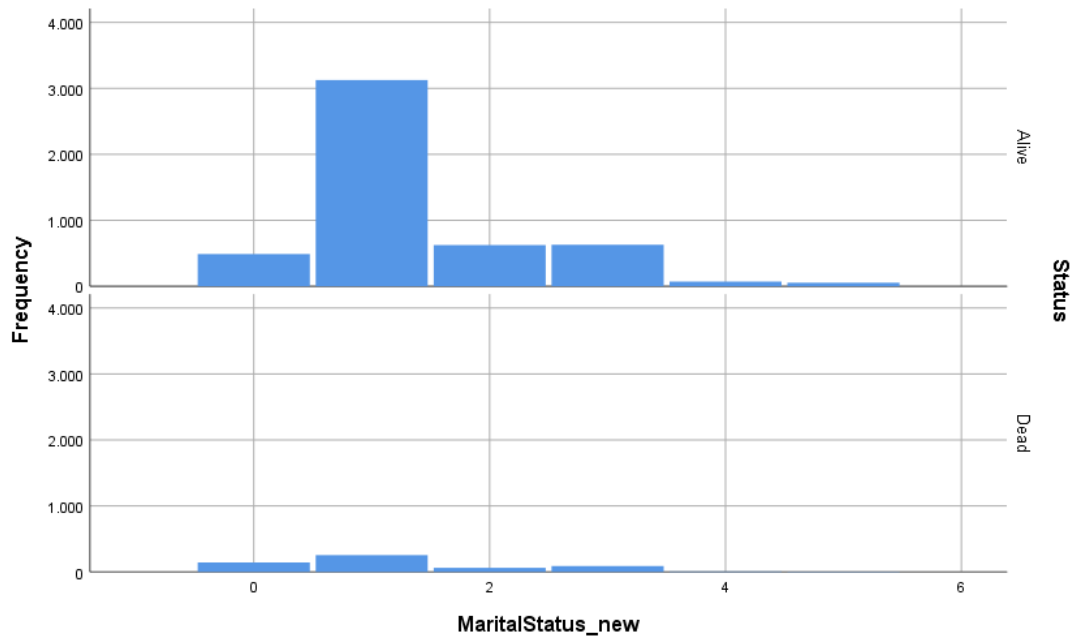
	Value	df	Asymptotic Significance (2- sided)	Exact Sig. (2- sided)
Pearson Chi-Square	137,841 ^a	5	,000	. ^b
Likelihood Ratio	114,446	5	,000	. ^b
Fisher's Exact Test	. ^b			. ^b
N of Valid Cases	5552			

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 5,82.

b. Cannot be computed because there is insufficient memory.

Πίνακας 3.14 : Έλεγχος χ^2 για τις μεταβλητές Status και Marital_Status.

Ο έλεγχος ανεξαρτησίας $\chi^2=137,841$ του Pearson έχει $p\text{-value}=0,000 < 0,05$, συνεπώς απορρίπτουμε τη μηδενική υπόθεση του ελέγχου για το αν η οικογενειακή κατάσταση ενός ανθρώπου συσχετίζεται με την θνησιμότητα από τον καρκίνο του μαστού. Άρα η οικογενειακή κατάσταση σχετίζεται με τη θνησιμότητα από καρκίνο του μαστού.



Διάγραμμα 3.5 : Ραβδόγραμμα μεταξύ των μεταβλητών “Marital_Status” και “Status”.

Όπου :

Widowed	0
Separated	4
Unmarried or Domestic Partner	5
Divorced	2
Married (including common law)	1
Single (never married)	3

Συσχέτιση θνησιμότητας από τον καρκίνο του μαστού με την μέτρηση για το κατά πόσο τα καρκινικά κύτταρα μοιάζουν με τα κανονικά :

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Status * Grade	5552	100,0%	0	0,0%	5552	100,0%

Πίνακας 3.15 : Περίληψη επεξεργασίας υπόθεσης «Συσχέτιση θνησιμότητας από τον καρκίνο του μαστού με την μέτρηση για το κατά πόσο τα καρκινικά κύτταρα μοιάζουν με τα κανονικά».

Status * Grade Crosstabulation

Status			Grade			Total
			Moderately differentiated; Grade II	Poorly differentiated; Grade III	Well differentiated; Grade I	
Alive	Count		2322	1992	681	4995
	Expected Count		2249,2	2092,6	653,2	4995,0
Dead	Count		178	334	45	557
	Expected Count		250,8	233,4	72,8	557,0
Total	Count		2500	2326	726	5552
	Expected Count		2500,0	2326,0	726,0	5552,0

Πίνακας 3.16 : «Διασταυρώσεις» των μεταβλητών Status και Grade.

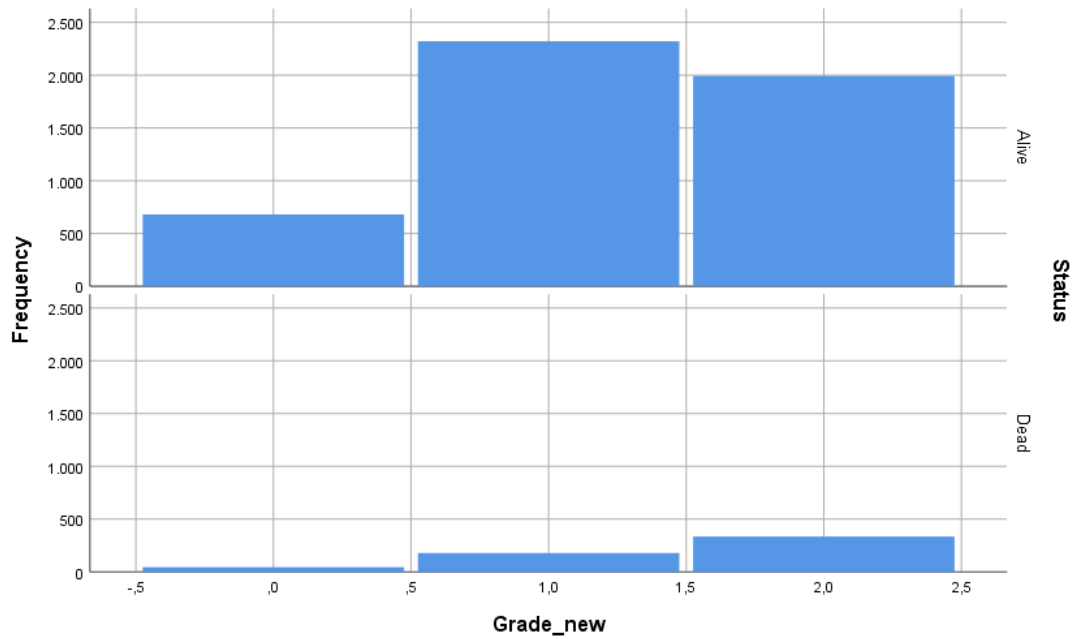
Chi-Square Tests

	Value	df	Asymptotic Significance (2- sided)	Exact Sig. (2- sided)
Pearson Chi-Square	83,567 ^a	2	,000	,000
Likelihood Ratio	82,540	2	,000	,000
Fisher's Exact Test	82,103			,000
N of Valid Cases	5552			

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 72,84.

Πίνακας 3.17 : Έλεγχος χ^2 για τις μεταβλητές Status και Grade.

Ο έλεγχος ανεξαρτησίας $\chi^2 = 83,567$ του Pearson έχει $p\text{-value} = 0,000 < 0,05$, συνεπώς απορρίπτουμε τη μηδενική υπόθεση του ελέγχου για το αν η παραπάνω μέτρηση συσχετίζεται με την θνησιμότητα από τον καρκίνο του μαστού. Άρα η μέτρηση αυτή σχετίζεται με τη θνησιμότητα από καρκίνο του μαστού.



Διάγραμμα 3.6 : Ραβδόγραμμα μεταξύ των μεταβλητών “Grade” και “Status”.

Όπου :

Moderately differentiated; Grade II	1
Poorly differentiated; Grade III	2
Well differentiated; Grade I	0

Συσχέτιση θνησιμότητας από τον καρκίνο του μαστού με την προέλευση πρωτογενούς :

Case Processing Summary

	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Status * Laterality	5552	100,0%	0	0,0%	5552	100,0%

Πίνακας 3.18 : Περίληψη επεξεργασίας υπόθεσης «Συσχέτιση θνησιμότητας από τον καρκίνο του μαστού με την προέλευση πρωτογενούς».

Status * Laterality Crosstabulation

Status			Laterality		Total
			Left - origin of primary	Right - origin of primary	
Alive	Count		2551	2444	4995
	Expected Count		2556,9	2438,1	4995,0
Dead	Count		291	266	557
	Expected Count		285,1	271,9	557,0
Total	Count		2842	2710	5552
	Expected Count		2842,0	2710,0	5552,0

Πίνακας 3.19 : «Διασταυρώσεις» των μεταβλητών Status και Laterality.

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	,276 ^a	1	,599	,623	,315
Continuity Correction ^b	,231	1	,631		
Likelihood Ratio	,276	1	,599	,623	,315
Fisher's Exact Test				,623	,315
N of Valid Cases	5552				

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 271,88.

b. Computed only for a 2x2 table

Πίνακας 3.20 : Έλεγχος χ^2 για τις μεταβλητές Status και Laterality.

Ο έλεγχος ανεξαρτησίας $\chi^2 = 0,276$ του Pearson έχει $p\text{-value} = 0,599 > 0,05$, συνεπώς δεχόμαστε τη μηδενική υπόθεση του ελέγχου για το αν η προέλευση πρωτογενούς συσχετίζεται με την θνησιμότητα από τον καρκίνο του μαστού. Άρα η προέλευση πρωτογενούς δεν σχετίζεται με τη θνησιμότητα από καρκίνο του μαστού.

Συσχέτιση θνησιμότητας από τον καρκίνο του μαστού με την έκταση της νόσου :

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Status * StageA	5552	100,0%	0	0,0%	5552	100,0%

Πίνακας 3.21 : Περίληψη επεξεργασίας υπόθεσης «Συσχέτιση θνησιμότητας από τον καρκίνο του μαστού με την έκταση της νόσου».

Status * StageA Crosstabulation

		StageA		Total	
		Distant	Regional		
Status	Alive	Count	206	4789	4995
		Expected Count	241,1	4753,9	4995,0
	Dead	Count	62	495	557
		Expected Count	26,9	530,1	557,0
Total		Count	268	5284	5552
		Expected Count	268,0	5284,0	5552,0

Πίνακας 3.22 : «Διασταυρώσεις» των μεταβλητών Status και StageA.

Chi-Square Tests

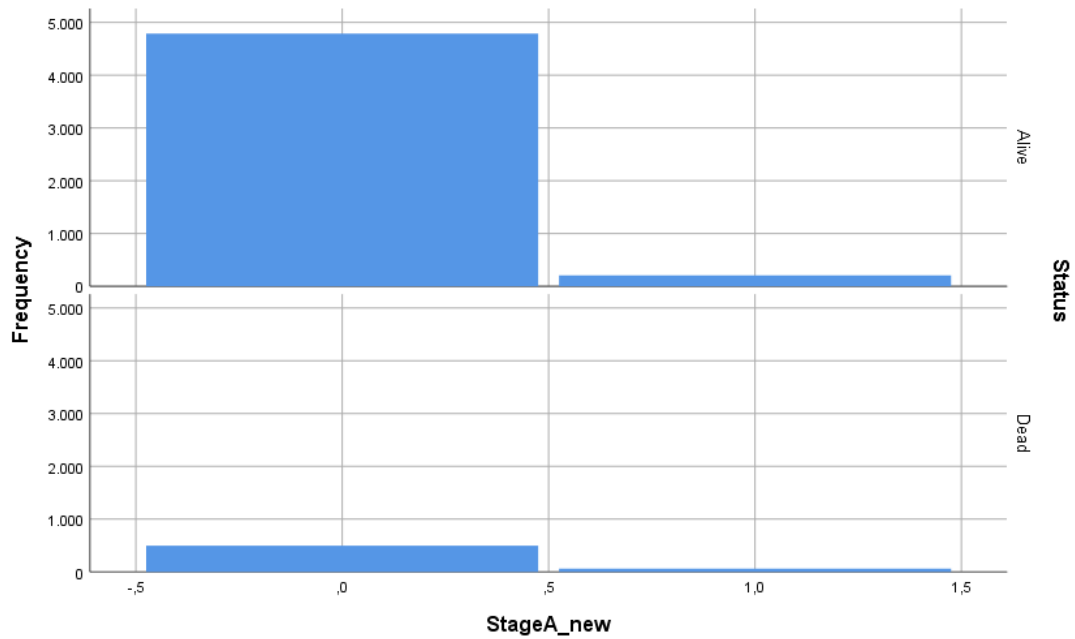
	Value	df	Asymptotic Significance (2- sided)	Exact Sig. (2- sided)	Exact Sig. (1- sided)
Pearson Chi-Square	53,555 ^a	1	,000	,000	,000
Continuity Correction ^b	52,040	1	,000		
Likelihood Ratio	41,394	1	,000	,000	,000
Fisher's Exact Test				,000	,000
N of Valid Cases	5552				

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 26,89.

b. Computed only for a 2x2 table

Πίνακας 3.23 : Έλεγχος χ^2 για τις μεταβλητές Status και StageA.

Ο έλεγχος ανεξαρτησίας $\chi^2 = 53,555$ του Pearson έχει $p\text{-value} = 0,000 < 0,05$, συνεπώς απορρίπτουμε τη μηδενική υπόθεση του ελέγχου για το αν η έκταση της νόσου συσχετίζεται με την θνησιμότητα από τον καρκίνο του μαστού. Άρα η έκταση της νόσου σχετίζεται με τη θνησιμότητα από καρκίνο του μαστού.



Διάγραμμα 3.7 : Ραβδόγραμμα μεταξύ των μεταβλητών “StageA” και “Status”.

Όπου :

Regional 0
Distant 1

Συσχέτιση θνησιμότητας από τον καρκίνο του μαστού με τους υποδοχείς οιστρογόνου :

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Status * ER_Status	5552	100,0%	0	0,0%	5552	100,0%

Πίνακας 3.24 : Περίληψη επεξεργασίας υπόθεσης «Συσχέτιση θνησιμότητας από τον καρκίνο του μαστού με τους υποδοχείς οιστρογόνου».

Status * ER_Status Crosstabulation

		ER_Status		Total	
		Negative	Positive		
Status	Alive	Count	718	4277	4995
		Expected Count	813,3	4181,7	4995,0
	Dead	Count	186	371	557
		Expected Count	90,7	466,3	557,0
Total		Count	904	4648	5552
		Expected Count	904,0	4648,0	5552,0

Πίνακας 3.25 : «Διασταυρώσεις» των μεταβλητών Status και ER_Status.

Chi-Square Tests

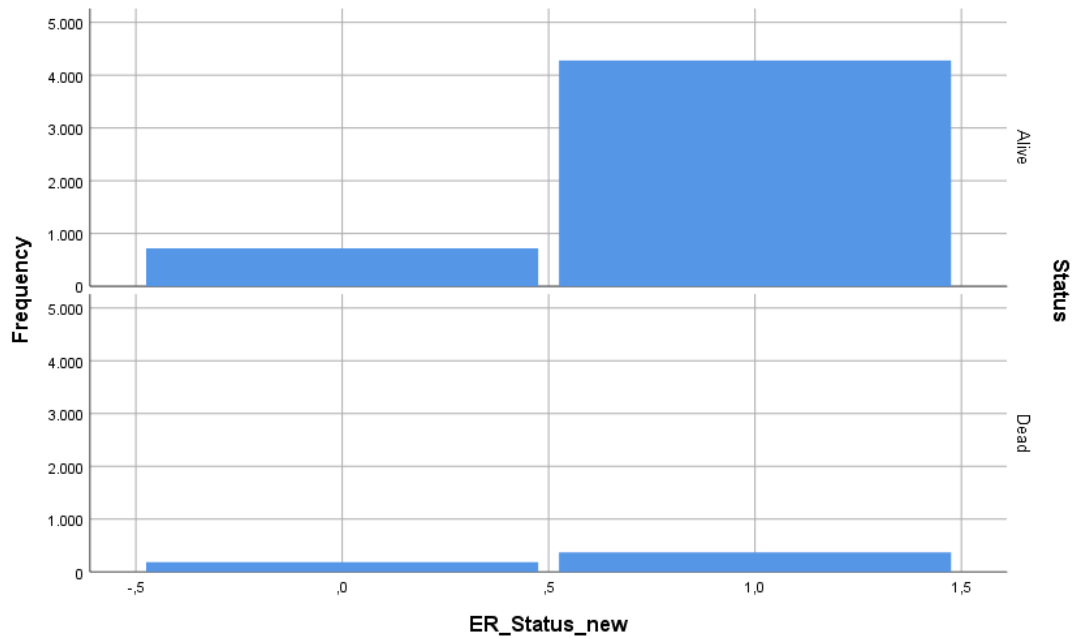
	Value	df	Asymptotic Significance (2- sided)	Exact Sig. (2- sided)	Exact Sig. (1- sided)
Pearson Chi-Square	132,976 ^a	1	,000	,000	,000
Continuity Correction ^b	131,584	1	,000		
Likelihood Ratio	111,331	1	,000	,000	,000
Fisher's Exact Test				,000	,000
N of Valid Cases	5552				

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 90,69.

b. Computed only for a 2x2 table

Πίνακας 3.26 : Έλεγχος χ^2 για τις μεταβλητές Status και ER_Status.

Ο έλεγχος ανεξαρτησίας $\chi^2 = 132,976$ του Pearson έχει $p\text{-value} = 0,000 < 0,05$, συνεπώς απορρίπτουμε τη μηδενική υπόθεση του ελέγχου για το αν οι υποδοχείς οιστρογόνου συσχετίζονται με την θνησιμότητα από τον καρκίνο του μαστού. Άρα οι υποδοχείς οιστρογόνου σχετίζονται με τη θνησιμότητα από καρκίνο του μαστού.



Διάγραμμα 3.8 : Ραβδόγραμμα μεταξύ των μεταβλητών “ER_Status” και “Status”.

Όπου :

Positive	1
Negative	0

Συσχέτιση θνησιμότητας από τον καρκίνο του μαστού με τους υποδοχείς προγεστερόνης :

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Status * PR_Status	5552	100,0%	0	0,0%	5552	100,0%

Πίνακας 3.27 : Περίληψη επεξεργασίας υπόθεσης «Συσχέτιση θνησιμότητας από τον καρκίνο του μαστού με τους υποδοχείς προγεστερόνης».

Status * PR_Status Crosstabulation

		PR_Status		Total	
		Negative	Positive		
Status	Alive	Count	1133	3862	4995
		Expected Count	1249,6	3745,4	4995,0
Status	Dead	Count	256	301	557
		Expected Count	139,4	417,6	557,0
Total	Count		1389	4163	5552
	Expected Count		1389,0	4163,0	5552,0

Πίνακας 3.28 : «Διασταυρώσεις» των μεταβλητών Status και PR_Status.

Chi-Square Tests

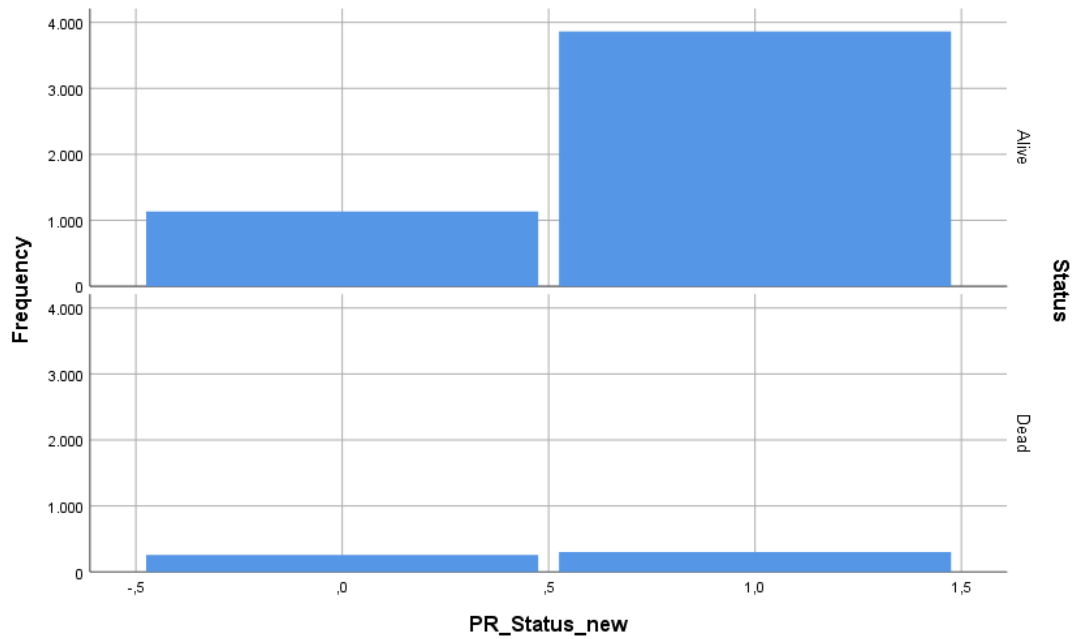
	Value	df	Asymptotic Significance (2- sided)	Exact Sig. (2- sided)	Exact Sig. (1- sided)
Pearson Chi-Square	144,749 ^a	1	,000	,000	,000
Continuity Correction ^b	143,511	1	,000		
Likelihood Ratio	129,057	1	,000	,000	,000
Fisher's Exact Test				,000	,000
N of Valid Cases	5552				

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 139,35.

b. Computed only for a 2x2 table

Πίνακας 3.29 : Έλεγχος χ^2 για τις μεταβλητές Status και PR_Status.

Ο έλεγχος ανεξαρτησίας $\chi^2 = 144,749$ του Pearson έχει $p\text{-value} = 0,000 < 0,05$, συνεπώς απορρίπτουμε τη μηδενική υπόθεση του ελέγχου για το αν οι υποδοχείς προγεστερόνης συσχετίζονται με την θνησιμότητα από τον καρκίνο του μαστού. Άρα οι υποδοχείς προγεστερόνης σχετίζονται με τη θνησιμότητα από καρκίνο του μαστού.



Διάγραμμα 3.9 : Ραβδόγραμμα μεταξύ των μεταβλητών “PR_Status” και “Status”.

Όπου :

Positive 1
 Negative 0

Συσχέτιση θνησιμότητας από τον καρκίνο του μαστού με τον δευτερεύον τύπο μαστού :

Case Processing Summary

	Valid		Cases Missing		Total	
	N	Percent	N	Percent	N	Percent
Status * BreastSubtype	5552	100,0%	0	0,0%	5552	100,0%

Πίνακας 3.30 : Περίληψη επεξεργασίας υπόθεσης «Συσχέτιση θνησιμότητας από τον καρκίνο του μαστού με τον δευτερεύον τύπο μαστού».

Status * BreastSubtype Crosstabulation

		BreastSubtype					
		Her2-/HR+	Her2+/HR-	Her2+/HR+	Triple Negative	Total	
Status	Alive	Count	3715	251	623	406	4995
		Expected Count	3647,3	251,0	602,8	493,9	4995,0
Status	Dead	Count	339	28	47	143	557
		Expected Count	406,7	28,0	67,2	55,1	557,0
Total		Count	4054	279	670	549	5552
		Expected Count	4054,0	279,0	670,0	549,0	5552,0

Πίνακας 3.31 : «Διασταυρώσεις» των μεταβλητών Status και BreastSubtype.

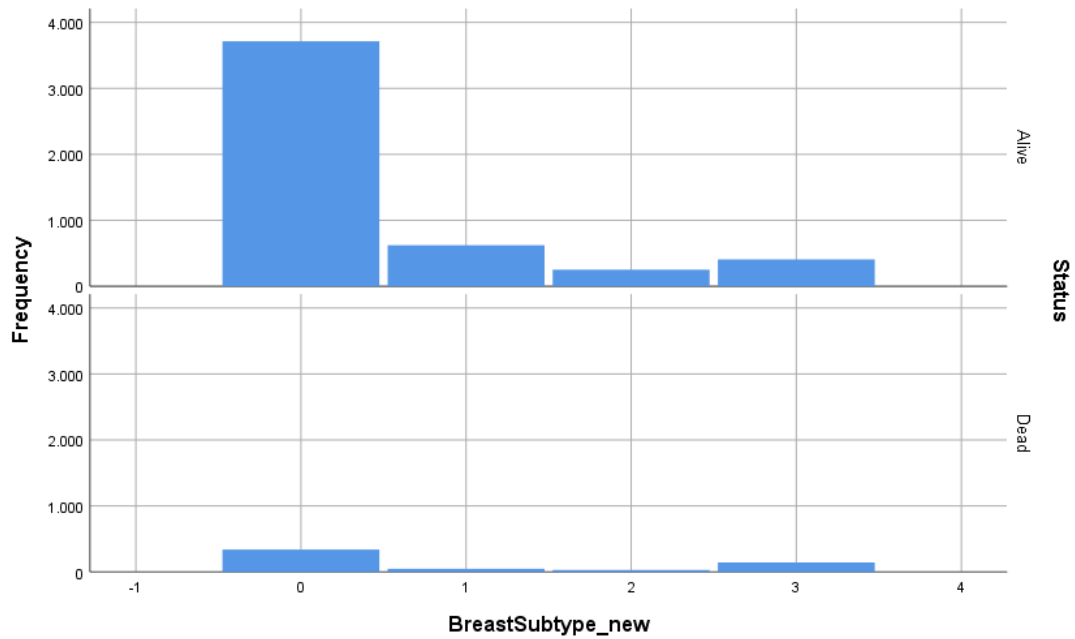
Chi-Square Tests

	Value	df	Asymptotic Significance (2- sided)	Exact Sig. (2- sided)
Pearson Chi-Square	175,292 ^a	3	,000	,000
Likelihood Ratio	134,380	3	,000	,000
Fisher's Exact Test	134,661			,000
N of Valid Cases	5552			

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 27,99.

Πίνακας 3.32 : Έλεγχος χ^2 για τις μεταβλητές Status και BreastSubtype.

Ο έλεγχος ανεξαρτησίας $\chi^2 = 175,292$ του Pearson έχει $p\text{-value} = 0,000 < 0,05$, συνεπώς απορρίπτουμε τη μηδενική υπόθεση του ελέγχου για το αν ο δευτερεύον τύπος μαστού συσχετίζεται με την θνησιμότητα από τον καρκίνο του μαστού. Άρα ο δευτερεύον τύπος μαστού σχετίζεται με τη θνησιμότητα από καρκίνο του μαστού.



Διάγραμμα 3.10 : Ραβδόγραμμα μεταξύ των μεταβλητών “BreastSubtype” και “Status”.

Όπου :

Her2-/HR+	0
Her2+/HR-	2
Her2+/HR+	1
Triple Negative	3

Συσχέτιση θνησιμότητας από τον καρκίνο του μαστού με το 6^ο στάδιο καρκίνου :

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Status * 6th_Stage	5552	100,0%	0	0,0%	5552	100,0%

Πίνακας 3.33 : Περίληψη επεξεργασίας υπόθεσης «Συσχέτιση θνησιμότητας από τον καρκίνο του μαστού με το 6^ο στάδιο».

Status * 6th_Stage Crosstabulation

			6th_Stage					Total	
			I	IIA	IIB	IIIA	IIIB	IIIC	
Status	Alive	Count	53	1595	1713	1025	220	389	4995
		Expected Count	54,9	1522,3	1684,2	1039,1	257,3	437,2	4995,0
	Dead	Count	8	97	159	130	66	97	557
		Expected Count	6,1	169,7	187,8	115,9	28,7	48,8	557,0
Total	Count		61	1692	1872	1155	286	486	5552
	Expected Count		61,0	1692,0	1872,0	1155,0	286,0	486,0	5552,0

Πίνακας 3.34 : «Διασταυρώσεις» των μεταβλητών Status και 6th Stage.

Chi-Square Tests

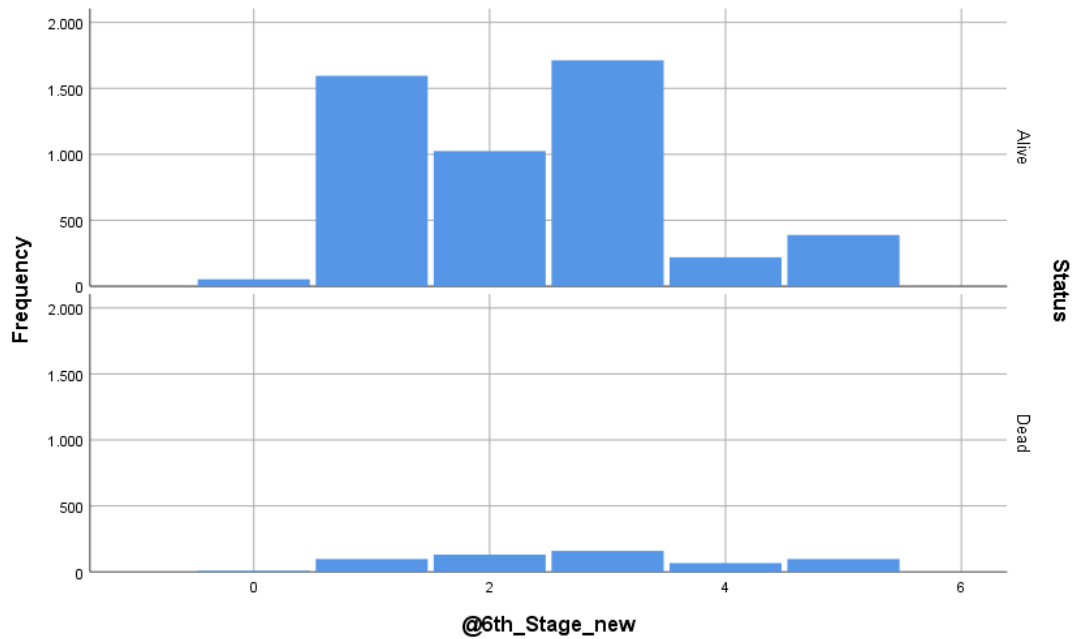
	Value	df	Asymptotic Significance (2- sided)	Exact Sig. (2- sided)
Pearson Chi-Square	149,095 ^a	5	,000	. ^b
Likelihood Ratio	131,472	5	,000	. ^b
Fisher's Exact Test	. ^b			. ^b
N of Valid Cases	5552			

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 6,12.

b. Cannot be computed because there is insufficient memory.

Πίνακας 3.35 : Έλεγχος χ^2 για τις μεταβλητές Status και 6th Stage.

Ο έλεγχος ανεξαρτησίας $\chi^2 = 149,095$ του Pearson έχει $p\text{-value} = 0,000 < 0,05$, συνεπώς απορρίπτουμε τη μηδενική υπόθεση του ελέγχου για το αν το 6^ο στάδιο καρκίνου συσχετίζεται με την θνησιμότητα από τον καρκίνο του μαστού. Άρα το 6^ο στάδιο καρκίνου σχετίζεται με τη θνησιμότητα από καρκίνο του μαστού.



Διάγραμμα 3.11 : Ραβδόγραμμα μεταξύ των μεταβλητών “6th Stage ” και “Status”.

Όπου :

IIB	3
IIIC	5
IIIB	4
I	0
IIIA	2
IIA	1

Συσχέτιση θνησιμότητας από τον καρκίνο του μαστού με το T στάδιο καρκίνου :

Case Processing Summary

	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Status * Stage_T	5552	100,0%	0	0,0%	5552	100,0%

Πίνακας 3.36 : Περίληψη επεξεργασίας υπόθεσης «Συσχέτιση θνησιμότητας από τον καρκίνο του μαστού με το T στάδιο».

Status * Stage_T Crosstabulation

		Stage_T										
		T1a	T1b	T1c	T2	T3	T4a	T4b	T4c	T4d	Total	
Status	Alive	Count	46	294	1383	2319	687	42	140	5	79	4995
		Expected Count	43,2	279,8	1316,2	2331,1	710,7	45,0	168,2	8,1	92,7	4995,0
Dead	Count	2	17	80	272	103	8	47	4	24	557	
		Expected Count	4,8	31,2	146,8	259,9	79,3	5,0	18,8	,9	10,3	557,0
Total	Count	48	311	1463	2591	790	50	187	9	103	5552	
		Expected Count	48,0	311,0	1463,0	2591,0	790,0	50,0	187,0	9,0	103,0	5552,0

Πίνακας 3.37 : «Διασταυρώσεις» των μεταβλητών Status και T Stage.

Chi-Square Tests

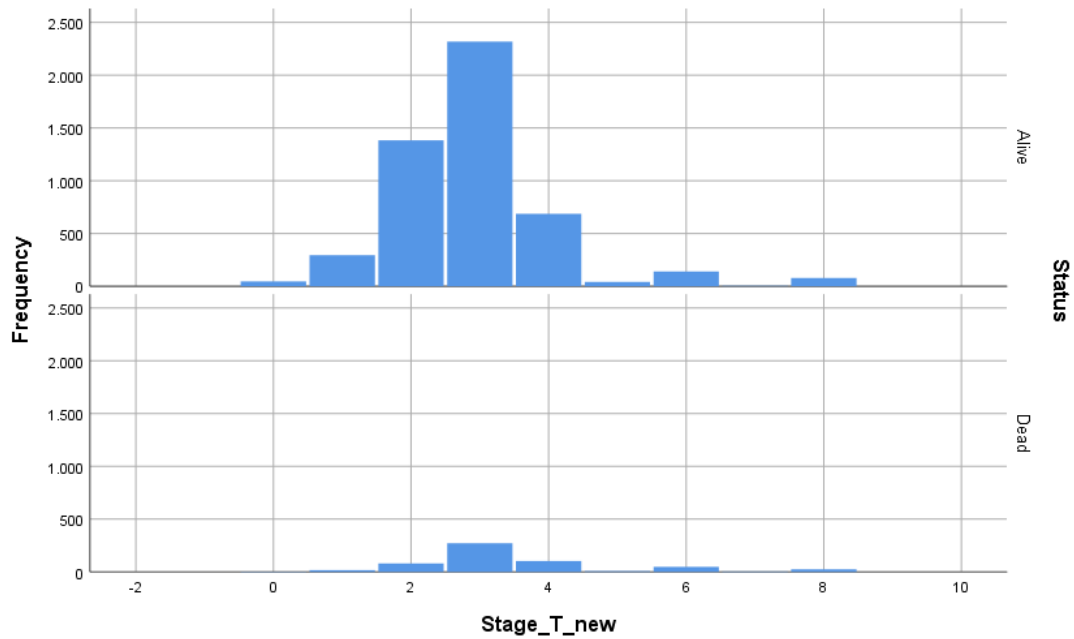
	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)
Pearson Chi-Square	132,427 ^a	8	,000	. ^b
Likelihood Ratio	117,358	8	,000	. ^b
Fisher's Exact Test	. ^b			. ^b
N of Valid Cases	5552			

a. 2 cells (11,1%) have expected count less than 5. The minimum expected count is ,90.

b. Cannot be computed because there is insufficient memory.

Πίνακας 3.38 : Έλεγχος χ^2 για τις μεταβλητές Status και T Stage.

Ο έλεγχος ανεξαρτησίας $\chi^2 = 132,427$ του Pearson έχει $p\text{-value} = 0,000 < 0,05$, συνεπώς απορρίπτουμε τη μηδενική υπόθεση του ελέγχου για το αν το T στάδιο καρκίνου συσχετίζεται με την θνησιμότητα από τον καρκίνο του μαστού. Άρα το T στάδιο καρκίνου σχετίζεται με τη θνησιμότητα από καρκίνο του μαστού.



Διάγραμμα 3.12 : Ραβδόγραμμα μεταξύ των μεταβλητών “T Stage” και “Status”.

Όπου :

T4d	8
T1b	1
T4b	6
T4c	7
T3	4
T1c	2
T4a	5
T1a	0
T2	3

Συσχέτιση θνησιμότητας από τον καρκίνο του μαστού με το N στάδιο καρκίνου :

Case Processing Summary

	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Status * Stage_N	5552	100,0%	0	0,0%	5552	100,0%

Πίνακας 3.39 : Περίληψη επεξεργασίας υπόθεσης «Συσχέτιση θνησιμότητας από τον καρκίνο του μαστού με το N στάδιο».

Status * Stage_N Crosstabulation

			Stage_N				Total
			N0	N1	N2	N3	
Status	Alive	Count	251	3675	680	389	4995
		Expected Count	271,7	3587,0	699,0	437,2	4995,0
	Dead	Count	51	312	97	97	557
		Expected Count	30,3	400,0	78,0	48,8	557,0
Total		Count	302	3987	777	486	5552
		Expected Count	302,0	3987,0	777,0	486,0	5552,0

Πίνακας 3.40 : «Διασταυρώσεις» των μεταβλητών Status και N Stage.

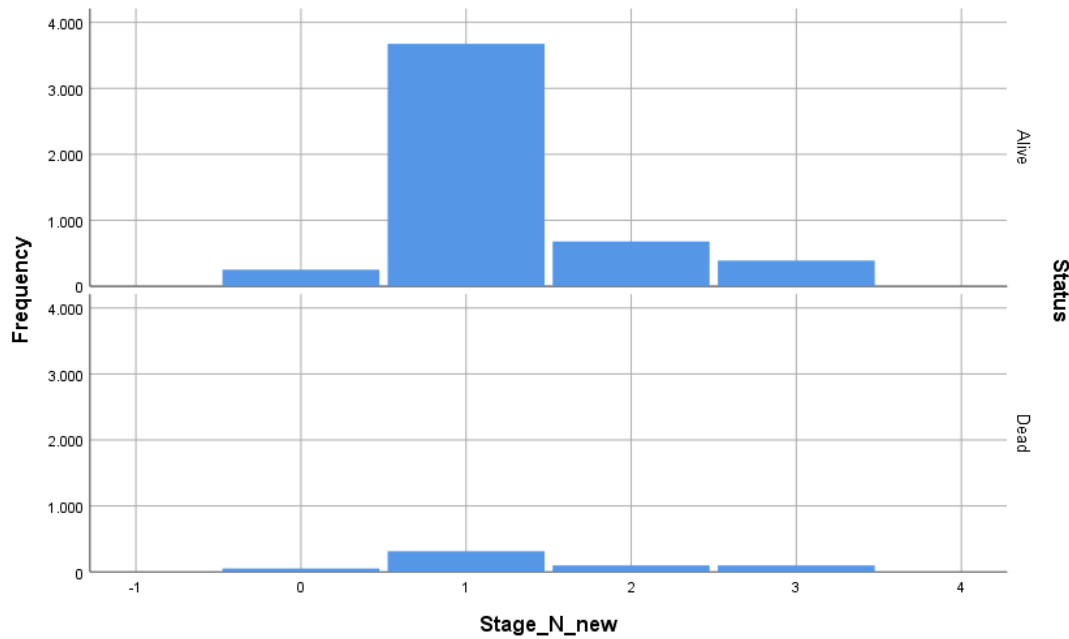
Chi-Square Tests

	Value	df	Asymptotic Significance (2- sided)	Exact Sig. (2- sided)
Pearson Chi-Square	95,468 ^a	3	,000	,000
Likelihood Ratio	83,757	3	,000	,000
Fisher's Exact Test	84,743			,000
N of Valid Cases	5552			

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 30,30.

Πίνακας 3.41 : Έλεγχος χ^2 για τις μεταβλητές Status και N Stage.

Ο έλεγχος ανεξαρτησίας $\chi^2 = 95,468$ του Pearson έχει $p\text{-value} = 0,000 < 0,05$, συνεπώς απορρίπτουμε τη μηδενική υπόθεση του ελέγχου για το αν το N στάδιο καρκίνου συσχετίζεται με την θνησιμότητα από τον καρκίνο του μαστού. Άρα το N στάδιο καρκίνου σχετίζεται με τη θνησιμότητα από καρκίνο του μαστού.



Διάγραμμα 3.13 : Ραβδόγραμμα μεταξύ των μεταβλητών “N Stage” και “Status”.

Όπου :

N1	1
N0	0
N2	2
N3	3

3.4 Έλεγχος συσχέτισης ποιοτικών και ποσοτικών μεταβλητών

Σύμφωνα με τον έλεγχο κανονικότητας των ποσοτικών μεταβλητών που προηγήθηκε, το δείγμα δεν ακολουθεί την κανονική κατανομή. Επομένως, για την συσχέτιση ποιοτικών και ποσοτικών μεταβλητών θα χρησιμοποιηθεί μη παραμετρικός έλεγχος υποθέσεων.

- **Μη παραμετρικός έλεγχος U των Mann-Whitney για ανεξάρτητα δείγματα.**

H_0 : Τα δύο ανεξάρτητα δείγματα έχουν την ίδια κατανομή.
έναντι της εναλλακτικής

H_1 : Τα δύο ανεξάρτητα δείγματα δεν έχουν την ίδια κατανομή.

- **Έλεγχος συσχέτισης της θνησιμότητας από καρκίνο του μαστού με ποσοτικές μεταβλητές.**

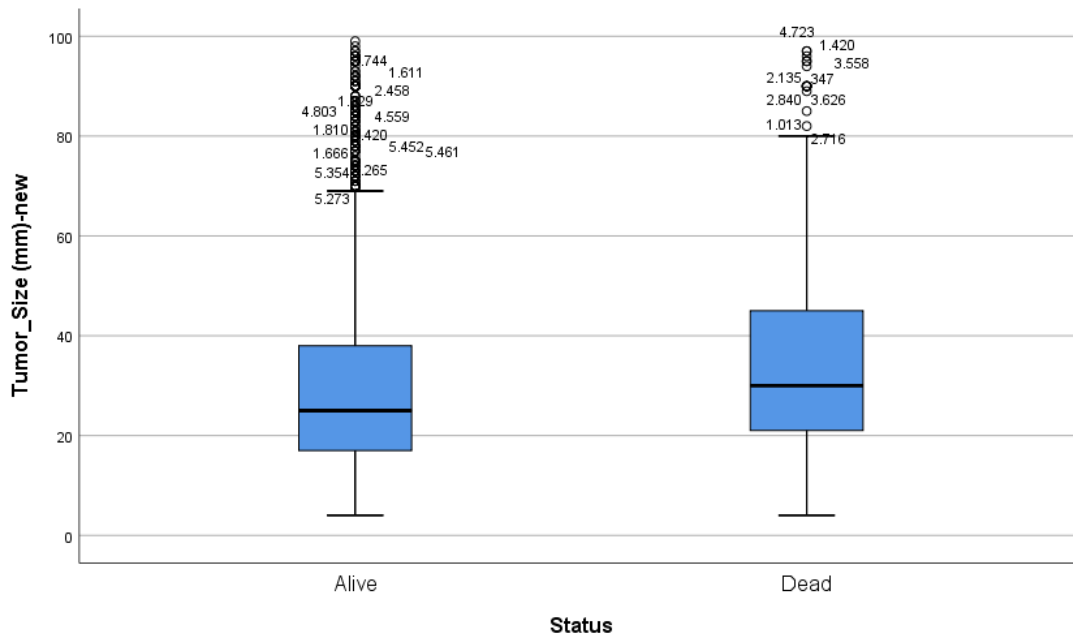
Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of YearOfDiagnosis is the same across categories of Status.	Independent-Samples Mann-Whitney U Test	,000	Reject the null hypothesis.
2	The distribution of Tumor_Size (mm)-new is the same across categories of Status.	Independent-Samples Mann-Whitney U Test	,000	Reject the null hypothesis.
3	The distribution of RegionalNodesExamined is the same across categories of Status.	Independent-Samples Mann-Whitney U Test	,026	Reject the null hypothesis.
4	The distribution of RegionalNodesPositive is the same across categories of Status.	Independent-Samples Mann-Whitney U Test	,000	Reject the null hypothesis.
5	The distribution of Survival_Months is the same across categories of Status.	Independent-Samples Mann-Whitney U Test	,000	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

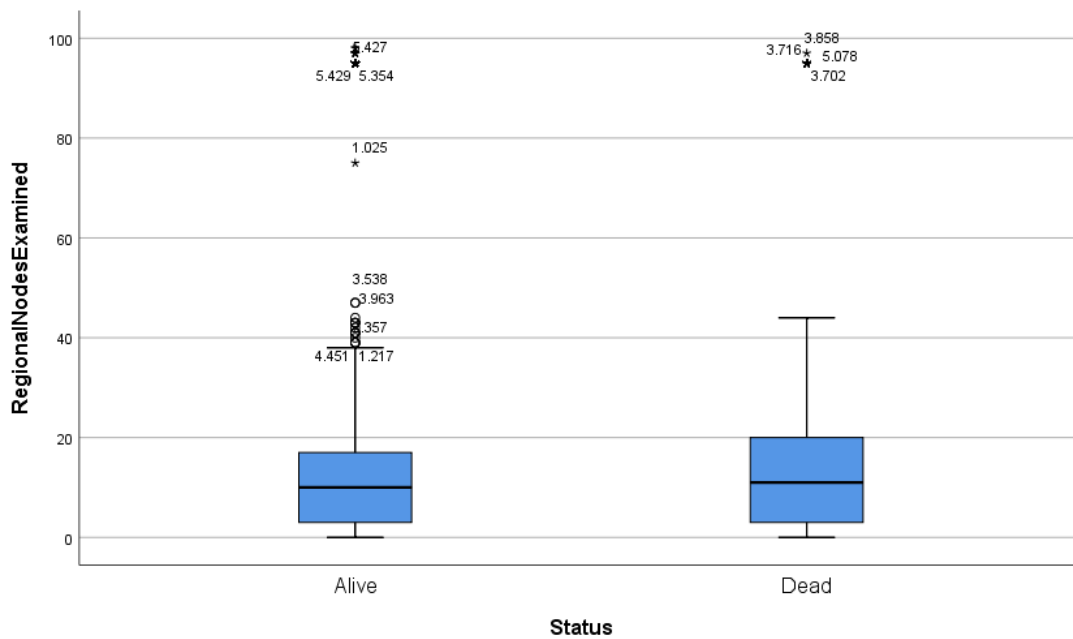
Πίνακας 3.42 : : Έλεγχος συσχετίσεων με τη μέθοδο του Mann-Whitney.

- ✚ Για την μεταβλητή Tumor_Size το $p\text{-value}=0,000 < 0,05$. Συνεπώς, σε επίπεδο σημαντικότητας 5% απορρίπτουμε την μηδενική υπόθεση. Αυτό σημαίνει ότι τα δύο δείγματα δεν έχουν την ίδια κατανομή.



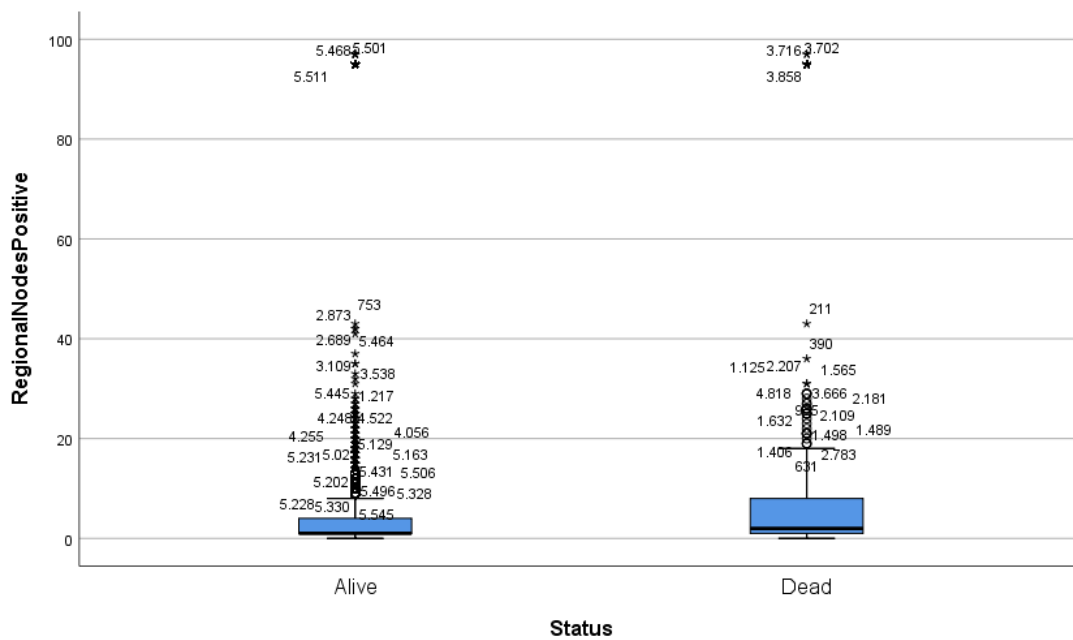
Διάγραμμα 3.14 : Θηκόγραμμα μεταξύ των μεταβλητών “Tumor_Size” και “Status”.

✚ Για την μεταβλητή RegionalNodesExamined το $p\text{-value}=0,026 < 0,05$. Συνεπώς, σε επίπεδο σημαντικότητας 5% απορρίπτουμε την μηδενική υπόθεση. Αυτό σημαίνει ότι τα δύο δείγματα δεν έχουν την ίδια κατανομή.



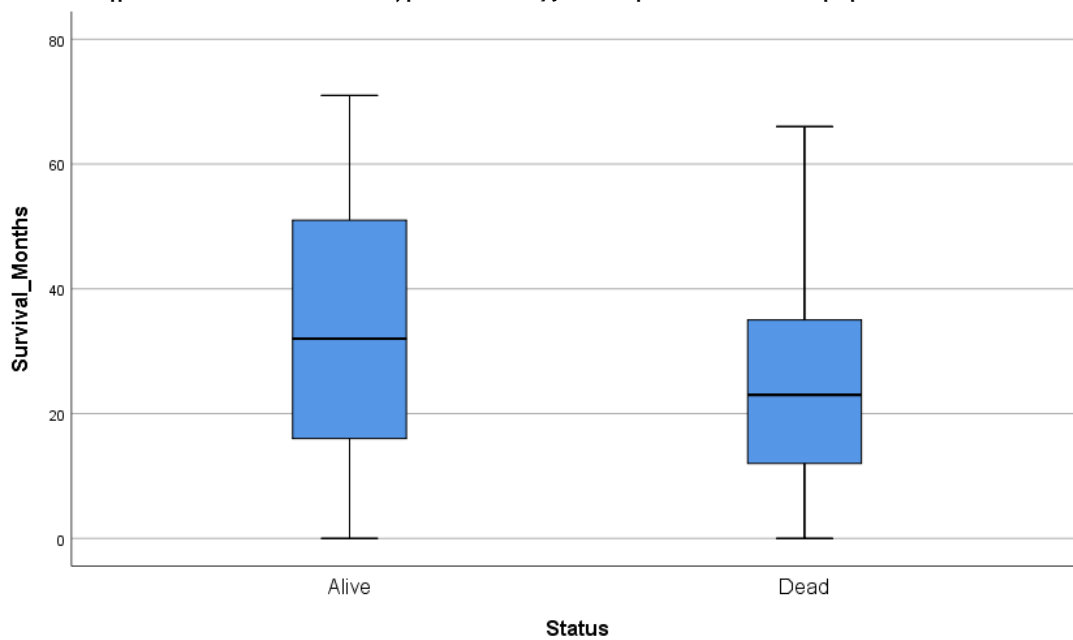
Διάγραμμα 3.15 : Θηκόγραμμα μεταξύ των μεταβλητών “Regional_Nodes_Examined” και “Status”.

🚩 Για την μεταβλητή RegionalNodesPositive το $p\text{-value}=0,000 < 0,05$. Συνεπώς, σε επίπεδο σημαντικότητας 5% απορρίπτουμε την μηδενική υπόθεση. Αυτό σημαίνει ότι τα δύο δείγματα δεν έχουν την ίδια κατανομή.



Διάγραμμα 3.16 : Θηκόγραμμα μεταξύ των μεταβλητών “Regional_Nodes_Positive” και “Status”.

🚩 Για την μεταβλητή Survival_Months το $p\text{-value}=0,000 < 0,05$. Συνεπώς, σε επίπεδο σημαντικότητας 5% απορρίπτουμε την μηδενική υπόθεση. Αυτό σημαίνει ότι τα δύο δείγματα δεν έχουν την ίδια κατανομή.



Διάγραμμα 3.17 : Θηκόγραμμα μεταξύ των μεταβλητών “Survival_Months” και “Status”.

- Μη παραμετρικός έλεγχος Kruskal-Wallis, U των Mann-Whitney για ανεξάρτητα δείγματα.

H_0 : Τα k δείγματα προέρχονται από τον ίδιο πληθυσμό.
έναντι της εναλλακτικής

H_1 : Τουλάχιστον 1 δείγμα προέρχεται από διαφορετικό πληθυσμό.

- Έλεγχος συσχέτισης των μηνών επιβίωσης με την ηλικία.

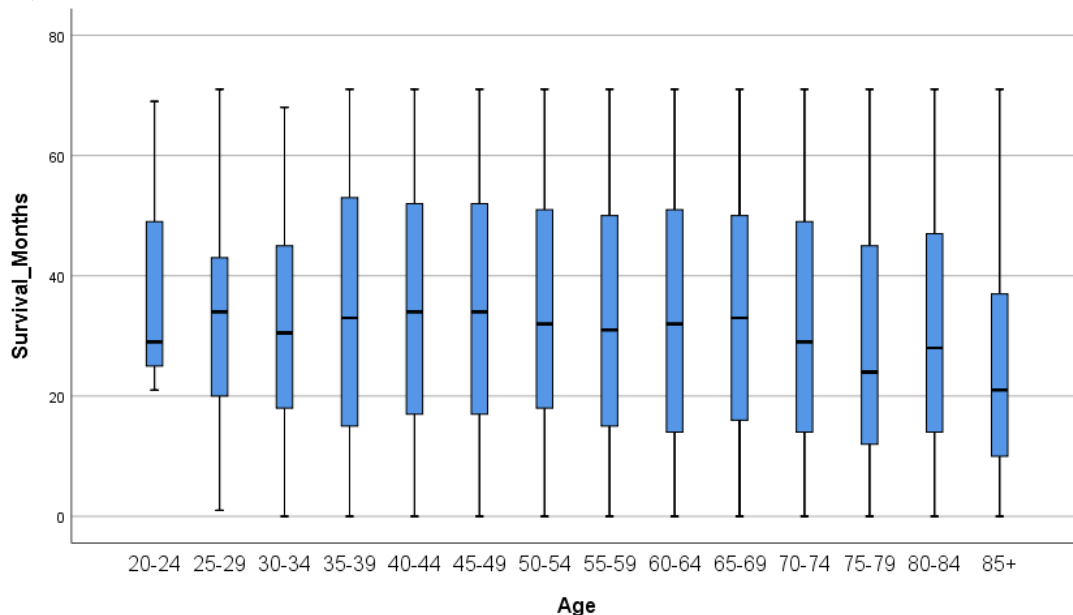
Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Survival_Months is the same across categories of Age.	Independent-Samples Kruskal-Wallis Test	,000	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

Πίνακας 3.43 : Έλεγχος συσχετίσεων με τη μέθοδο του Kruskal - Wallis – μήνες επιβίωσης με την ηλικία.

Για την μεταβλητή Survival_Months το $p\text{-value}=0,000 < 0,05$. Συνεπώς, σε επίπεδο σημαντικότητας 5% απορρίπτουμε την μηδενική υπόθεση. Αυτό σημαίνει ότι τα δύο δείγματα δεν προέρχονται από τον ίδιο πληθυσμό.



Διάγραμμα 3.18 : Θηκόγραμμα μεταξύ των μεταβλητών "Survival_Months" και "Age".

Παρακάτω παρατηρούμε ποια levels της μεταβλητής Age συσχετίζονται μεταξύ τους:

Multiple Comparisons

Dependent Variable: Rank of Survival_Months

Tukey HSD

(I) Age_new	(J) Age_new	Mean Difference			95% Confidence Interval	
		(I-J)	Std. Error	Sig.	Lower Bound	Upper Bound
0	1	458,722222	951,587776	1,000	-2733,97504	3651,41949
	2	603,339744	931,941469	1,000	-2523,44168	3730,12117
	3	428,048409	927,087839	1,000	-2682,44847	3538,54529
	4	373,151613	924,338288	1,000	-2728,12018	3474,42340
	5	428,040900	923,488188	1,000	-2670,37870	3526,46050
	6	439,556777	923,267379	1,000	-2658,12198	3537,23553
	7	527,127628	923,443730	1,000	-2571,14281	3625,39806
	8	545,394293	923,161712	1,000	-2551,92994	3642,71852
	9	468,172035	923,488188	1,000	-2630,24756	3566,59163
	10	642,923905	924,443854	1,000	-2458,70207	3744,54988
	11	833,022917	925,679753	1,000	-2272,74966	3938,79549
	12	725,784970	927,520276	1,000	-2386,16279	3837,73273
	13	1154,715210	928,055664	,994	-1959,02885	4268,45927
1	0	-458,722222	951,587776	1,000	-3651,41949	2733,97504
	2	144,617521	276,017242	1,000	-781,45521	1070,69025
	3	-30,673813	259,157247	1,000	-900,17913	838,83151
	4	-85,570609	249,142303	1,000	-921,47453	750,33331
	5	-30,681322	245,969611	1,000	-855,94046	794,57781
	6	-19,165446	245,139288	1,000	-841,63874	803,30785
	7	68,405405	245,802643	1,000	-756,29353	893,10434
	8	86,672071	244,741013	1,000	-734,46496	907,80910
	9	9,449813	245,969611	1,000	-815,80932	834,70895
	10	184,201683	249,533678	1,000	-653,01535	1021,41871
	11	374,300694	254,074048	,972	-478,14985	1226,75124
	12	267,062748	260,699978	,999	-607,61863	1141,74412
	13	695,992988	262,598423	,309	-185,05791	1577,04389
2	0	-603,339744	931,941469	1,000	-3730,12117	2523,44168
	1	-144,617521	276,017242	1,000	-1070,69025	781,45521
	3	-175,291334	173,660832	,999	-757,94538	407,36271
	4	-230,188131	158,327170	,975	-761,39580	301,01954
	5	-175,298844	153,286187	,997	-689,59338	338,99569
	6	-163,782967	151,950243	,999	-673,59524	346,02930
	7	-76,212116	153,018120	1,000	-589,60725	437,18302
	8	-57,945450	151,306871	1,000	-565,59913	449,70823
	9	-135,167709	153,286187	1,000	-649,46224	379,12683
	10	39,584161	158,942323	1,000	-493,68742	572,85574

	11	229,683173	165,979599	,984	-327,19936	786,56571
	12	122,445227	175,954778	1,000	-467,90530	712,79575
	13	551,375467	178,755524	,110	-48,37191	1151,12285
3	0	-428,048409	927,087839	1,000	-3538,54529	2682,44847
	1	30,673813	259,157247	1,000	-838,83151	900,17913
	2	175,291334	173,660832	,999	-407,36271	757,94538
	4	-54,896797	126,666705	1,000	-479,87960	370,08601
	5	-,007510	120,306344	1,000	-403,65050	403,63548
	6	11,508367	118,599483	1,000	-386,40789	409,42462
	7	99,079218	119,964605	1,000	-303,41719	501,57563
	8	117,345884	117,774065	,999	-277,80099	512,49276
	9	40,123625	120,306344	1,000	-363,51937	443,76662
	10	214,875496	127,434780	,921	-212,68430	642,43529
	11	404,974507	136,110942	,147	-51,69491	861,64392
	12	297,736561	148,111597	,762	-199,19657	794,66969
	13	726,666801*	151,428197	,000	218,60606	1234,72754
4	0	-373,151613	924,338288	1,000	-3474,42340	2728,12018
	1	85,570609	249,142303	1,000	-750,33331	921,47453
	2	230,188131	158,327170	,975	-301,01954	761,39580
	3	54,896797	126,666705	1,000	-370,08601	479,87960
	5	54,889287	96,866015	1,000	-270,10844	379,88701
	6	66,405164	94,737774	1,000	-251,45204	384,26237
	7	153,976015	96,441250	,947	-169,59657	477,54860
	8	172,242680	93,702395	,859	-142,14070	486,62606
	9	95,020422	96,866015	,999	-229,97730	420,01814
	10	269,772292	105,588974	,370	-84,49206	624,03665
	11	459,871304*	115,912022	,006	70,97182	848,77079
	12	352,633357	129,793889	,269	-82,84155	788,10826
	13	781,563597*	133,566115	,000	333,43240	1229,69480
5	0	-428,040900	923,488188	1,000	-3526,46050	2670,37870
	1	30,681322	245,969611	1,000	-794,57781	855,94046
	2	175,298844	153,286187	,997	-338,99569	689,59338
	3	,007510	120,306344	1,000	-403,63548	403,65050
	4	-54,889287	96,866015	1,000	-379,88701	270,10844
	6	11,515877	86,048871	1,000	-277,18896	300,22071
	7	99,086728	87,920857	,998	-195,89886	394,07232
	8	117,353393	84,907603	,984	-167,52235	402,22914
	9	40,131135	88,386578	1,000	-256,41701	336,67928
	10	214,883005	97,868247	,633	-113,47734	543,24335
	11	404,982017*	108,925476	,015	39,52328	770,44076

	12	297,744070	123,594563	,474	-116,93131	712,41945
	13	726,674311*	127,550262	,000	298,72706	1154,62156
6	0	-439,556777	923,267379	1,000	-3537,23553	2658,12198
	1	19,165446	245,139288	1,000	-803,30785	841,63874
	2	163,782967	151,950243	,999	-346,02930	673,59524
	3	-11,508367	118,599483	1,000	-409,42462	386,40789
	4	-66,405164	94,737774	1,000	-384,26237	251,45204
	5	-11,515877	86,048871	1,000	-300,22071	277,18896
	7	87,570851	85,570428	,999	-199,52875	374,67045
	8	105,837517	82,471341	,992	-170,86426	382,53929
	9	28,615258	86,048871	1,000	-260,08958	317,32010
	10	203,367128	95,762283	,685	-117,92744	524,66169
	11	393,466140*	107,037284	,017	34,34252	752,58976
	12	286,228194	121,933740	,519	-122,87492	695,33130
	13	715,158434*	125,941615	,000	292,60840	1137,70847
7	0	-527,127628	923,443730	1,000	-3625,39806	2571,14281
	1	-68,405405	245,802643	1,000	-893,10434	756,29353
	2	76,212116	153,018120	1,000	-437,18302	589,60725
	3	-99,079218	119,964605	1,000	-501,57563	303,41719
	4	-153,976015	96,441250	,947	-477,54860	169,59657
	5	-99,086728	87,920857	,998	-394,07232	195,89886
	6	-87,570851	85,570428	,999	-374,67045	199,52875
	8	18,266665	84,422693	1,000	-264,98214	301,51547
	9	-58,955593	87,920857	1,000	-353,94118	236,03000
	10	115,796277	97,447852	,996	-211,15358	442,74614
	11	305,895289	108,547912	,216	-58,29668	670,08725
	12	198,657343	123,261940	,943	-214,90205	612,21673
	13	627,587583*	127,227982	,000	200,72162	1054,45354
8	0	-545,394293	923,161712	1,000	-3642,71852	2551,92994
	1	-86,672071	244,741013	1,000	-907,80910	734,46496
	2	57,945450	151,306871	1,000	-449,70823	565,59913
	3	-117,345884	117,774065	,999	-512,49276	277,80099
	4	-172,242680	93,702395	,859	-486,62606	142,14070
	5	-117,353393	84,907603	,984	-402,22914	167,52235
	6	-105,837517	82,471341	,992	-382,53929	170,86426
	7	-18,266665	84,422693	1,000	-301,51547	264,98214
	9	-77,222258	84,907603	1,000	-362,09800	207,65348
	10	97,529612	94,738101	,999	-220,32869	415,38792
	11	287,628624	106,121973	,273	-68,42401	643,68126
	12	180,390677	121,131045	,970	-226,01929	586,80064
	13	609,320917*	125,164625	,000	189,37778	1029,26405

9	0	-468,172035	923,488188	1,000	-3566,59163	2630,24756
	1	-9,449813	245,969611	1,000	-834,70895	815,80932
	2	135,167709	153,286187	1,000	-379,12683	649,46224
	3	-40,123625	120,306344	1,000	-443,76662	363,51937
	4	-95,020422	96,866015	,999	-420,01814	229,97730
	5	-40,131135	88,386578	1,000	-336,67928	256,41701
	6	-28,615258	86,048871	1,000	-317,32010	260,08958
	7	58,955593	87,920857	1,000	-236,03000	353,94118
	8	77,222258	84,907603	1,000	-207,65348	362,09800
	10	174,751870	97,868247	,883	-153,60847	503,11221
	11	364,850882	108,925476	,051	-,60786	730,30962
	12	257,612935	123,594563	,712	-157,06245	672,28832
	13	686,543176*	127,550262	,000	258,59592	1114,49043
	10	0	-642,923905	924,443854	1,000	-3744,54988
1		-184,201683	249,533678	1,000	-1021,41871	653,01535
2		-39,584161	158,942323	1,000	-572,85574	493,68742
3		-214,875496	127,434780	,921	-642,43529	212,68430
4		-269,772292	105,588974	,370	-624,03665	84,49206
5		-214,883005	97,868247	,633	-543,24335	113,47734
6		-203,367128	95,762283	,685	-524,66169	117,92744
7		-115,796277	97,447852	,996	-442,74614	211,15358
8		-97,529612	94,738101	,999	-415,38792	220,32869
9		-174,751870	97,868247	,883	-503,11221	153,60847
11		190,099012	116,750871	,939	-201,61492	581,81294
12		82,861065	130,543566	1,000	-355,12910	520,85123
13		511,791305*	134,294736	,010	61,21549	962,36712
11		0	-833,022917	925,679753	1,000	-3938,79549
	1	-374,300694	254,074048	,972	-1226,75124	478,14985
	2	-229,683173	165,979599	,984	-786,56571	327,19936
	3	-404,974507	136,110942	,147	-861,64392	51,69491
	4	-459,871304*	115,912022	,006	-848,77079	-70,97182
	5	-404,982017*	108,925476	,015	-770,44076	-39,52328
	6	-393,466140*	107,037284	,017	-752,58976	-34,34252
	7	-305,895289	108,547912	,216	-670,08725	58,29668
	8	-287,628624	106,121973	,273	-643,68126	68,42401
	9	-364,850882	108,925476	,051	-730,30962	-,60786
	10	-190,099012	116,750871	,939	-581,81294	201,61492
	12	-107,237946	139,025854	1,000	-573,68726	359,21137
	13	321,692294	142,553995	,588	-156,59438	799,97897
	12	0	-725,784970	927,520276	1,000	-3837,73273

1	-267,062748	260,699978	,999	-1141,74412	607,61863	
2	-122,445227	175,954778	1,000	-712,79575	467,90530	
3	-297,736561	148,111597	,762	-794,66969	199,19657	
4	-352,633357	129,793889	,269	-788,10826	82,84155	
5	-297,744070	123,594563	,474	-712,41945	116,93131	
6	-286,228194	121,933740	,519	-695,33130	122,87492	
7	-198,657343	123,261940	,943	-612,21673	214,90205	
8	-180,390677	121,131045	,970	-586,80064	226,01929	
9	-257,612935	123,594563	,712	-672,28832	157,06245	
10	-82,861065	130,543566	1,000	-520,85123	355,12910	
11	107,237946	139,025854	1,000	-359,21137	573,68726	
13	428,930240	154,053556	,233	-87,93891	945,79939	
13	0	-1154,715210	928,055664	,994	-4268,45927	1959,02885
1	-695,992988	262,598423	,309	-1577,04389	185,05791	
2	-551,375467	178,755524	,110	-1151,12285	48,37191	
3	-726,666801*	151,428197	,000	-1234,72754	-218,60606	
4	-781,563597*	133,566115	,000	-1229,69480	-333,43240	
5	-726,674311*	127,550262	,000	-1154,62156	-298,72706	
6	-715,158434*	125,941615	,000	-1137,70847	-292,60840	
7	-627,587583*	127,227982	,000	-1054,45354	-200,72162	
8	-609,320917*	125,164625	,000	-1029,26405	-189,37778	
9	-686,543176*	127,550262	,000	-1114,49043	-258,59592	
10	-511,791305*	134,294736	,010	-962,36712	-61,21549	
11	-321,692294	142,553995	,588	-799,97897	156,59438	
12	-428,930240	154,053556	,233	-945,79939	87,93891	

*. The mean difference is significant at the 0.05 level.

Πίνακας 3.44 : levels της μεταβλητής Age που συσχετίζονται μεταξύ τους.

ο Έλεγχος συσχέτισης των μηνών επιβίωσης με την φυλή.

Hypothesis Test Summary				
	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Survival_Months is the same across categories of Race.	Independent-Samples Kruskal-Wallis Test	,282	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

Πίνακας 3.45 : Έλεγχος συσχέτισεων με τη μέθοδο του Kruskal - Wallis - μήνες επιβίωσης με την φυλή.

Για την μεταβλητή *Survival_Months* το $p\text{-value}=0,282 > 0,05$. Συνεπώς, σε επίπεδο σημαντικότητας 5% δεχόμαστε την μηδενική υπόθεση. Αυτό σημαίνει ότι τα δύο δείγματα προέρχονται από τον ίδιο πληθυσμό.

- ο Έλεγχος συσχέτισης των μηνών επιβίωσης με την οικογενειακή κατάσταση.

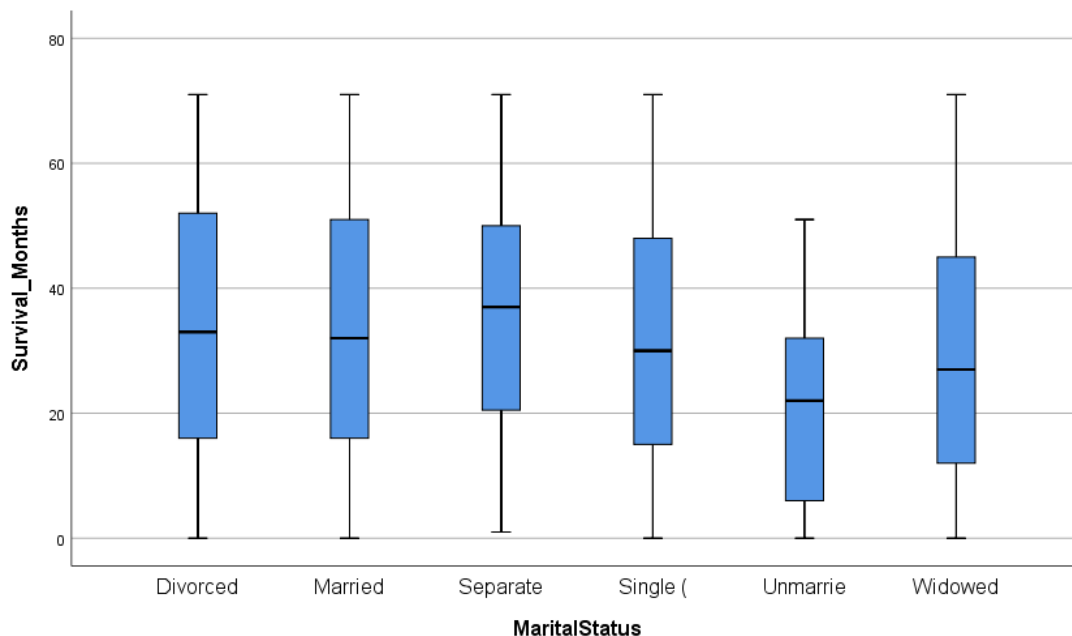
Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of <i>Survival_Months</i> is the same across categories of <i>MaritalStatus</i> .	Independent-Samples Kruskal-Wallis Test	,000	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

Πίνακας 3.46 : Έλεγχος συσχετίσεων με τη μέθοδο του Kruskal - Wallis – μήνες επιβίωσης με την οικογενειακή κατάσταση.

Για την μεταβλητή *Survival_Months* το $p\text{-value}=0,000 < 0,05$. Συνεπώς, σε επίπεδο σημαντικότητας 5% απορρίπτουμε την μηδενική υπόθεση. Αυτό σημαίνει ότι τα δύο δείγματα δεν προέρχονται από τον ίδιο πληθυσμό.



Διάγραμμα 3.19 : Θηκόγραμμα μεταξύ των μεταβλητών “*Survival_Months*” και “*Marital_Status*”.

Παρακάτω παρατηρούμε ποια levels της μεταβλητής *Marital_Status* συσχετίζονται μεταξύ τους.

Multiple Comparisons

Dependent Variable: Rank of Survival_Months

Tukey HSD

(I) MaritalStatus_new	(J) MaritalStatus_new	Mean Difference	Std. Error	Sig.	95% Confidence Interval	
		(I-J)			Lower Bound	Upper Bound
0	1	-301,107578*	69,277287	,000	-498,59663	-103,61853
	2	-354,551054*	88,181588	,001	-605,93067	-103,17144
	3	-191,065056	87,202404	,242	-439,65331	57,52320
	4	-471,578576	189,597627	,128	-1012,06531	68,90816
	5	599,216036	219,199069	,069	-25,65565	1224,08773
1	0	301,107578*	69,277287	,000	103,61853	498,59663
	2	-53,443475	66,977141	,968	-244,37549	137,48854
	3	110,042523	65,682602	,548	-77,19914	297,28419
	4	-170,470997	180,713046	,935	-685,63042	344,68843
	5	900,323615*	211,561284	,000	297,22499	1503,42224
2	0	354,551054*	88,181588	,001	103,17144	605,93067
	1	53,443475	66,977141	,968	-137,48854	244,37549
	3	163,485998	85,386499	,393	-79,92565	406,89764
	4	-117,027522	188,769317	,990	-655,15299	421,09795
	5	953,767090*	218,483013	,000	330,93666	1576,59752
3	0	191,065056	87,202404	,242	-57,52320	439,65331
	1	-110,042523	65,682602	,548	-297,28419	77,19914
	2	-163,485998	85,386499	,393	-406,89764	79,92565
	4	-280,513520	188,313892	,671	-817,34071	256,31367
	5	790,281092*	218,089646	,004	168,57204	1411,99015
4	0	471,578576	189,597627	,128	-68,90816	1012,06531
	1	170,470997	180,713046	,935	-344,68843	685,63042
	2	117,027522	188,769317	,990	-421,09795	655,15299
	3	280,513520	188,313892	,671	-256,31367	817,34071
	5	1070,794612*	275,510607	,001	285,39539	1856,19384
5	0	-599,216036	219,199069	,069	-1224,08773	25,65565
	1	-900,323615*	211,561284	,000	-1503,42224	-297,22499
	2	-953,767090*	218,483013	,000	-1576,59752	-330,93666
	3	-790,281092*	218,089646	,004	-1411,99015	-168,57204
	4	-1070,794612*	275,510607	,001	-1856,19384	-285,39539

*. The mean difference is significant at the 0.05 level.

Πίνακας 3.47 : levels της μεταβλητής Marital_Status που συσχετίζονται μεταξύ τους.

- ο Έλεγχος συσχέτισης των μηνών επιβίωσης με μία μέτρηση του κατά πόσο τα καρκινικά κύτταρα μοιάζουν με τα κανονικά.

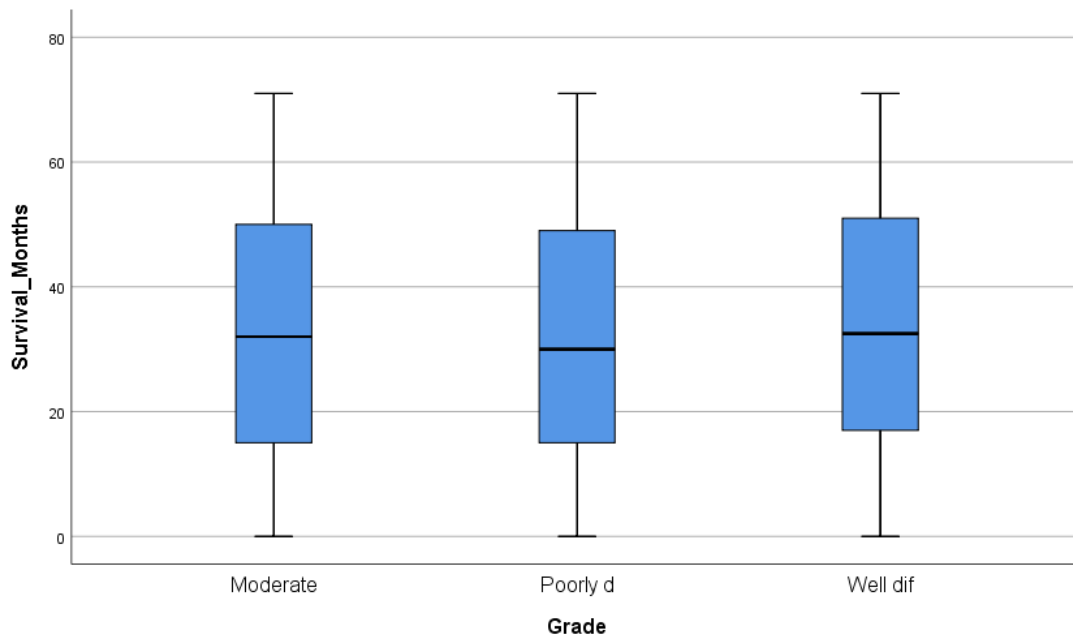
Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Survival_Months is the same across categories of Grade.	Independent-Samples Kruskal-Wallis Test	,042	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

Πίνακας 3.48 : Έλεγχος συσχετίσεων με τη μέθοδο του Kruskal - Wallis - μήνες επιβίωσης με μία μέτρηση του κατά πόσο τα καρκινικά κύτταρα μοιάζουν με τα κανονικά.

Για την μεταβλητή Survival_Months το $p\text{-value}=0,042 < 0,05$. Συνεπώς, σε επίπεδο σημαντικότητας 5% απορρίπτουμε την μηδενική υπόθεση. Αυτό σημαίνει ότι τα δύο δείγματα δεν προέρχονται από τον ίδιο πληθυσμό.



Διάγραμμα 3.20 : Θηκόγραμμα μεταξύ των μεταβλητών "Survival_Months" και "Grade".

Παρακάτω παρατηρούμε ποια levels της μεταβλητής Grade συσχετίζονται μεταξύ τους:

Multiple Comparisons

Dependent Variable: Rank of Survival_Months

Tukey HSD

(I) Grade_new	(J) Grade_new	Mean Difference			95% Confidence Interval	
		(I-J)	Std. Error	Sig.	Lower Bound	Upper Bound
0	1	63,020678	67,542086	,619	-95,32032	221,36168
	2	150,963347	68,108247	,068	-8,70492	310,63162
1	0	-63,020678	67,542086	,619	-221,36168	95,32032
	2	87,942669	46,152943	,137	-20,25510	196,14044
2	0	-150,963347	68,108247	,068	-310,63162	8,70492
	1	-87,942669	46,152943	,137	-196,14044	20,25510

Πίνακας 3.49 : levels της μεταβλητής Grade που συσχετίζονται μεταξύ τους.

- Έλεγχος συσχέτισης των μηνών επιβίωσης με την προέλευση πρωτογενούς.

Test Statistics^a

	Survival_Months
Mann-Whitney U	3824093,500
Wilcoxon W	7497498,500
Z	-,449
Asymp. Sig. (2-tailed)	,653

a. Grouping Variable: Laterality_new

Πίνακας 3.50 : Έλεγχος συσχετίσεων με τη μέθοδο του Mann Whitney - μήνες επιβίωσης με την προέλευση πρωτογενούς.

Για την μεταβλητή Survival_Months το $p\text{-value}=0,653 > 0,05$. Συνεπώς, σε επίπεδο σημαντικότητας 5% δεχόμαστε την μηδενική υπόθεση. Αυτό σημαίνει ότι τα δύο δείγματα προέρχονται από τον ίδιο πληθυσμό.

- Έλεγχος συσχέτισης των μηνών επιβίωσης με το στάδιο A.

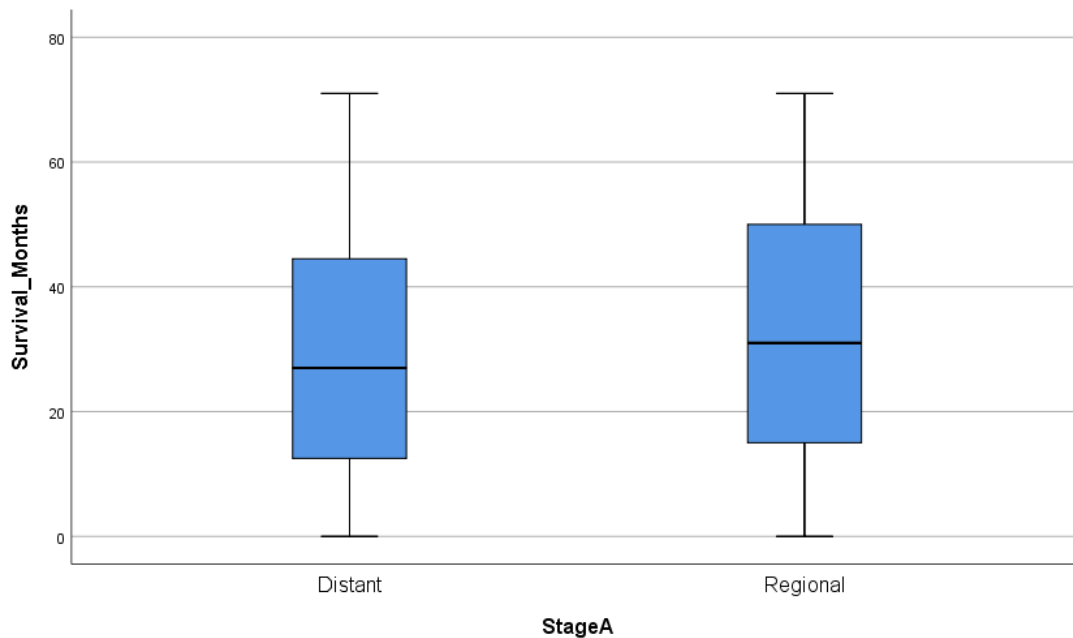
Test Statistics^a

	Survival_Months
Mann-Whitney U	631304,500
Wilcoxon W	667350,500
Z	-2,999
Asymp. Sig. (2-tailed)	,003

a. Grouping Variable: StageA_new

Πίνακας 3.51 : Έλεγχος συσχετίσεων με τη μέθοδο του Mann Whitney - μήνες επιβίωσης με το στάδιο A.

Για την μεταβλητή Survival_Months το $p\text{-value}=0,003 < 0,05$. Συνεπώς, σε επίπεδο σημαντικότητας 5% απορρίπτουμε την μηδενική υπόθεση. Αυτό σημαίνει ότι τα δύο δείγματα δεν προέρχονται από τον ίδιο πληθυσμό.



Διάγραμμα 3.21 : Θηκόγραμμα μεταξύ των μεταβλητών "Survival_Months" και "StageA".

- Έλεγχος συσχέτισης των μηνών επιβίωσης με τους υποδοχείς οιστρογόνου.

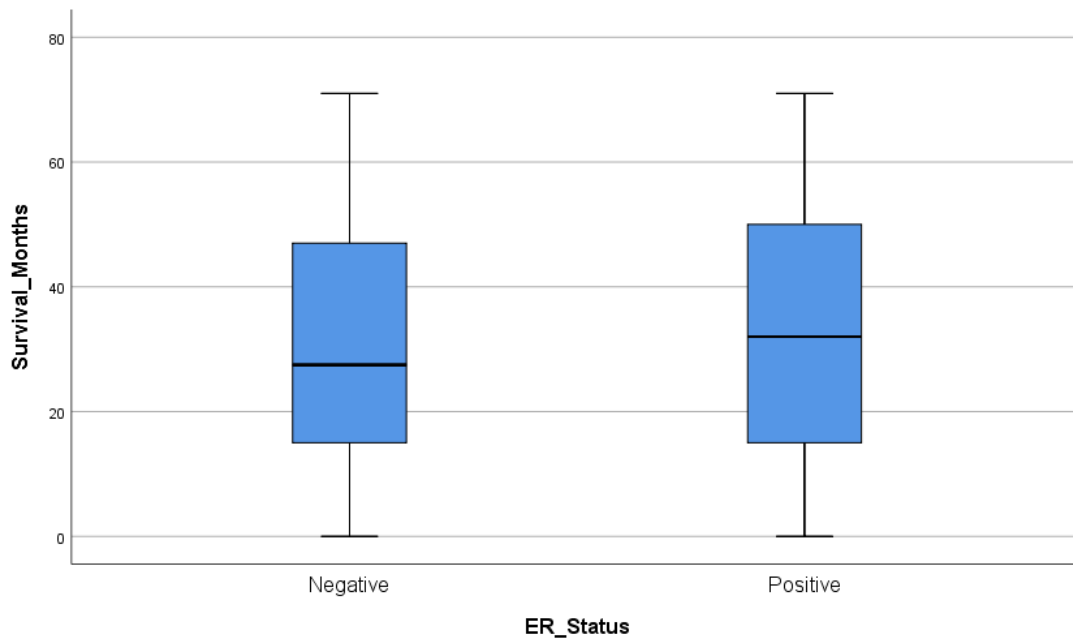
Test Statistics^a

	Survival_Months
Mann-Whitney U	1980375,500
Wilcoxon W	2389435,500
Z	-2,733
Asymp. Sig. (2-tailed)	,006

a. Grouping Variable: ER_Status_new

Πίνακας 3.52 : Έλεγχος συσχετίσεων με τη μέθοδο του Mann Whitney – μήνες επιβίωσης με τους υποδοχείς οιστρογόνου.

Για την μεταβλητή Survival_Months το $p\text{-value}=0,006 < 0,05$. Συνεπώς, σε επίπεδο σημαντικότητας 5% απορρίπτουμε την μηδενική υπόθεση. Αυτό σημαίνει ότι τα δύο δείγματα δεν προέρχονται από τον ίδιο πληθυσμό.



Διάγραμμα 3.22 : Θηκόγραμμα μεταξύ των μεταβλητών “Survival_Months” και “ER_Status”.

- Έλεγχος συσχέτισης των μηνών επιβίωσης με τους υποδοχείς προγεστερόνης.

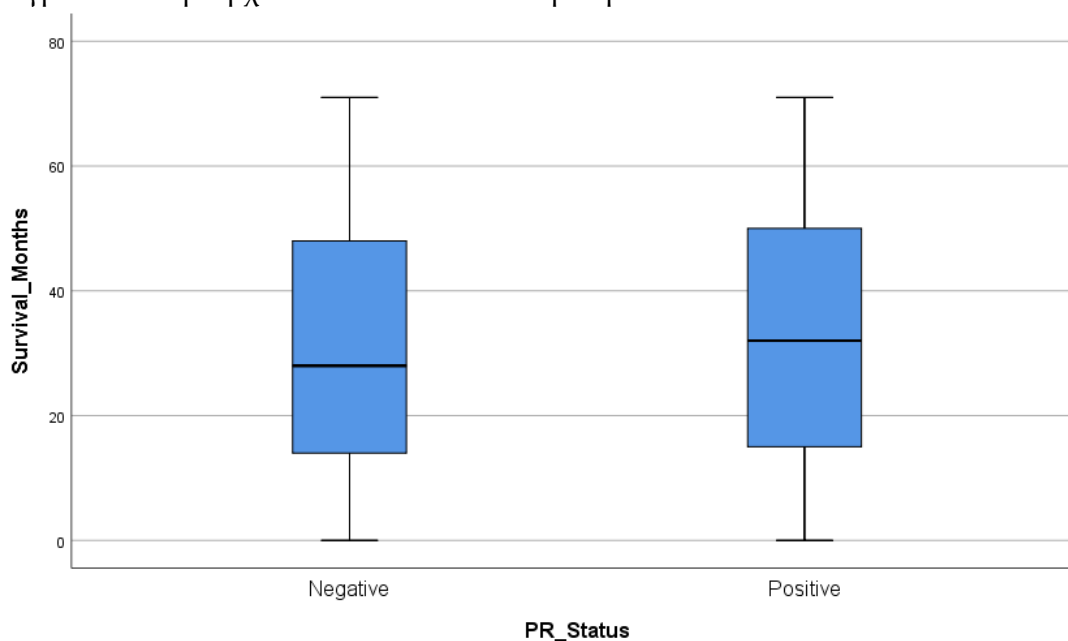
Test Statistics^a

	Survival_Months
Mann-Whitney U	2735844,000
Wilcoxon W	3701199,000
Z	-3,004
Asymp. Sig. (2-tailed)	,003

a. Grouping Variable: PR_Status_new

Πίνακας 3.53 : Έλεγχος συσχετίσεων με τη μέθοδο του Mann Whitney – μήνες επιβίωσης με τους υποδοχείς προγεστερόνης.

Για την μεταβλητή Survival_Months το $p\text{-value}=0,003 < 0,05$. Συνεπώς, σε επίπεδο σημαντικότητας 5% απορρίπτουμε την μηδενική υπόθεση. Αυτό σημαίνει ότι τα δύο δείγματα δεν προέρχονται από τον ίδιο πληθυσμό.



Διάγραμμα 3.23 : Θηκόγραμμα μεταξύ των μεταβλητών “Survival_Months” και “PR_Status”.

- Έλεγχος συσχέτισης των μηνών επιβίωσης με τον δευτερεύον τύπο μαστού.

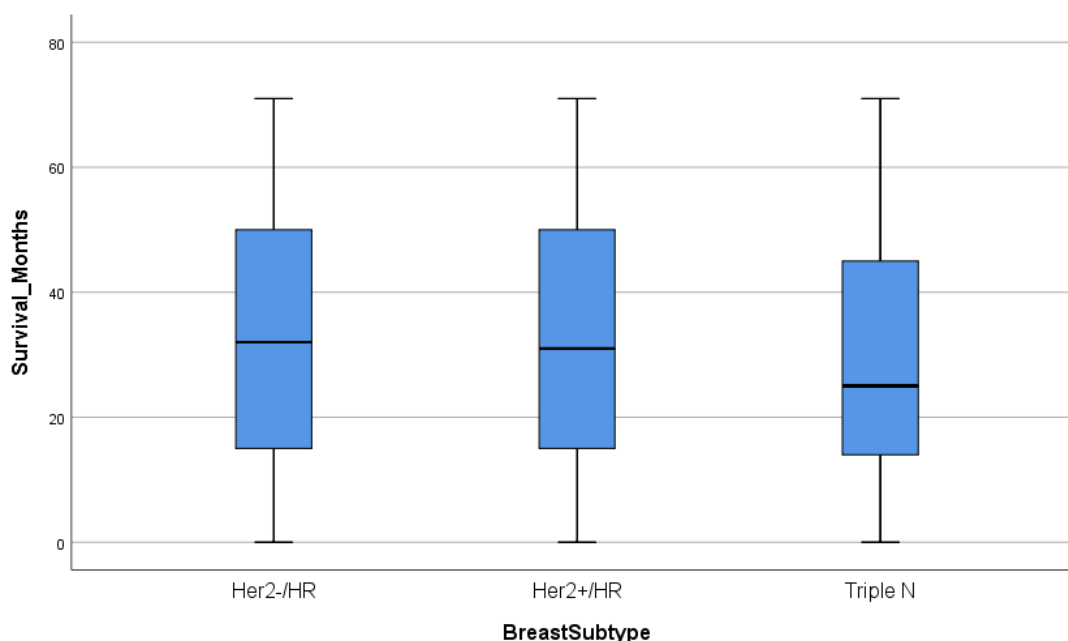
Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Survival_Months is the same across categories of BreastSubtype.	Independent-Samples Kruskal-Wallis Test	,001	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

Πίνακας 3.54 : Έλεγχος συσχετίσεων με τη μέθοδο του Kruskal - Wallis - μήνες επιβίωσης με τον δευτερεύον τύπο μαστού.

Για την μεταβλητή Survival_Months το $p\text{-value}=0,001 < 0,05$. Συνεπώς, σε επίπεδο σημαντικότητας 5% απορρίπτουμε την μηδενική υπόθεση. Αυτό σημαίνει ότι τα δύο δείγματα δεν προέρχονται από τον ίδιο πληθυσμό.



Διάγραμμα 3.24 : Θηκόγραμμα μεταξύ των μεταβλητών "Survival_Months" και "BreastSubtype".

Παρακάτω παρατηρούμε ποια levels της μεταβλητής BreastSubtype συσχετίζονται μεταξύ τους:

Multiple Comparisons

Dependent Variable: Rank of Survival_Months

Tukey HSD

(I) BreastSubtype_new	(J) BreastSubtype_new	Mean Difference	Std. Error	Sig.	95% Confidence Interval	
		(I-J)			Lower Bound	Upper Bound
0	1	55,473174	66,762520	,840	-116,09385	227,04019
	2	-55,422589	99,084887	,944	-310,05197	199,20679
	3	278,827631*	72,803036	,001	91,73763	465,91763
1	0	-55,473174	66,762520	,840	-227,04019	116,09385
	2	-110,895763	114,064540	,765	-404,02001	182,22848
	3	223,354457	92,158485	,073	-13,47538	460,18430
2	0	55,422589	99,084887	,944	-199,20679	310,05197
	1	110,895763	114,064540	,765	-182,22848	404,02001
	3	334,250220*	117,702028	,023	31,77832	636,72212
3	0	-278,827631*	72,803036	,001	-465,91763	-91,73763
	1	-223,354457	92,158485	,073	-460,18430	13,47538
	2	-334,250220*	117,702028	,023	-636,72212	-31,77832

*. The mean difference is significant at the 0.05 level.

Πίνακας 3.55 : levels της μεταβλητής BreastSubtype που συσχετίζονται μεταξύ τους.

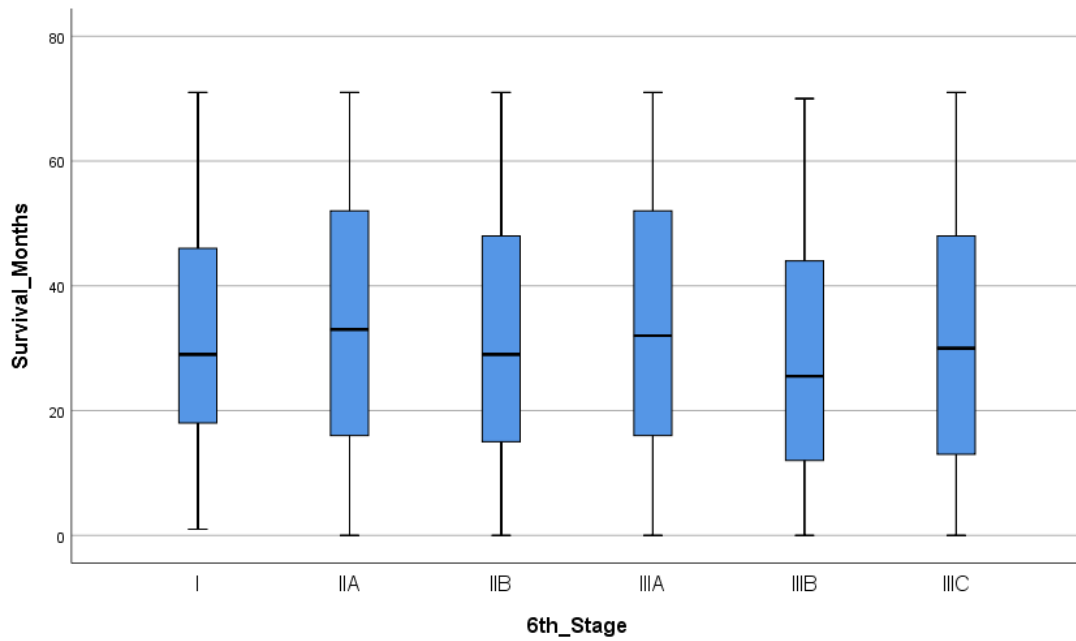
- Έλεγχος συσχέτισης των μηνών επιβίωσης με το 6^ο Στάδιο.

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Survival_Months is the same across categories of 6th_Stage.	Independent-Samples Kruskal-Wallis Test	,000	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

Πίνακας 3.56 : Έλεγχος συσχετίσεων με τη μέθοδο του Kruskal - Wallis - μήνες επιβίωσης με το 6^ο στάδιο.

Για την μεταβλητή Survival_Months το $p\text{-value}=0,000 < 0,05$. Συνεπώς, σε επίπεδο σημαντικότητας 5% απορρίπτουμε την μηδενική υπόθεση. Αυτό σημαίνει ότι τα δύο δείγματα δεν προέρχονται από τον ίδιο πληθυσμό.



Διάγραμμα 3.25 : Θηκόγραμμα μεταξύ των μεταβλητών “Survival_Months” και “6th Stage”.

Παρακάτω παρατηρούμε ποια levels της μεταβλητής 6th Stage συσχετίζονται μεταξύ τους.

Multiple Comparisons

Dependent Variable: Rank of Survival_Months

Tukey HSD

(I) @6th_Stage_new	(J) @6th_Stage_new	Mean Difference	Std. Error	Sig.	95% Confidence Interval	
		(I-J)			Lower Bound	Upper Bound
0	1	-83,947681	208,401803	,999	-678,03956	510,14420
	2	-73,677376	210,080861	,999	-672,55575	525,20100
	3	88,892961	208,052864	,998	-504,20420	681,99012
	4	367,670010	225,523817	,578	-275,23167	1010,57169
	5	119,558777	217,213204	,994	-499,65180	738,76935
1	0	83,947681	208,401803	,999	-510,14420	678,03956
	2	10,270305	61,034891	1,000	-163,72211	184,26272
	3	172,840641*	53,640320	,016	19,92796	325,75332
	4	451,617691*	102,236402	,000	160,17196	743,06342
	5	203,506457	82,297423	,132	-31,09915	438,11206
2	0	73,677376	210,080861	,999	-525,20100	672,55575
	1	-10,270305	61,034891	1,000	-184,26272	163,72211
	3	162,570337	59,832604	,072	-7,99472	333,13539
	4	441,347386*	105,616943	,000	140,26473	742,43004
	5	193,236153	86,461104	,222	-53,23887	439,71118
3	0	-88,892961	208,052864	,998	-681,99012	504,20420
	1	-172,840641*	53,640320	,016	-325,75332	-19,92796
	2	-162,570337	59,832604	,072	-333,13539	7,99472
	4	278,777049	101,523222	,067	-10,63562	568,18972
	5	30,665816	81,409756	,999	-201,40931	262,74094
4	0	-367,670010	225,523817	,578	-1010,57169	275,23167
	1	-451,617691*	102,236402	,000	-743,06342	-160,17196
	2	-441,347386*	105,616943	,000	-742,43004	-140,26473
	3	-278,777049	101,523222	,067	-568,18972	10,63562
	5	-248,111233	119,174437	,297	-587,84229	91,61983
5	0	-119,558777	217,213204	,994	-738,76935	499,65180
	1	-203,506457	82,297423	,132	-438,11206	31,09915
	2	-193,236153	86,461104	,222	-439,71118	53,23887
	3	-30,665816	81,409756	,999	-262,74094	201,40931
	4	248,111233	119,174437	,297	-91,61983	587,84229

*. The mean difference is significant at the 0.05 level.

Πίνακας 3.57 : levels της μεταβλητής 6th Stage που συσχετίζονται μεταξύ τους.

- Έλεγχος συσχέτισης των μηνών επιβίωσης με το Στάδιο T.

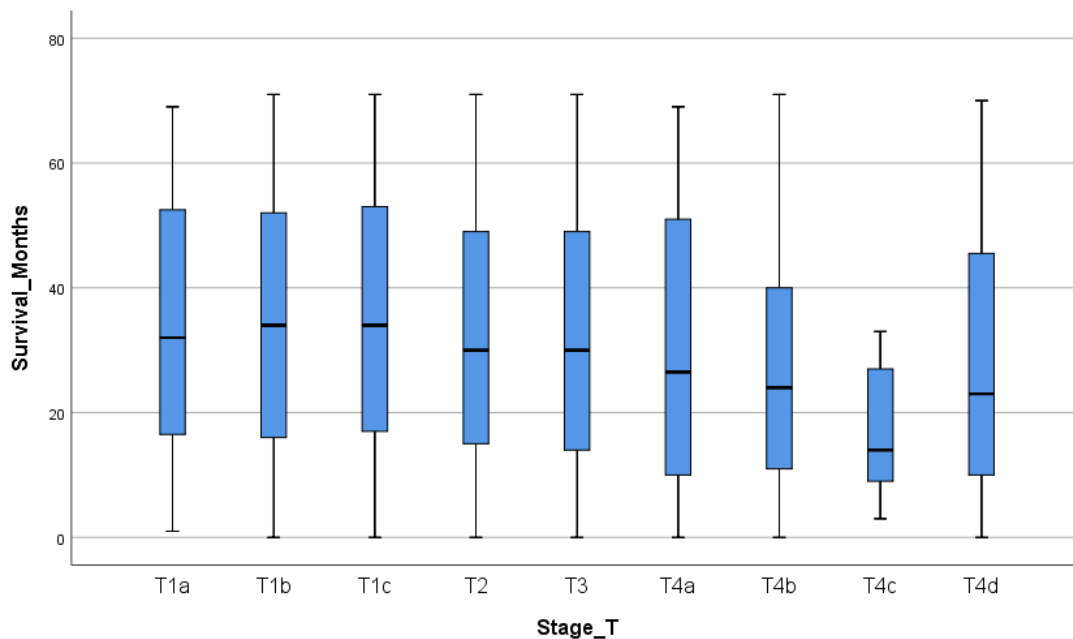
Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Survival_Months is the same across categories of Stage_T.	Independent-Samples Kruskal-Wallis Test	,000	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

Πίνακας 3.58 : Έλεγχος συσχετίσεων με τη μέθοδο του Kruskal - Wallis - μήνες επιβίωσης με το T στάδιο.

Για την μεταβλητή Survival_Months το $p\text{-value}=0,000 < 0,05$. Συνεπώς, σε επίπεδο σημαντικότητας 5% απορρίπτουμε την μηδενική υπόθεση. Αυτό σημαίνει ότι τα δύο δείγματα δεν προέρχονται από τον ίδιο πληθυσμό.



Διάγραμμα 3.26 : Θηκόγραμμα μεταξύ των μεταβλητών “Survival_Months” και “T Stage”.

Παρακάτω παρατηρούμε ποια levels της μεταβλητής Stage_T συσχετίζονται μεταξύ τους:

Multiple Comparisons

Dependent Variable: Rank of Survival_Months

Tukey HSD

(I) Stage_T_new	(J) Stage_T_new	Mean Difference		Sig.	95% Confidence Interval	
		(I-J)	Std. Error		Lower Bound	Upper Bound
0	1	-26,386689	247,498113	1,000	-794,36448	741,59110
	2	-59,134491	234,107286	1,000	-785,56102	667,29204
	3	159,935349	232,482812	,999	-561,45050	881,32119
	4	215,872442	237,253881	,992	-520,31786	952,06274
	5	324,590417	322,502360	,985	-676,12286	1325,30370
	6	588,841967	258,236716	,354	-212,45732	1390,14125
	7	1367,399306	579,723673	,307	-431,46250	3166,26111
	8	525,170611	278,916832	,626	-340,29833	1390,63956
1	0	26,386689	247,498113	1,000	-741,59110	794,36448
	2	-32,747802	99,655192	1,000	-341,97430	276,47869
	3	186,322039	95,776791	,582	-110,86992	483,51399
	4	242,259131	106,837914	,362	-89,25509	573,77335
	5	350,977106	243,172178	,881	-403,57746	1105,53167
	6	615,228657 [*]	147,685941	,001	156,96447	1073,49285
	7	1393,785995	539,633647	,194	-280,67790	3068,24989
	8	551,557300	181,437413	,060	-11,43650	1114,55110
2	0	59,134491	234,107286	1,000	-667,29204	785,56102
	1	32,747802	99,655192	1,000	-276,47869	341,97430
	3	219,069840 [*]	52,192907	,001	57,11712	381,02256
	4	275,006933 [*]	70,464499	,003	56,35812	493,65575
	5	383,724908	229,529113	,764	-328,49572	1095,94553
	6	647,976459 [*]	123,943708	,000	263,38357	1032,56935
	7	1426,533797	533,624754	,158	-229,28472	3082,35231
	8	584,305102 [*]	162,697389	,010	79,46093	1089,14928
3	0	-159,935349	232,482812	,999	-881,32119	561,45050
	1	-186,322039	95,776791	,582	-483,51399	110,86992
	2	-219,069840 [*]	52,192907	,001	-381,02256	-57,11712
	4	55,937093	64,863565	,995	-145,33223	257,20641
	5	164,655068	227,872004	,998	-542,42361	871,73374
	6	428,906618 [*]	120,847339	,012	53,92165	803,89159
	7	1207,463956	532,914078	,363	-446,14935	2861,07727
	8	365,235262	160,351105	,356	-132,32848	862,79900
4	0	-215,872442	237,253881	,992	-952,06274	520,31786
	1	-242,259131	106,837914	,362	-573,77335	89,25509
	2	-275,006933 [*]	70,464499	,003	-493,65575	-56,35812

	3	-55,937093	64,863565	,995	-257,20641	145,33223
	5	108,717975	232,737613	1,000	-613,45851	830,89446
	6	372,969525	129,789158	,095	-29,76159	775,70064
	7	1151,526864	535,012674	,437	-508,59831	2811,65204
	8	309,298169	167,193370	,648	-209,49687	828,09321
5	0	-324,590417	322,502360	,985	-1325,30370	676,12286
	1	-350,977106	243,172178	,881	-1105,53167	403,57746
	2	-383,724908	229,529113	,764	-1095,94553	328,49572
	3	-164,655068	227,872004	,998	-871,73374	542,42361
	4	-108,717975	232,737613	1,000	-830,89446	613,45851
	6	264,251551	254,093671	,982	-524,19202	1052,69512
	7	1042,808889	577,890067	,679	-750,36330	2835,98108
	8	200,580194	275,085426	,998	-653,00004	1054,16042
6	0	-588,841967	258,236716	,354	-1390,14125	212,45732
	1	-615,228657*	147,685941	,001	-1073,49285	-156,96447
	2	-647,976459*	123,943708	,000	-1032,56935	-263,38357
	3	-428,906618*	120,847339	,012	-803,89159	-53,92165
	4	-372,969525	129,789158	,095	-775,70064	29,76159
	5	-264,251551	254,093671	,982	-1052,69512	524,19202
	7	778,557338	544,642413	,887	-911,44858	2468,56325
	8	-63,671357	195,832633	1,000	-671,33301	543,99030
7	0	-1367,399306	579,723673	,307	-3166,26111	431,46250
	1	-1393,785995	539,633647	,194	-3068,24989	280,67790
	2	-1426,533797	533,624754	,158	-3082,35231	229,28472
	3	-1207,463956	532,914078	,363	-2861,07727	446,14935
	4	-1151,526864	535,012674	,437	-2811,65204	508,59831
	5	-1042,808889	577,890067	,679	-2835,98108	750,36330
	6	-778,557338	544,642413	,887	-2468,56325	911,44858
	8	-842,228695	554,746569	,847	-2563,58744	879,13005
8	0	-525,170611	278,916832	,626	-1390,63956	340,29833
	1	-551,557300	181,437413	,060	-1114,55110	11,43650
	2	-584,305102*	162,697389	,010	-1089,14928	-79,46093
	3	-365,235262	160,351105	,356	-862,79900	132,32848
	4	-309,298169	167,193370	,648	-828,09321	209,49687
	5	-200,580194	275,085426	,998	-1054,16042	653,00004
	6	63,671357	195,832633	1,000	-543,99030	671,33301
	7	842,228695	554,746569	,847	-879,13005	2563,58744

*. The mean difference is significant at the 0.05 level.

Πίνακας 3.59 : levels της μεταβλητής Stage_T που συσχετίζονται μεταξύ τους.

- Έλεγχος συσχέτισης των μηνών επιβίωσης με το Στάδιο N.

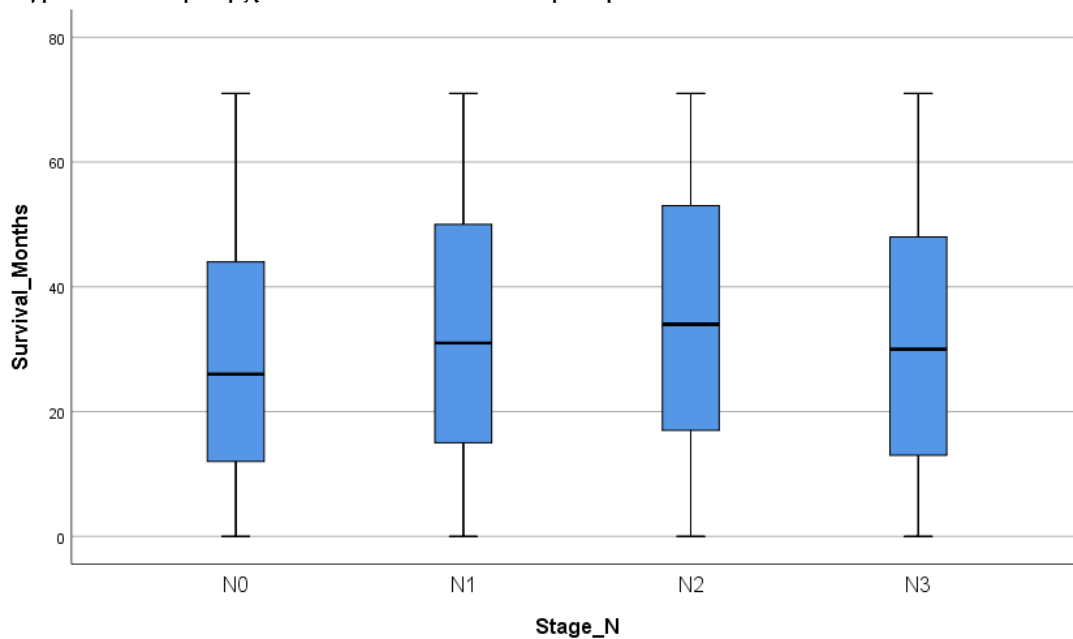
Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Survival_Months is the same across categories of Stage_N.	Independent-Samples Kruskal-Wallis Test	,001	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

Πίνακας 3.60 : Έλεγχος συσχετίσεων με τη μέθοδο του Kruskal - Wallis - μήνες επιβίωσης με το N στάδιο.

Για την μεταβλητή Survival_Months το $p\text{-value}=0,001 < 0,05$. Συνεπώς, σε επίπεδο σημαντικότητας 5% απορρίπτουμε την μηδενική υπόθεση. Αυτό σημαίνει ότι τα δύο δείγματα δεν προέρχονται από τον ίδιο πληθυσμό.



Διάγραμμα 3.27 : Θηκόγραμμα μεταξύ των μεταβλητών “Survival_Months” και “N Stage”.

Παρακάτω παρατηρούμε ποια levels της μεταβλητής Stage_N συσχετίζονται μεταξύ τους:

Multiple Comparisons

Dependent Variable: Rank of Survival_Months

Tukey HSD

(I) Stage_N_new	(J) Stage_N_new	Mean Difference	Std. Error	Sig.	95% Confidence Interval	
		(I-J)			Lower Bound	Upper Bound
0	1	-288,745099*	95,528194	,013	-534,23445	-43,25575
	2	-425,753088*	108,536765	,001	-704,67200	-146,83418
	3	-181,756118	117,279373	,408	-483,14187	119,62964
1	0	288,745099*	95,528194	,013	43,25575	534,23445
	2	-137,007989	62,767146	,128	-298,30766	24,29168
	3	106,988980	76,902123	,505	-90,63490	304,61286
2	0	425,753088*	108,536765	,001	146,83418	704,67200
	1	137,007989	62,767146	,128	-24,29168	298,30766
	3	243,996969*	92,566355	,042	6,11898	481,87495
3	0	181,756118	117,279373	,408	-119,62964	483,14187
	1	-106,988980	76,902123	,505	-304,61286	90,63490
	2	-243,996969*	92,566355	,042	-481,87495	-6,11898

*. The mean difference is significant at the 0.05 level.

Πίνακας 3.61 : levels της μεταβλητής Stage_N που συσχετίζονται μεταξύ τους.

ΚΕΦΑΛΑΙΟ 4

ΠΟΛΛΑΠΛΗ ΠΑΛΙΝΔΡΟΜΗΣΗ

4.1 Εισαγωγή

Στο κεφάλαιο αυτό θα δούμε την μέθοδο της παλινδρόμησης, με την οποία θα αναπτύξουμε σχέση μεταξύ της εξαρτημένης μεταβλητής *survival_months* και των ανεξάρτητων μεταβλητών *Tumor_Size*, *Regional_Nodes_Examined*, *Age*, *Marital_status*, *Grade*, *Stage_A*, *ER_Status*, *PR_Status*, *BreastSubtype*, *6th_Stage*, *Stage_T*, *Stage_N* για τον πληθυσμό που ορίζεται από τη σχέση *Status=1*, δηλαδή για τον πληθυσμό που έχει πεθάνει από καρκίνο του μαστού. Οι ανεξάρτητες μεταβλητές που επιλέχθηκαν είναι με βάση τη σχέση των μεταβλητών με την εξαρτημένη μεταβλητή *survival_months*. Στο συμπέρασμα αυτό καταλήξαμε από τη μελέτη που πραγματοποιήσαμε στο 3^ο Κεφάλαιο.

4.2 Πολλαπλή Παλινδρόμηση

Σε αρκετά προβλήματα η μεταβλητή απόκρισης *Y* μπορεί να θεωρηθεί ότι επηρεάζεται από περισσότερες από μια ερμηνευτικές μεταβλητές, έστω X_1, X_2, \dots, X_{p-1} . Μπορούμε να χρησιμοποιήσουμε ένα γραμμικό μοντέλο για να διερευνήσουμε την εξάρτηση της *Y* από τις X_1, X_2, \dots, X_{p-1} . Το μοντέλο αυτό, το οποίο αποτελεί γενίκευση του μοντέλου $Y = b_0 + b_1 X + \varepsilon$, έχει τη μορφή:

$$Y = b_0 + b_1 X_1 + b_2 X_2 + \dots + b_{p-1} X_{p-1} + \varepsilon,$$

όπου τα «σφάλματα» $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ θεωρούνται ανεξάρτητες τ.μ. από την $N(0,1)$ ενώ οι ερμηνευτικές μεταβλητές X_1, X_2, \dots, X_{p-1} , όπως και στο απλό γραμμικό μοντέλο, δεν θεωρούνται τυχαίες. Το παραπάνω μοντέλο γράφεται με τη βοήθεια πινάκων στην απλούστερη μορφή:

$$Y = X * b + \varepsilon$$

Όπου

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}, \quad X = \begin{bmatrix} 1 & X_{11} & \dots & X_{1,p-1} \\ 1 & X_{21} & & X_{2,p-1} \\ \vdots & \vdots & & \vdots \\ 1 & X_{n1} & \dots & X_{n,p-1} \end{bmatrix}, \quad b = \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_{p-1} \end{bmatrix}, \quad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}.$$

Ο πίνακας *X* καλείται και πίνακας σχεδιασμού. Στο πολλαπλό μοντέλο θεωρούμε ότι τα σημεία δεν βρίσκονται «κοντά» σε μία ευθεία αλλά «κοντά» σε ένα (υπερ)επίπεδο στις *p* διαστάσεις. Η εύρεση του βέλτιστου επιπέδου για την προσέγγιση των

διαθέσιμων σημείων γίνεται με ελαχιστοποίηση του αθροίσματος τετραγώνων των σφαλμάτων:

$$SSE(\beta) = \sum_{i=1}^n \varepsilon_i^2 = \sum_{i=1}^n \left[Y_i - (\hat{\beta}_0 + \hat{\beta}_1 X_{i1} + \hat{\beta}_2 X_{i2} + \hat{\beta}_3 X_{i3} + \varepsilon_i \dots + \hat{\beta}_{p-1} X_{i,p-1}) \right]^2$$

$$= (Y - X\hat{\beta})^T (Y - X\hat{\beta})$$

Για να προσδιορίσουμε τις τιμές των b_0, b_1, \dots, b_{p-1} λύνουμε το σύστημα που θα προκύψει αν θέσουμε τις μερικές παραγώγους ίσες με μηδέν.

4.2.1 Υποθέσεις του πολλαπλού γραμμικού μοντέλου

1. Κανονικότητα: Για κάθε συνδυασμό X_1, X_2, \dots, X_{p-1} η μεταβλητή Y κατανέμεται κανονικά:

$$Y_i \sim N(\mu_{Y|X_1, \dots, X_{p-1}}, \sigma^2) \text{ ή αλλιώς } \varepsilon_i \sim N(0, \sigma^2)$$

2. Ομοσκεδαστικότητα: Η διασπορά των Y_i είναι ίδια για κάθε συνδυασμούς X_1, X_2, \dots, X_{p-1} με:

$$\sigma_{Y/X_1, X_2, \dots, X_{p-1}}^2 \equiv \sigma^2 \text{ ή αλλιώς } \sigma_{Y/X_1, X_2, \dots, X_{p-1}}^2 \equiv \sigma^2$$

3. Ανεξαρτησία: Οι παρατηρήσεις Y_i πρέπει να είναι στατιστικά ανεξάρτητες μεταξύ τους.

4.3 Η Μέθοδος Stepwise Regression

Η μέθοδος της βηματικής παλινδρόμησης (stepwise regression) είναι μία μέθοδος επιλογής ενός "καλού" υποσύνολου ανεξαρτήτων μεταβλητών. Η μέθοδος αυτή είναι παρόμοια με την μέθοδο της προοδευτικής προσθήκης μεταβλητών. Η διαφορά των δύο μεθόδων έγκειται στο γεγονός ότι για κάθε διαδοχικό βήμα η υπόθεση $H_0: \beta_j = 0$ ελέγχεται για όλες τις δυνατές ανεξάρτητες μεταβλητές ώστε να αποκλείονται εκείνες για τις οποίες οι τιμές της στατιστικής συνάρτησης $|T_j|$ είναι μικρότερες από ένα προκαθορισμένο κρίσιμο επίπεδο. Η επόμενη μεταβλητή προστίθεται στο υποσύνολο με την ίδια διαδικασία χρησιμοποίησης του κριτηρίου του μεγίστου συντελεστή συσχέτισης όπως στη μέθοδο της προοδευτικής προσθήκης μεταβλητών. Αυτή η βηματική επιλογή συνεχίζεται μέχρις ότου φθάσουμε σε ένα υποσύνολο μεταβλητών για το οποίο καμιά από τις μεταβλητές που περιέχει το υποσύνολο αυτό δεν έχουν τιμή για τη στατιστική συνάρτηση $|T_j|$ μικρότερη από κάποια συγκεκριμένη κρίσιμη τιμή της μεταβλητής t και δεν υπάρχουν άλλες μεταβλητές που θα πρέπει να αξιολογηθούν για να περιληφθούν στο μοντέλο.

4.3.1 Η μέθοδος Stepwise για τον πληθυσμό ατόμων που πέθαναν από καρκίνο του μαστού - Status = 1 .

Variables Entered/Removed ^a			
Model	Variables Entered	Variables Removed	Method
1	Age_14		. Stepwise (Criteria: Probability-of-F- to-enter <= ,050, Probability-of-F- to-remove >= ,100).
2	Stage_T_3		. Stepwise (Criteria: Probability-of-F- to-enter <= ,050, Probability-of-F- to-remove >= ,100).
3	MaritalStatus_6		. Stepwise (Criteria: Probability-of-F- to-enter <= ,050, Probability-of-F- to-remove >= ,100).
4	BreastSubtype_4		. Stepwise (Criteria: Probability-of-F- to-enter <= ,050, Probability-of-F- to-remove >= ,100).
5	Age_12		. Stepwise (Criteria: Probability-of-F- to-enter <= ,050, Probability-of-F- to-remove >= ,100).

6	Stage_N_3		. Stepwise (Criteria: Probability-of-F- to-enter <= ,050, Probability-of-F- to-remove >= ,100).
7	Stage_T_7		. Stepwise (Criteria: Probability-of-F- to-enter <= ,050, Probability-of-F- to-remove >= ,100).
8	Age_11		. Stepwise (Criteria: Probability-of-F- to-enter <= ,050, Probability-of-F- to-remove >= ,100).
9	Stage_T_2		. Stepwise (Criteria: Probability-of-F- to-enter <= ,050, Probability-of-F- to-remove >= ,100).
10	Age_13		. Stepwise (Criteria: Probability-of-F- to-enter <= ,050, Probability-of-F- to-remove >= ,100).
11	Stage_T_9		. Stepwise (Criteria: Probability-of-F- to-enter <= ,050, Probability-of-F- to-remove >= ,100).

a. Dependent Variable: Survival_Months

Πίνακας 4.1 : Η μέθοδος Stepwise για τον πληθυσμό ατόμων που πέθαναν από καρκίνο του μαστού - Status = 1.

Από τον παραπάνω πίνακα παρατηρούμε πως οι μεταβλητές Age_14, Stage_T3, Marital_Status_6, BreastSubtype_4, Age_12, Stage_N3, Stage_T7, Age_11, Stage_T2, Age_13 και Stage_T9 μπήκαν στο μοντέλο με την μέθοδο stepwise.

		Unstandardized Coefficients		Standardized Coefficients		
Model		B	Std. Error	Beta	t	Sig.
1	(Constant)	33,021	,277		119,069	,000
	Age_14	-8,031	1,440	-,075	-5,578	,000
2	(Constant)	32,217	,322		99,953	,000
	Age_14	-7,693	1,438	-,072	-5,348	,000
	Stage_T_3	3,004	,617	,065	4,867	,000
3	(Constant)	32,326	,323		100,167	,000
	Age_14	-7,813	1,436	-,073	-5,440	,000
	Stage_T_3	3,075	,616	,067	4,989	,000
	MaritalStatus_6	-11,820	2,668	-,059	-4,430	,000
4	(Constant)	32,696	,338		96,650	,000
	Age_14	-7,887	1,435	-,073	-5,496	,000
	Stage_T_3	2,910	,617	,063	4,713	,000
	MaritalStatus_6	-11,731	2,665	-,059	-4,401	,000
	BreastSubtype_4	-3,284	,910	-,048	-3,609	,000
5	(Constant)	32,956	,345		95,426	,000
	Age_14	-8,143	1,435	-,076	-5,674	,000
	Stage_T_3	2,909	,617	,063	4,717	,000
	MaritalStatus_6	-11,839	2,662	-,059	-4,446	,000
	BreastSubtype_4	-3,334	,909	-,049	-3,667	,000
	Age_12	-4,240	1,162	-,049	-3,648	,000
6	(Constant)	32,571	,367		88,743	,000
	Age_14	-8,100	1,434	-,075	-5,648	,000
	Stage_T_3	3,068	,618	,066	4,962	,000
	MaritalStatus_6	-11,767	2,661	-,059	-4,423	,000
	BreastSubtype_4	-3,317	,908	-,049	-3,652	,000
	Age_12	-4,208	1,162	-,048	-3,623	,000
	Stage_N_3	2,408	,782	,041	3,079	,002
7	(Constant)	32,752	,371		88,181	,000
	Age_14	-7,718	1,438	-,072	-5,366	,000
	Stage_T_3	2,869	,621	,062	4,617	,000

	MaritalStatus_6	-11,724	2,659	-,059	-4,410	,000
	BreastSubtype_4	-3,265	,908	-,048	-3,597	,000
	Age_12	-4,172	1,161	-,048	-3,594	,000
	Stage_N_3	2,457	,782	,042	3,143	,002
	Stage_T_7	-4,655	1,513	-,041	-3,077	,002
8	(Constant)	32,946	,380		86,627	,000
	Age_14	-7,931	1,441	-,074	-5,505	,000
	Stage_T_3	2,935	,622	,064	4,721	,000
	MaritalStatus_6	-11,894	2,658	-,059	-4,474	,000
	BreastSubtype_4	-3,307	,908	-,049	-3,644	,000
	Age_12	-4,384	1,164	-,050	-3,767	,000
	Stage_N_3	2,464	,781	,042	3,154	,002
	Stage_T_7	-4,563	1,513	-,040	-3,017	,003
	Age_11	-2,349	,997	-,032	-2,357	,018
9	(Constant)	32,676	,397		82,412	,000
	Age_14	-7,842	1,440	-,073	-5,445	,000
	Stage_T_3	3,185	,630	,069	5,055	,000
	MaritalStatus_6	-11,858	2,657	-,059	-4,462	,000
	BreastSubtype_4	-3,181	,909	-,047	-3,501	,000
	Age_12	-4,402	1,163	-,050	-3,784	,000
	Stage_N_3	2,649	,785	,045	3,376	,001
	Stage_T_7	-4,351	1,515	-,039	-2,873	,004
	Age_11	-2,405	,997	-,032	-2,413	,016
	Stage_T_2	2,870	1,197	,032	2,397	,017
10	(Constant)	32,835	,403		81,495	,000
	Age_14	-8,004	1,442	-,074	-5,552	,000
	Stage_T_3	3,174	,630	,069	5,038	,000
	MaritalStatus_6	-12,002	2,657	-,060	-4,517	,000
	BreastSubtype_4	-3,186	,908	-,047	-3,507	,000
	Age_12	-4,552	1,165	-,052	-3,908	,000
	Stage_N_3	2,609	,785	,045	3,325	,001
	Stage_T_7	-4,244	1,515	-,038	-2,801	,005
	Age_11	-2,555	,999	-,034	-2,558	,011
	Stage_T_2	2,818	1,197	,032	2,355	,019
	Age_13	-3,024	1,379	-,029	-2,193	,028
11	(Constant)	32,954	,407		81,034	,000
	Age_14	-8,011	1,441	-,074	-5,558	,000
	Stage_T_3	3,056	,632	,066	4,835	,000
	MaritalStatus_6	-12,075	2,657	-,060	-4,545	,000
	BreastSubtype_4	-3,117	,909	-,046	-3,430	,001

Age_12	-4,594	1,165	-,053	-3,944	,000
Stage_N_3	2,614	,784	,045	3,332	,001
Stage_T_7	-4,366	1,516	-,039	-2,881	,004
Age_11	-2,570	,998	-,034	-2,575	,010
Stage_T_2	2,702	1,198	,031	2,256	,024
Age_13	-3,053	1,379	-,030	-2,214	,027
Stage_T_9	-4,268	2,011	-,028	-2,122	,034

a. Dependent Variable: Survival_Months

Πίνακας 4.2 : Χαρακτηριστικά των μεταβλητών που μπήκαν στο μοντέλο με τη μέθοδο Stepwise.

Σύμφωνα με τον παραπάνω πίνακα το μοντέλο που προκύπτει είναι:

$$\text{Survival_months} = 32,954 - 8,011 * \text{Age_14} + 3,056 * \text{Stage_T3} - 12,075 * \text{MaritalStatus_6} - 3,117 * \text{BreastSubtype_4} - 4,594 * \text{Age_12} + 2,614 * \text{Stage_N3} - 4,366 * \text{Stage_T7} - 2,570 * \text{Age_11} + 2,702 * \text{Stage_T2} - 3,053 * \text{Age_13} - 4,268 * \text{Stage_T9}$$

Έλεγχος κανονικότητας των σφαλμάτων :

One-Sample Kolmogorov-Smirnov Test

		Standardized Residual	Unstandardized Residual
N		5552	5552
Normal Parameters ^{a,b}	Mean	,0000000	,0000000
	Std. Deviation	,99900870	20,07685790
Most Extreme Differences	Absolute	,067	,067
	Positive	,067	,067
	Negative	-,054	-,054
Test Statistic		,067	,067
Asymp. Sig. (2-tailed)		,000 ^c	,000 ^c

a. Test distribution is Normal.

b. Calculated from data.

c. Lilliefors Significance Correction.

Πίνακας 4.3 : Έλεγχος κανονικότητας των σφαλμάτων.

Από τον παραπάνω πίνακα παρατηρούμε πως τα σφάλματα με τη μέθοδο της γραμμικής παλινδρόμησης δεν ακολουθούν την κανονική κατανομή σε επίπεδο σημαντικότητας 5%. Συνεπώς δεν μπορούμε να συνεχίσουμε σε στατιστική συμπερασματολογία.

ΚΕΦΑΛΑΙΟ 5

ΛΟΓΙΣΤΙΚΗ ΠΑΛΙΝΔΡΟΜΗΣΗ

5.1 Εισαγωγή

Στο κεφάλαιο αυτό θα δούμε την μέθοδο της λογιστικής παλινδρόμησης, με την οποία θα αναπτύξουμε σχέση μεταξύ της εξαρτημένης μεταβλητής Status και των ανεξάρτητων μεταβλητών Age, Marital_status, Grade, Stage_A, ER_Status, PR_Status, BreastSubtype, 6th_Stage, Stage_T, Stage_N, Tumor_Size, Regional_Nodes_Examined, RegionalNodesPositive. Οι ανεξάρτητες μεταβλητές που επιλέχθηκαν είναι με βάση τη σχέση των μεταβλητών με την εξαρτημένη μεταβλητή Status. Στο συμπέρασμα αυτό καταλήξαμε από τη μελέτη που πραγματοποιήσαμε στο 3ο Κεφάλαιο.

5.2 Λογιστική Παλινδρόμηση

Η λογιστική παλινδρόμηση στη βασική της μορφή χρησιμοποιεί εξαρτημένη μεταβλητή η οποία μπορεί να λάβει δύο πιθανές τιμές, όπως η επιτυχία / αποτυχία που αντιπροσωπεύεται από μια δίτιμη μεταβλητή (δείκτρια), όπου οι δύο τιμές φέρουν την ένδειξη "0" ή "1". Στο λογιστικό μοντέλο, ο λογάριθμος των αποδόσεων (log-odds) για την τιμή με την ένδειξη "1" είναι ένας γραμμικός συνδυασμός μίας ή περισσότερων ανεξάρτητων μεταβλητών δηλαδή "προβλέψεις". Οι ανεξάρτητες μεταβλητές μπορούν καθεμία να είναι μια δυαδική μεταβλητή (δύο κατηγορίες, κωδικοποιημένες από μια μεταβλητή δείκτρια) ή μια συνεχής μεταβλητή με οποιαδήποτε πραγματική τιμή. Η συνάρτηση που μετατρέπει τις αποδόσεις σε πιθανότητες είναι η λογιστική συνάρτηση, εξ ου και το όνομα logit. Το καθοριστικό χαρακτηριστικό του λογιστικού μοντέλου είναι ότι η αύξηση μίας από τις ανεξάρτητες μεταβλητές κλιμακώνει πολλαπλασιαστικά τις πιθανότητες του δεδομένου αποτελέσματος, με κάθε ανεξάρτητη μεταβλητή να έχει τη δική της παράμετρο.

Το γραμμικό λογιστικό μοντέλο για την πρόβλεψη πραγματοποίησης της πιθανότητας της εξαρτημένης μεταβλητής p_i σε σχέση με τις k ανεξάρτητες μεταβλητές $x_{1i}, x_{2i}, \dots, x_{ki}$ δίνεται από την σχέση:

$$\log_i t(p_i) = \log(p_i / (1-p_i)) = \beta_0 + \beta_1 * x_{1i} + \beta_2 * x_{2i} + \dots + \beta_k * x_{ki} ,$$

$$p_i = \exp\{\beta_0 + \beta_1 * x_{1i} + \beta_2 * x_{2i} + \dots + \beta_k * x_{ki}\} / [1 + \exp\{\beta_0 + \beta_1 * x_{1i} + \beta_2 * x_{2i} + \dots + \beta_k * x_{ki}\}] ,$$

$$p_i = e^{n_i} / (1 + e^{n_i}) , \text{ Όπου, } n_i = \sum_j \beta_j x_{ji} .$$

Πρέπει να αναφέρουμε ότι η σχέση μεταξύ των ανεξάρτητων μεταβλητών και της πιθανότητας είναι μη γραμμική, οι εκτιμητές της πιθανότητας βρίσκονται μεταξύ 0 και 1, και τέλος, η εκτίμηση των παραμέτρων β_i θα γίνει με τη μέθοδο της μέγιστης πιθανοφάνειας. Η συνάρτηση πιθανοφάνειας δίνεται από τη σχέση:

$$L(\beta) = \prod_{i=1}^n (p_i^{y_i} (1 - p_i)^{n - y_i})$$

όπου Y_i είναι η τυχαία μεταβλητή του υποδείγματος. Η συνάρτηση πιθανοφάνειας εξαρτάται από τα p_i τα οποία, με την σειρά τους, εξαρτώνται από τα β , άρα η συνάρτηση πιθανοφάνειας εξαρτάται τελικώς, από τα β .

5.3 Εφαρμογή του μοντέλου της λογιστικής παλινδρόμησης με εξαρτημένη μεταβλητή τη μεταβλητή "Status".

Case Processing Summary

Unweighted Cases ^a		N	Percent
Selected Cases	Included in Analysis	5552	100,0
	Missing Cases	0	,0
	Total	5552	100,0
Unselected Cases		0	,0
Total		5552	100,0

a. If weight is in effect, see classification table for the total number of cases.

Πίνακας 5.1 : Περίληψη επεξεργασίας του μοντέλου της λογιστικής παλινδρόμησης με εξαρτημένη μεταβλητή τη μεταβλητή "Status".

Όπως παρατηρούμε από τον παραπάνω πίνακα, δεν υπάρχουν missing values.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	15,242	8	,055

Πίνακας 5.2 : Έλεγχος Hosmer and Lemeshow.

H_0 : Το μοντέλο έχει καλή προσαρμογή. Έναντι της εναλλακτικής έναντι της εναλλακτικής

H_1 : Το μοντέλο δεν έχει καλή προσαρμογή

Το p-value του ελέγχου ισούται με $0.055 > 0.05$, συνεπώς σε επίπεδο σημαντικότητας 5% δεχόμαστε την μηδενική υπόθεση, άρα το μοντέλο έχει καλή προσαρμογή.

Classification Table^a

	Observed	Predicted		Percentage Correct	
		Alive	Dead		
Step 1	Status	Alive	4947	48	99,0
		Dead	498	59	10,6
Overall Percentage					90,2

a. The cut value is ,500

Πίνακας 5.3 : Classification table με εξαρτημένη μεταβλητή το "Status".

Από το classification table, παρατηρούμε πως το μοντέλο προβλέπει 107 άτομα να πεθάνουν από καρκίνο του μαστού και 5445 να μην πεθάνουν από αυτή την αιτία.

Συγκεκριμένα, παρατηρούμε πως έχουν προβλεφθεί σωστά 59 άτομα που απεβίωσαν κατά τη διάρκεια της έρευνας από καρκίνο του μαστού ενώ έχουν προβλεφθεί λανθασμένα 48 άτομα που δεν απεβίωσαν κατά τη διάρκεια της έρευνας από καρκίνο του μαστού. Επίσης, έχουν προβλεφθεί σωστά 4947 άτομα που δεν πέθαναν από καρκίνο του μαστού και λανθασμένα 498 άτομα που απεβίωσαν από καρκίνο του μαστού.

Το ποσοστό ορθής ταξινόμησης είναι 90,2%.

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 ^a	Age_1	-,133	1,256	,011	1	,916	,876
	Age_2	-2,284	,548	17,358	1	,000	,102
	Age_3	-2,216	,373	35,326	1	,000	,109
	Age_4	-2,983	,372	64,246	1	,000	,051
	Age_5	-2,701	,289	87,593	1	,000	,067
	Age_6	-2,457	,251	95,621	1	,000	,086
	Age_7	-2,359	,240	96,962	1	,000	,095
	Age_8	-2,387	,245	94,710	1	,000	,092
	Age_9	-2,008	,224	80,599	1	,000	,134
	Age_10	-2,202	,237	86,465	1	,000	,111
	Age_11	-1,771	,237	55,764	1	,000	,170
	Age_12	-1,007	,222	20,610	1	,000	,365
	Age_13	-,970	,233	17,252	1	,000	,379
	MaritalStatus_1	,250	,555	,202	1	,653	1,284
	MaritalStatus_2	-,091	,541	,028	1	,866	,913
	MaritalStatus_3	-,101	,556	,033	1	,857	,904
	MaritalStatus_4	,446	,551	,654	1	,419	1,562
	MaritalStatus_5	,117	,670	,031	1	,861	1,125
	Grade_1	-,439	,190	5,363	1	,021	,645
	Grade_2	-,609	,121	25,417	1	,000	,544
BreastSubtype_1	,247	,373	,439	1	,507	1,281	

BreastSubtype_2	-,214	,381	,315	1	,575	,808
BreastSubtype_3	-1,312	,242	29,331	1	,000	,269
@6th_Stage_1	-,302	,468	,416	1	,519	,739
Stage_T_1	-1,382	,825	2,809	1	,094	,251
Stage_T_2	-1,036	,462	5,022	1	,025	,355
Stage_T_3	-1,123	,390	8,275	1	,004	,325
Stage_T_4	-,665	,363	3,351	1	,067	,514
Stage_T_5	-,420	,374	1,257	1	,262	,657
Stage_T_6	-,583	,545	1,144	1	,285	,558
Stage_T_7	-,021	,397	,003	1	,957	,979
Stage_T_8	,672	,814	,682	1	,409	1,959
Stage_N_1	-,634	,257	6,091	1	,014	,530
Stage_N_2	-,755	,177	18,266	1	,000	,470
Stage_N_3	-,398	,200	3,972	1	,046	,671
StageA_new	,169	,263	,412	1	,521	1,184
ER_Status_new	-1,004	,331	9,222	1	,002	,366
PR_Status_new	-,574	,148	14,969	1	,000	,563
Tumor_Size (mm)-new	-,003	,004	,717	1	,397	,997
RegionalNodesExamined	,014	,003	17,176	1	,000	1,014
RegionalNodesPositive	-,007	,003	5,623	1	,018	,993
Constant	2,575	,884	8,492	1	,004	13,138

a. Variable(s) entered on step 1: Age_1, Age_2, Age_3, Age_4, Age_5, Age_6, Age_7, Age_8, Age_9, Age_10, Age_11, Age_12, Age_13, MaritalStatus_1, MaritalStatus_2, MaritalStatus_3, MaritalStatus_4, MaritalStatus_5, Grade_1, Grade_2, BreastSubtype_1, BreastSubtype_2, BreastSubtype_3, @6th_Stage_1, Stage_T_1, Stage_T_2, Stage_T_3, Stage_T_4, Stage_T_5, Stage_T_6, Stage_T_7, Stage_T_8, Stage_N_1, Stage_N_2, Stage_N_3, StageA_new, ER_Status_new, PR_Status_new, Tumor_Size (mm)-new, RegionalNodesExamined, RegionalNodesPositive.

Πίνακας 5.4 : Ερμηνεία μεταβλητών.

Το τεστ του Wald (στήλη " Wald ") χρησιμοποιείται για τον προσδιορισμό της στατιστικής σημασίας για καθεμία από τις ανεξάρτητες μεταβλητές. Η στατιστική σημασία του τεστ βρίσκεται στη στήλη " Sig. ". Σύμφωνα με τον έλεγχο του Wald λοιπόν προκύπτει ότι οι μεταβλητές Age2, Age3, Age4, Age5, Age6, Age7, Age8, Age9, Age10, Age11, Age12, Grade_1, Grade_2, BreastSubtype_3, Stage_T2, Stage_T3, Stage_N1, Stage_N2, Stage_N3, ER_Status, PR_Status, Regional_NodesExamined, Regional_Nodes_Positive και ο σταθερός όρος προκύπτουν στατιστικά σημαντικές για την πιθανότητα επιβίωσης.

$$\begin{aligned} \ln \left(\frac{p}{1-p} \right) = & 2,575 - 2.284 * \text{Age2} - 2.216 * \text{Age3} - 2.983 * \text{Age4} - 2.701 * \text{Age5} - 2.457 * \text{Age5} \\ & - 2,359 * \text{Age6} - 2.387 * \text{Age7} - 2.008 * \text{Age8} - 2.202 * \text{Age9} - 1.771 * \text{Age10} - 1.007 * \text{Age11} - \\ & 0.970 * \text{Age12} - 0.439 * \text{Grade}_1 - 0.609 * \text{Grade}_2 - 1.312 * \text{BreastSubtype}_3 - \\ & 1.036 * \text{Stage}_T2 - 1.123 * \text{Stage}_T3 - 0.634 * \text{Stage}_N1 - 0.755 * \text{Stage}_N2 - \\ & 0.398 * \text{Stage}_N3 - 1.004 * \text{ER}_\text{Status} - 0.574 * \text{PR}_\text{Status} + \\ & 0.014 * \text{RegionalNodesExamined} - 0.007 * \text{RegionalNodesPositive} \end{aligned}$$

Σύμφωνα με αυτό το μοντέλο έχουμε:

Όταν κάποιος ανήκει στην ηλικιακή ομάδα 25-29 το odds μειώνεται πολλαπλασιαστικά κατά 0,102, δηλαδή αυξάνεται η πιθανότητα αποβίωσης και μειώνεται η πιθανότητα επιβίωσης, σε σχέση με την ηλικιακή ομάδα 85+.

Όταν κάποιος ανήκει στην ηλικιακή ομάδα 30-34 το odds μειώνεται πολλαπλασιαστικά κατά 0,109, δηλαδή αυξάνεται η πιθανότητα αποβίωσης και μειώνεται η πιθανότητα επιβίωσης, σε σχέση με την ηλικιακή ομάδα 85+.

Όταν κάποιος ανήκει στην ηλικιακή ομάδα 35-39 το odds μειώνεται πολλαπλασιαστικά κατά 0,051, δηλαδή αυξάνεται η πιθανότητα αποβίωσης και μειώνεται η πιθανότητα επιβίωσης, σε σχέση με την ηλικιακή ομάδα 85+.

Όταν κάποιος ανήκει στην ηλικιακή ομάδα 40-44 το odds μειώνεται πολλαπλασιαστικά κατά 0,067, δηλαδή αυξάνεται η πιθανότητα αποβίωσης και μειώνεται η πιθανότητα επιβίωσης, σε σχέση με την ηλικιακή ομάδα 85+.

Όταν κάποιος ανήκει στην ηλικιακή ομάδα 45-49 το odds μειώνεται πολλαπλασιαστικά κατά 0,086, δηλαδή αυξάνεται η πιθανότητα αποβίωσης και μειώνεται η πιθανότητα επιβίωσης, σε σχέση με την ηλικιακή ομάδα 85+.

Όταν κάποιος ανήκει στην ηλικιακή ομάδα 50-54 το odds μειώνεται πολλαπλασιαστικά κατά 0,095, δηλαδή αυξάνεται η πιθανότητα αποβίωσης και μειώνεται η πιθανότητα επιβίωσης, σε σχέση με την ηλικιακή ομάδα 85+.

Όταν κάποιος ανήκει στην ηλικιακή ομάδα 55-59 το odds μειώνεται πολλαπλασιαστικά κατά 0,092, δηλαδή αυξάνεται η πιθανότητα αποβίωσης και μειώνεται η πιθανότητα επιβίωσης, σε σχέση με την ηλικιακή ομάδα 85+.

Όταν κάποιος ανήκει στην ηλικιακή ομάδα 60-64 το odds μειώνεται πολλαπλασιαστικά κατά 0,134, δηλαδή αυξάνεται η πιθανότητα αποβίωσης και μειώνεται η πιθανότητα επιβίωσης, σε σχέση με την ηλικιακή ομάδα 85+.

Όταν κάποιος ανήκει στην ηλικιακή ομάδα 65-69 το odds μειώνεται πολλαπλασιαστικά κατά 0,111, δηλαδή αυξάνεται η πιθανότητα αποβίωσης και μειώνεται η πιθανότητα επιβίωσης, σε σχέση με την ηλικιακή ομάδα 85+.

Όταν κάποιος ανήκει στην ηλικιακή ομάδα 70-74 το odds μειώνεται πολλαπλασιαστικά κατά 0,170, δηλαδή αυξάνεται η πιθανότητα αποβίωσης και μειώνεται η πιθανότητα επιβίωσης, σε σχέση με την ηλικιακή ομάδα 85+.

Όταν κάποιος ανήκει στην ηλικιακή ομάδα 75-79 το odds μειώνεται πολλαπλασιαστικά κατά 0,365, δηλαδή αυξάνεται η πιθανότητα αποβίωσης και μειώνεται η πιθανότητα επιβίωσης, σε σχέση με την ηλικιακή ομάδα 85+.

Όταν κάποιος ανήκει στην ηλικιακή ομάδα 80-84 το odds μειώνεται πολλαπλασιαστικά κατά 0,379, δηλαδή αυξάνεται η πιθανότητα αποβίωσης και μειώνεται η πιθανότητα επιβίωσης, σε σχέση με την ηλικιακή ομάδα 85+.

Όταν τα καρκινικά κύτταρα είναι βαθμού I το odds μειώνεται πολλαπλασιαστικά κατά 0,645, δηλαδή αυξάνεται η πιθανότητα αποβίωσης και μειώνεται η πιθανότητα επιβίωσης, σε σχέση με το βαθμό III.

Όταν τα καρκινικά κύτταρα είναι βαθμού II το odds μειώνεται πολλαπλασιαστικά κατά 0,544, δηλαδή αυξάνεται η πιθανότητα αποβίωσης και μειώνεται η πιθανότητα επιβίωσης, σε σχέση με το βαθμό III.

Όταν ο δευτερεύον τύπος μαστού είναι «Τριπλός Αρνητικός» το odds μειώνεται πολλαπλασιαστικά κατά 0,269, δηλαδή αυξάνεται η πιθανότητα αποβίωσης και μειώνεται η πιθανότητα επιβίωσης, σε σχέση με το HER2 - enriched.

Όταν το Stage T έχει την τιμή «T1b» το odds μειώνεται πολλαπλασιαστικά κατά 0,355, δηλαδή αυξάνεται η πιθανότητα αποβίωσης και μειώνεται η πιθανότητα επιβίωσης, σε σχέση με το T4d.

Όταν το Stage T έχει την τιμή «T1c» το odds μειώνεται πολλαπλασιαστικά κατά 0,325, δηλαδή αυξάνεται η πιθανότητα αποβίωσης και μειώνεται η πιθανότητα επιβίωσης, σε σχέση με το T4d.

Όταν το Stage N έχει την τιμή «N0» το odds μειώνεται πολλαπλασιαστικά κατά 0,530, δηλαδή αυξάνεται η πιθανότητα αποβίωσης και μειώνεται η πιθανότητα επιβίωσης, σε σχέση με το N3.

Όταν το Stage N έχει την τιμή «N1» το odds μειώνεται πολλαπλασιαστικά κατά 0,470, δηλαδή αυξάνεται η πιθανότητα αποβίωσης και μειώνεται η πιθανότητα επιβίωσης, σε σχέση με το N3.

Όταν το Stage N έχει την τιμή «N2» το odds μειώνεται πολλαπλασιαστικά κατά 0,671, δηλαδή αυξάνεται η πιθανότητα αποβίωσης και μειώνεται η πιθανότητα επιβίωσης, σε σχέση με το N3.

Όταν το ER_Status είναι θετικό, το odds μειώνεται πολλαπλασιαστικά κατά 0,366, δηλαδή αυξάνεται η πιθανότητα αποβίωσης και μειώνεται η πιθανότητα επιβίωσης.

Όταν το PR_Status είναι θετικό, το odds μειώνεται πολλαπλασιαστικά κατά 0,563, δηλαδή αυξάνεται η πιθανότητα αποβίωσης και μειώνεται η πιθανότητα επιβίωσης.

Για αύξηση μίας μονάδας της μεταβλητής Regional_Nodes_Examined, το odds αυξάνεται πολλαπλασιαστικά κατά 1,014, δηλαδή αυξάνεται η πιθανότητα αποβίωσης και μειώνεται η πιθανότητα επιβίωσης.

Για αύξηση μίας μονάδας της μεταβλητής Regional_Nodes_Positive, το odds μειώνεται πολλαπλασιαστικά κατά 0,993, δηλαδή αυξάνεται η πιθανότητα αποβίωσης και μειώνεται η πιθανότητα επιβίωσης.

ΚΕΦΑΛΑΙΟ 6

ΜΟΝΤΕΛΟ ΤΟΥ COX

6.1 Εισαγωγή

Στο κεφάλαιο αυτό θα εφαρμόσουμε την μέθοδο Cox. Θεωρούμε ότι η μεταβλητή *Survival_months* είναι ο χρόνος μέχρι να συμβεί το γεγονός του θανάτου και η μεταβλητή *status* (0 = λογοκριμένη και 1 = full). Οι μεταβλητές *Age*, *Marital_status*, *Grade*, *Stage_A*, *ER_Status*, *PR_Status*, *BreastSubtype*, *6th_Stage*, *Stage_T*, *Stage_N*, *Tumor_Size*, *Regional_Nodes_Examined*, *RegionalNodesPositive* είναι οι ανεξάρτητες μεταβλητές.

6.2 Μοντέλο Κινδύνων του Cox

Το μοντέλο κινδύνων του Cox αναπαριστάται από τον γνωστό λόγο των κινδύνων:

$$h(tt; XX) = h_0(tt) e^{\sum_{ii=1}^{nn} b_{ii} X_{ii}}$$

Αυτό το μοντέλο δίνει μια έκφραση της συνάρτησης κινδύνου στον χρόνο *t* για μια συγκεκριμένη περίπτωση με τις συγκεκριμένες τιμές των συνοδευτικών μεταβλητών *X* που είναι στο μοντέλο. Οι μεταβλητές είναι μια συλλογή από τις παραμέτρους που έχουν μια σημαντική επίδραση στο ρυθμό ρίσκου του κάθε ατόμου. Στην παραπάνω σχέση η συνάρτηση βάσης $h_0(tt)$ εξαρτάται από τον χρόνο ενώ το εκθετικό μέρος είναι ανεξάρτητο του χρόνου. Στην περίπτωση που οι *X* δεν είναι ανεξάρτητες από τον χρόνο ακολουθούμε άλλη μεθοδολογία που λέγεται και εκτεταμένο μοντέλο του Cox. Οι συντελεστές συσχέτισης σε κάθε περίπτωση (b_{ii}) εκτιμούνται χωρίς υποθέσεις σχετικά με την συνάρτηση κινδύνου βάσης $h_0(tt)$ ή τον πληθυσμό που πήραμε τα δεδομένα. Αυτή η ιδιότητα κάνει το μοντέλο κινδύνων του Cox πολύ εύρωστη μέθοδο.

6.3 Θνησιμότητα από τον καρκίνο του μαστού

Θεωρούμε ότι η μεταβλητή *Survival_months* είναι ο χρόνος μέχρι να συμβεί το γεγονός του θανάτου και η μεταβλητή *status* με την τιμή 0 είναι λογοκριμένη και με την 1 full. Οι μεταβλητές *Age*, *Marital_status*, *Grade*, *Stage_A*, *ER_Status*, *PR_Status*, *BreastSubtype*, *6th_Stage*, *Stage_T*, *Stage_N*, *Tumor_Size*, *Regional_Nodes_Examined*, *RegionalNodesPositive* είναι οι ανεξάρτητες μεταβλητές.

Case Processing Summary

		N	Percent
Cases available in analysis	Event ^a	557	10,0%
	Censored	4995	90,0%
	Total	5552	100,0%
Cases dropped	Cases with missing values	0	0,0%
	Cases with negative time	0	0,0%
	Censored cases before the earliest event in a stratum	0	0,0%
	Total	0	0,0%
Total		5552	100,0%

a. Dependent Variable: Survival_Months

Πίνακας 6.1 : Περίληψη επεξεργασίας του μοντέλου του Cox με εξαρτημένη μεταβλητή τη μεταβλητή "Survival_Months".

Ο παραπάνω πίνακας δίνει πληροφορίες σχετικά με τα δεδομένα. Οι 557 από τις συνολικά 5552 περιπτώσεις έχουν εμφανίσει το ενδεχόμενο (θάνατος από καρκίνο του μαστού) ενώ οι υπόλοιπες 4995 είναι λογοκριμένοι χρόνοι επιβίωσης. Δεν υπάρχει καμία ελλείπουσα τιμή, κανένας αρνητικός χρόνος και καμία λογοκριμένη περίπτωση πριν τον πρώτο θάνατο από καρκίνο του μαστού σε κάποιο από τα στρώματα.

Omnibus Tests of Model Coefficients^a

-2 Log Likelihood	Overall (score)			Change From Previous Step			Change From Previous Block		
	Chi-square	df	Sig.	Chi-square	df	Sig.	Chi-square	df	Sig.
8298,831	966,219	41	,000	662,138	41	,000	662,138	41	,000

a. Beginning Block Number 1. Method = Enter

Πίνακας 6.2 : Δοκιμές γενικών συντελεστών μοντέλου.

Για τον έλεγχο : $H_0 : \beta_1 = \dots = \beta_j = 0$ έναντι της εναλλακτικής $H_1 : \beta_i \neq 0$ ή \dots ή $\beta_j \neq 0$, $i = 1$ και $j = 13$

Έχουμε ότι η στατιστική συνάρτηση score ισούται με 966,219 με p-value = 0,000 και η στατιστική συνάρτηση LR ισούται με 498,440 με p-value = 0,000. Με βάση και τους δύο ελέγχους απορρίπτουμε τη μηδενική υπόθεση σε επίπεδο σημαντικότητας 5%. Αυτό σημαίνει ότι το μοντέλο είναι στατιστικά σημαντικό.

Variables in the Equation^b

	B	SE	Wald	df	Sig.	Exp(B)	95,0% CI for Exp(B)	
							Lower	Upper
Age_1	,452	1,018	,197	1	,657	1,571	,214	11,548
Age_2	1,847	,482	14,654	1	,000	6,340	2,463	16,322
Age_3	1,897	,325	34,080	1	,000	6,663	3,525	12,596
Age_4	2,550	,336	57,596	1	,000	12,813	6,631	24,758
Age_5	2,364	,251	88,561	1	,000	10,635	6,500	17,401
Age_6	2,036	,208	95,487	1	,000	7,657	5,090	11,518
Age_7	1,975	,199	98,342	1	,000	7,207	4,878	10,648
Age_8	1,993	,205	94,587	1	,000	7,340	4,912	10,968
Age_9	1,651	,183	81,435	1	,000	5,213	3,642	7,462
Age_10	1,858	,200	86,085	1	,000	6,410	4,329	9,491
Age_11	1,387	,198	48,905	1	,000	4,005	2,715	5,908
Age_12	,571	,176	10,567	1	,001	1,771	1,255	2,499
Age_13	,672	,182	13,605	1	,000	1,958	1,370	2,799
Age_14			.	0 ^a	.			
MaritalStatus_1	,498	,518	,926	1	,336	1,646	,596	4,542
MaritalStatus_2	,701	,507	1,909	1	,167	2,015	,746	5,442
MaritalStatus_3	,671	,520	1,666	1	,197	1,957	,706	5,423
MaritalStatus_4	,266	,515	,266	1	,606	1,305	,475	3,583
MaritalStatus_5	,472	,619	,581	1	,446	1,603	,476	5,399
MaritalStatus_6			.	0 ^a	.			
Grade_1	,480	,175	7,561	1	,006	1,616	1,148	2,275
Grade_2	,592	,108	30,053	1	,000	1,808	1,463	2,235
Grade_3			.	0 ^a	.			
StageA_new	,154	,216	,507	1	,476	1,166	,764	1,779
ER_Status_new	,972	,276	12,402	1	,000	2,644	1,539	4,543
PR_Status_new	,447	,130	11,878	1	,001	1,564	1,213	2,016
BreastSubtype_1	-,426	,314	1,837	1	,175	,653	,353	1,209
BreastSubtype_2	,042	,318	,018	1	,894	1,043	,559	1,947
BreastSubtype_3	1,201	,210	32,707	1	,000	3,322	2,202	5,013
BreastSubtype_4			.	0 ^a	.			
@6th_Stage_1	,376	,411	,839	1	,360	1,456	,651	3,256
Stage_T_1	1,697	,777	4,774	1	,029	5,458	1,191	25,011
Stage_T_2	1,431	,396	13,052	1	,000	4,184	1,925	9,094
Stage_T_3	1,535	,327	22,081	1	,000	4,641	2,447	8,803
Stage_T_4	1,000	,301	11,034	1	,001	2,719	1,507	4,905
Stage_T_5	,740	,318	5,410	1	,020	2,097	1,124	3,913
Stage_T_6	,686	,463	2,200	1	,138	1,986	,802	4,920

Stage_T_7	,238	,324	,537	1	,464	1,268	,672	2,394
Stage_T_8	-,414	,588	,498	1	,481	,661	,209	2,090
Stage_T_9			.	0 ^a	.			
Stage_N_1	,538	,217	6,164	1	,013	1,713	1,120	2,620
Stage_N_2	,668	,148	20,496	1	,000	1,951	1,461	2,606
Stage_N_3	,417	,168	6,134	1	,013	1,518	1,091	2,111
Stage_N_4			.	0 ^a	.			
Tumor_Size (mm)-new	-,003	,003	1,262	1	,261	,997	,990	1,003
RegionalNodesExamined	,012	,003	14,865	1	,000	1,012	1,006	1,018
RegionalNodesPositive	-,002	,002	,624	1	,430	,998	,993	1,003

a. Degree of freedom reduced because of constant or linearly dependent covariates

b. Constant or Linearly Dependent Covariates Age_14 = 13 - Age_1 - Age_2 - Age_3 - Age_4 - Age_5 - Age_6 - Age_7 - Age_8 - Age_9 - Age_10 - Age_11 - Age_12 - Age_13 ; MaritalStatus_6 = 5 - MaritalStatus_1 - MaritalStatus_2 - MaritalStatus_3 - MaritalStatus_4 - MaritalStatus_5 ; Grade_3 = 2 - Grade_1 - Grade_2 ; BreastSubtype_4 = 3 - BreastSubtype_1 - BreastSubtype_2 - BreastSubtype_3 ; Stage_T_9 = 8 - Stage_T_1 - Stage_T_2 - Stage_T_3 - Stage_T_4 - Stage_T_5 - Stage_T_6 - Stage_T_7 - Stage_T_8 ; Stage_N_4 = 3 - Stage_N_1 - Stage_N_2 - Stage_N_3 ;

Πίνακας 6.3 : Ερμηνεία μεταβλητών.

Από τον παραπάνω πίνακα παρατηρούμε ότι:

Για τον έλεγχο : $H_0 : \beta_1=0$ έναντι της εναλλακτικής $H_1 : \beta_1 \neq 0$

Παρατηρούμε ότι η στατιστική συνάρτηση ισούται με 0,197 με p-value = 0,657, συνεπώς σε επίπεδο σημαντικότητας 5% δεχόμαστε τη μηδενική υπόθεση. Αυτό σημαίνει ότι η μεταβλητή Age_1 δεν είναι στατιστικά σημαντική.

Για τον έλεγχο : $H_0 : \beta_2=0$ έναντι της εναλλακτικής $H_1 : \beta_2 \neq 0$

Παρατηρούμε ότι η στατιστική συνάρτηση ισούται με 14,654 με p-value = 0,000, συνεπώς σε επίπεδο σημαντικότητας 5% απορρίπτουμε τη μηδενική υπόθεση. Αυτό σημαίνει ότι η μεταβλητή Age_2 είναι στατιστικά σημαντική. Το β_2 ισούται 1,847.

Για τον έλεγχο : $H_0 : \beta_3=0$ έναντι της εναλλακτικής $H_1 : \beta_3 \neq 0$

Παρατηρούμε ότι η στατιστική συνάρτηση ισούται με 34,080 με p-value = 0,000, συνεπώς σε επίπεδο σημαντικότητας 5% απορρίπτουμε τη μηδενική υπόθεση. Αυτό σημαίνει ότι η μεταβλητή Age_3 είναι στατιστικά σημαντική. Το β_3 ισούται 1,897.

Για τον έλεγχο : $H_0 : \beta_4=0$ έναντι της εναλλακτικής $H_1 : \beta_4 \neq 0$

Παρατηρούμε ότι η στατιστική συνάρτηση ισούται με 54,596 με p-value = 0,000, συνεπώς σε επίπεδο σημαντικότητας 5% απορρίπτουμε τη μηδενική υπόθεση. Αυτό σημαίνει ότι η μεταβλητή Age_4 είναι στατιστικά σημαντική. Το β_4 ισούται 2,550.

Για τον έλεγχο : $H_0 : \beta_5=0$ έναντι της εναλλακτικής $H_1 : \beta_5 \neq 0$

Παρατηρούμε ότι η στατιστική συνάρτηση ισούται με 88,561 με $p\text{-value} = 0,000$, συνεπώς σε επίπεδο σημαντικότητας 5% απορρίπτουμε τη μηδενική υπόθεση. Αυτό σημαίνει ότι η μεταβλητή Age_5 είναι στατιστικά σημαντική. Το β_5 ισούται 2,364.

Για τον έλεγχο : $H_0 : \beta_6=0$ έναντι της εναλλακτικής $H_1 : \beta_6 \neq 0$

Παρατηρούμε ότι η στατιστική συνάρτηση ισούται με 95,487 με $p\text{-value} = 0,000$, συνεπώς σε επίπεδο σημαντικότητας 5% απορρίπτουμε τη μηδενική υπόθεση. Αυτό σημαίνει ότι η μεταβλητή Age_6 είναι στατιστικά σημαντική. Το β_6 ισούται 2,036.

Για τον έλεγχο : $H_0 : \beta_7=0$ έναντι της εναλλακτικής $H_1 : \beta_7 \neq 0$

Παρατηρούμε ότι η στατιστική συνάρτηση ισούται με 98,342 με $p\text{-value} = 0,000$, συνεπώς σε επίπεδο σημαντικότητας 5% απορρίπτουμε τη μηδενική υπόθεση. Αυτό σημαίνει ότι η μεταβλητή Age_7 είναι στατιστικά σημαντική. Το β_7 ισούται 1,975.

Για τον έλεγχο : $H_0 : \beta_8=0$ έναντι της εναλλακτικής $H_1 : \beta_8 \neq 0$

Παρατηρούμε ότι η στατιστική συνάρτηση ισούται με 94,587 με $p\text{-value} = 0,000$, συνεπώς σε επίπεδο σημαντικότητας 5% απορρίπτουμε τη μηδενική υπόθεση. Αυτό σημαίνει ότι η μεταβλητή Age_8 είναι στατιστικά σημαντική. Το β_8 ισούται 1,975.

Για τον έλεγχο : $H_0 : \beta_9=0$ έναντι της εναλλακτικής $H_1 : \beta_9 \neq 0$

Παρατηρούμε ότι η στατιστική συνάρτηση ισούται με 81,435 με $p\text{-value} = 0,000$, συνεπώς σε επίπεδο σημαντικότητας 5% απορρίπτουμε τη μηδενική υπόθεση. Αυτό σημαίνει ότι η μεταβλητή Age_9 είναι στατιστικά σημαντική. Το β_9 ισούται 1,651.

Για τον έλεγχο : $H_0 : \beta_{10}=0$ έναντι της εναλλακτικής $H_1 : \beta_{10} \neq 0$

Παρατηρούμε ότι η στατιστική συνάρτηση ισούται με 86,085 με $p\text{-value} = 0,000$, συνεπώς σε επίπεδο σημαντικότητας 5% απορρίπτουμε τη μηδενική υπόθεση. Αυτό σημαίνει ότι η μεταβλητή Age_10 είναι στατιστικά σημαντική. Το β_{10} ισούται 1,858.

Για τον έλεγχο : $H_0 : \beta_{11}=0$ έναντι της εναλλακτικής $H_1 : \beta_{11} \neq 0$

Παρατηρούμε ότι η στατιστική συνάρτηση ισούται με 48,905 με $p\text{-value} = 0,000$, συνεπώς σε επίπεδο σημαντικότητας 5% απορρίπτουμε τη μηδενική υπόθεση. Αυτό σημαίνει ότι η μεταβλητή Age_11 είναι στατιστικά σημαντική. Το β_{11} ισούται 1,387.

Για τον έλεγχο : $H_0 : \beta_{12}=0$ έναντι της εναλλακτικής $H_1 : \beta_{12} \neq 0$

Παρατηρούμε ότι η στατιστική συνάρτηση ισούται με 10,567 με $p\text{-value} = 0,001$, συνεπώς σε επίπεδο σημαντικότητας 5% απορρίπτουμε τη μηδενική υπόθεση. Αυτό σημαίνει ότι η μεταβλητή Age_12 είναι στατιστικά σημαντική. Το β_{12} ισούται 0,571.

Για τον έλεγχο : $H_0 : \beta_{13}=0$ έναντι της εναλλακτικής $H_1 : \beta_{13} \neq 0$

Παρατηρούμε ότι η στατιστική συνάρτηση ισούται με 13,605 με $p\text{-value} = 0,000$, συνεπώς σε επίπεδο σημαντικότητας 5% απορρίπτουμε τη μηδενική υπόθεση. Αυτό σημαίνει ότι η μεταβλητή Age_13 είναι στατιστικά σημαντική. Το β_{13} ισούται 0,672.

Για τον έλεγχο : $H_0 : \beta_{14}=0$ έναντι της εναλλακτικής $H_1 : \beta_{14} \neq 0$

Παρατηρούμε ότι η στατιστική συνάρτηση ισούται με 0,926 με $p\text{-value} = 0,336$, συνεπώς σε επίπεδο σημαντικότητας 5% δεχόμαστε τη μηδενική υπόθεση. Αυτό σημαίνει ότι η μεταβλητή MaritalStatus_1 δεν είναι στατιστικά σημαντική.

Για τον έλεγχο : $H_0 : \beta_{15}=0$ έναντι της εναλλακτικής $H_1 : \beta_{15} \neq 0$

Παρατηρούμε ότι η στατιστική συνάρτηση ισούται με 1,909 με p-value = 0,167, συνεπώς σε επίπεδο σημαντικότητας 5% δεχόμαστε τη μηδενική υπόθεση. Αυτό σημαίνει ότι η μεταβλητή MaritalStatus_2 δεν είναι στατιστικά σημαντική.

Για τον έλεγχο : $H_0 : \beta_{16}=0$ έναντι της εναλλακτικής $H_1 : \beta_{16} \neq 0$

Παρατηρούμε ότι η στατιστική συνάρτηση ισούται με 1,666 με p-value = 0,197, συνεπώς σε επίπεδο σημαντικότητας 5% δεχόμαστε τη μηδενική υπόθεση. Αυτό σημαίνει ότι η μεταβλητή MaritalStatus_3 δεν είναι στατιστικά σημαντική.

Για τον έλεγχο : $H_0 : \beta_{17}=0$ έναντι της εναλλακτικής $H_1 : \beta_{17} \neq 0$

Παρατηρούμε ότι η στατιστική συνάρτηση ισούται με 0,266 με p-value = 0,606, συνεπώς σε επίπεδο σημαντικότητας 5% δεχόμαστε τη μηδενική υπόθεση. Αυτό σημαίνει ότι η μεταβλητή MaritalStatus_4 δεν είναι στατιστικά σημαντική.

Για τον έλεγχο : $H_0 : \beta_{18}=0$ έναντι της εναλλακτικής $H_1 : \beta_{18} \neq 0$

Παρατηρούμε ότι η στατιστική συνάρτηση ισούται με 0,581 με p-value = 0,446, συνεπώς σε επίπεδο σημαντικότητας 5% δεχόμαστε τη μηδενική υπόθεση. Αυτό σημαίνει ότι η μεταβλητή MaritalStatus_5 δεν είναι στατιστικά σημαντική.

Για τον έλεγχο : $H_0 : \beta_{19}=0$ έναντι της εναλλακτικής $H_1 : \beta_{19} \neq 0$

Παρατηρούμε ότι η στατιστική συνάρτηση ισούται με 7,561 με p-value = 0,006, συνεπώς σε επίπεδο σημαντικότητας 5% απορρίπτουμε τη μηδενική υπόθεση. Αυτό σημαίνει ότι η μεταβλητή Grade_1 είναι στατιστικά σημαντική. Το β_{19} ισούται 0,480.

Για τον έλεγχο : $H_0 : \beta_{20}=0$ έναντι της εναλλακτικής $H_1 : \beta_{20} \neq 0$

Παρατηρούμε ότι η στατιστική συνάρτηση ισούται με 30,053 με p-value = 0,000, συνεπώς σε επίπεδο σημαντικότητας 5% απορρίπτουμε τη μηδενική υπόθεση. Αυτό σημαίνει ότι η μεταβλητή Grade_2 είναι στατιστικά σημαντική. Το β_{20} ισούται 0,592.

Για τον έλεγχο : $H_0 : \beta_{21}=0$ έναντι της εναλλακτικής $H_1 : \beta_{21} \neq 0$

Παρατηρούμε ότι η στατιστική συνάρτηση ισούται με 0,507 με p-value = 0,476, συνεπώς σε επίπεδο σημαντικότητας 5% δεχόμαστε τη μηδενική υπόθεση. Αυτό σημαίνει ότι η μεταβλητή StageA δεν είναι στατιστικά σημαντική.

Για τον έλεγχο : $H_0 : \beta_{22}=0$ έναντι της εναλλακτικής $H_1 : \beta_{22} \neq 0$

Παρατηρούμε ότι η στατιστική συνάρτηση ισούται με 12,402 με p-value = 0,000, συνεπώς σε επίπεδο σημαντικότητας 5% απορρίπτουμε τη μηδενική υπόθεση. Αυτό σημαίνει ότι η μεταβλητή ER_Status είναι στατιστικά σημαντική. Το β_{22} ισούται 0,972.

Για τον έλεγχο : $H_0 : \beta_{23}=0$ έναντι της εναλλακτικής $H_1 : \beta_{23} \neq 0$

Παρατηρούμε ότι η στατιστική συνάρτηση ισούται με 11,878 με p-value = 0,001, συνεπώς σε επίπεδο σημαντικότητας 5% απορρίπτουμε τη μηδενική υπόθεση. Αυτό σημαίνει ότι η μεταβλητή PR_Status είναι στατιστικά σημαντική. Το β_{23} ισούται 0,447.

Για τον έλεγχο : $H_0 : \beta_{24}=0$ έναντι της εναλλακτικής $H_1 : \beta_{24} \neq 0$

Παρατηρούμε ότι η στατιστική συνάρτηση ισούται με 1,837 με p-value = 0,175, συνεπώς σε επίπεδο σημαντικότητας 5% δεχόμαστε τη μηδενική υπόθεση. Αυτό σημαίνει ότι η μεταβλητή BreastSubtype_1 δεν είναι στατιστικά σημαντική.

Για τον έλεγχο : $H_0 : \beta_{25}=0$ έναντι της εναλλακτικής $H_1 : \beta_{25} \neq 0$

Παρατηρούμε ότι η στατιστική συνάρτηση ισούται με 0,018 με $p\text{-value} = 0,894$, συνεπώς σε επίπεδο σημαντικότητας 5% δεχόμαστε τη μηδενική υπόθεση. Αυτό σημαίνει ότι η μεταβλητή BreastSubtype_2 δεν είναι στατιστικά σημαντική.

Για τον έλεγχο : $H_0 : \beta_{26}=0$ έναντι της εναλλακτικής $H_1 : \beta_{26} \neq 0$
Παρατηρούμε ότι η στατιστική συνάρτηση ισούται με 32,707 με $p\text{-value} = 0,000$, συνεπώς σε επίπεδο σημαντικότητας 5% απορρίπτουμε τη μηδενική υπόθεση. Αυτό σημαίνει ότι η μεταβλητή BreastSubtype_3 είναι στατιστικά σημαντική. Το β_{26} ισούται 1,201.

Για τον έλεγχο : $H_0 : \beta_{27}=0$ έναντι της εναλλακτικής $H_1 : \beta_{27} \neq 0$
Παρατηρούμε ότι η στατιστική συνάρτηση ισούται με 0,839 με $p\text{-value} = 0,360$, συνεπώς σε επίπεδο σημαντικότητας 5% δεχόμαστε τη μηδενική υπόθεση. Αυτό σημαίνει ότι η μεταβλητή @6th_Stage δεν είναι στατιστικά σημαντική.

Για τον έλεγχο : $H_0 : \beta_{28}=0$ έναντι της εναλλακτικής $H_1 : \beta_{28} \neq 0$
Παρατηρούμε ότι η στατιστική συνάρτηση ισούται με 4,774 με $p\text{-value} = 0,029$, συνεπώς σε επίπεδο σημαντικότητας 5% απορρίπτουμε τη μηδενική υπόθεση. Αυτό σημαίνει ότι η μεταβλητή Stage_T1 είναι στατιστικά σημαντική. Το β_{28} ισούται 1,697.

Για τον έλεγχο : $H_0 : \beta_{29}=0$ έναντι της εναλλακτικής $H_1 : \beta_{29} \neq 0$
Παρατηρούμε ότι η στατιστική συνάρτηση ισούται με 13,052 με $p\text{-value} = 0,000$, συνεπώς σε επίπεδο σημαντικότητας 5% απορρίπτουμε τη μηδενική υπόθεση. Αυτό σημαίνει ότι η μεταβλητή Stage_T2 είναι στατιστικά σημαντική. Το β_{29} ισούται 1,431.

Για τον έλεγχο : $H_0 : \beta_{30}=0$ έναντι της εναλλακτικής $H_1 : \beta_{30} \neq 0$
Παρατηρούμε ότι η στατιστική συνάρτηση ισούται με 22,081 με $p\text{-value} = 0,000$, συνεπώς σε επίπεδο σημαντικότητας 5% απορρίπτουμε τη μηδενική υπόθεση. Αυτό σημαίνει ότι η μεταβλητή Stage_T3 είναι στατιστικά σημαντική. Το β_{30} ισούται 1,535.

Για τον έλεγχο : $H_0 : \beta_{31}=0$ έναντι της εναλλακτικής $H_1 : \beta_{31} \neq 0$
Παρατηρούμε ότι η στατιστική συνάρτηση ισούται με 11,034 με $p\text{-value} = 0,001$, συνεπώς σε επίπεδο σημαντικότητας 5% απορρίπτουμε τη μηδενική υπόθεση. Αυτό σημαίνει ότι η μεταβλητή Stage_T4 είναι στατιστικά σημαντική. Το β_{31} ισούται 1,000.

Για τον έλεγχο : $H_0 : \beta_{32}=0$ έναντι της εναλλακτικής $H_1 : \beta_{32} \neq 0$
Παρατηρούμε ότι η στατιστική συνάρτηση ισούται με 5,410 με $p\text{-value} = 0,020$, συνεπώς σε επίπεδο σημαντικότητας 5% απορρίπτουμε τη μηδενική υπόθεση. Αυτό σημαίνει ότι η μεταβλητή Stage_T5 είναι στατιστικά σημαντική. Το β_{32} ισούται 0,740.

Για τον έλεγχο : $H_0 : \beta_{33}=0$ έναντι της εναλλακτικής $H_1 : \beta_{33} \neq 0$
Παρατηρούμε ότι η στατιστική συνάρτηση ισούται με 2,200 με $p\text{-value} = 0,138$, συνεπώς σε επίπεδο σημαντικότητας 5% δεχόμαστε τη μηδενική υπόθεση. Αυτό σημαίνει ότι η μεταβλητή Stage_T6 δεν είναι στατιστικά σημαντική.

Για τον έλεγχο : $H_0 : \beta_{34}=0$ έναντι της εναλλακτικής $H_1 : \beta_{34} \neq 0$
Παρατηρούμε ότι η στατιστική συνάρτηση ισούται με 0,537 με $p\text{-value} = 0,464$, συνεπώς σε επίπεδο σημαντικότητας 5% δεχόμαστε τη μηδενική υπόθεση. Αυτό σημαίνει ότι η μεταβλητή Stage_T7 δεν είναι στατιστικά σημαντική.

Για τον έλεγχο : $H_0 : \beta_{35}=0$ έναντι της εναλλακτικής $H_1 : \beta_{35} \neq 0$

Παρατηρούμε ότι η στατιστική συνάρτηση ισούται με 0,498 με p-value = 0,481, συνεπώς σε επίπεδο σημαντικότητας 5% δεχόμαστε τη μηδενική υπόθεση. Αυτό σημαίνει ότι η μεταβλητή Stage_T8 δεν είναι στατιστικά σημαντική.

Για τον έλεγχο : $H_0 : \beta_{36}=0$ έναντι της εναλλακτικής $H_1 : \beta_{36} \neq 0$

Παρατηρούμε ότι η στατιστική συνάρτηση ισούται με 6,164 με p-value = 0,013, συνεπώς σε επίπεδο σημαντικότητας 5% απορρίπτουμε τη μηδενική υπόθεση. Αυτό σημαίνει ότι η μεταβλητή Stage_N1 είναι στατιστικά σημαντική. Το β_{36} ισούται 0,538.

Για τον έλεγχο : $H_0 : \beta_{37}=0$ έναντι της εναλλακτικής $H_1 : \beta_{37} \neq 0$

Παρατηρούμε ότι η στατιστική συνάρτηση ισούται με 20,496 με p-value = 0,000, συνεπώς σε επίπεδο σημαντικότητας 5% απορρίπτουμε τη μηδενική υπόθεση. Αυτό σημαίνει ότι η μεταβλητή Stage_N2 είναι στατιστικά σημαντική. Το β_{37} ισούται 0,668.

Για τον έλεγχο : $H_0 : \beta_{38}=0$ έναντι της εναλλακτικής $H_1 : \beta_{38} \neq 0$

Παρατηρούμε ότι η στατιστική συνάρτηση ισούται με 6,134 με p-value = 0,013, συνεπώς σε επίπεδο σημαντικότητας 5% απορρίπτουμε τη μηδενική υπόθεση. Αυτό σημαίνει ότι η μεταβλητή Stage_N3 είναι στατιστικά σημαντική. Το β_{38} ισούται 0,417.

Για τον έλεγχο : $H_0 : \beta_{39}=0$ έναντι της εναλλακτικής $H_1 : \beta_{39} \neq 0$

Παρατηρούμε ότι η στατιστική συνάρτηση ισούται με 1,262 με p-value = 0,261, συνεπώς σε επίπεδο σημαντικότητας 5% δεχόμαστε τη μηδενική υπόθεση. Αυτό σημαίνει ότι η μεταβλητή Tumor_Size δεν είναι στατιστικά σημαντική.

Για τον έλεγχο : $H_0 : \beta_{40}=0$ έναντι της εναλλακτικής $H_1 : \beta_{40} \neq 0$

Παρατηρούμε ότι η στατιστική συνάρτηση ισούται με 14,865 με p-value = 0,000, συνεπώς σε επίπεδο σημαντικότητας 5% απορρίπτουμε τη μηδενική υπόθεση. Αυτό σημαίνει ότι η μεταβλητή Regional_Nodes_Examined είναι στατιστικά σημαντική. Το β_{40} ισούται 0,012.

Για τον έλεγχο : $H_0 : \beta_{41}=0$ έναντι της εναλλακτικής $H_1 : \beta_{41} \neq 0$

Παρατηρούμε ότι η στατιστική συνάρτηση ισούται με 0,624 με p-value = 0,430, συνεπώς σε επίπεδο σημαντικότητας 5% δεχόμαστε τη μηδενική υπόθεση. Αυτό σημαίνει ότι η μεταβλητή Regional_Nodes_Positive δεν είναι στατιστικά σημαντική.

Το μοντέλο που περιγράφει τα παραπάνω είναι :

$$\begin{aligned} h(t/Z) = h_0(t) \exp(& 1,847*Age_2 + 1,897*Age_3 + 2,550*Age_4 + 2,364*Age_5 + \\ & 2,036*Age_6 + 1,975*Age_7 + 1,993*Age_8 + 1,651*Age_9 + 1,858*Age_10 \\ & + 1,387*Age_11 + 0,571*Age_12 + 0,672*Age_13 + 0,480*Grade_1 + 0,592*Grade_2 \\ & + 0,972*ER_Status + 0,447*PR_Status + 1,201*BreastSubtype_3 + 1,697*Stage_T1 \\ & + 1,431*Stage_T2 + 1,535*Stage_T3 + 1,000*Stage_T4 + 0,740*Stage_T5 \\ & + 0,538*Stage_N1 + 0,66*Stage_N2 + 0,417*Stage_N3 + \\ & 0,012*Regional_Nodes_Examined) , \text{ για κάθε } t \geq 0 \end{aligned}$$

Από το παραπάνω μοντέλο συμπεραίνουμε ότι:

Όταν μία γυναίκα ανήκει στην ηλικιακή ομάδα 25-29, η συνάρτηση κινδύνου πολλαπλασιάζεται με τον εκθετικό συντελεστή $e^{1,847}=6,340$ σε σχέση με την ηλικιακή ομάδα 85+.

Όταν μία γυναίκα ανήκει στην ηλικιακή ομάδα 30-34, η συνάρτηση κινδύνου πολλαπλασιάζεται με τον εκθετικό συντελεστή $e^{1,897}=6,663$ σε σχέση με την ηλικιακή ομάδα 85+.

Όταν μία γυναίκα ανήκει στην ηλικιακή ομάδα 35-39, η συνάρτηση κινδύνου πολλαπλασιάζεται με τον εκθετικό συντελεστή $e^{2,550}=12,813$ σε σχέση με την ηλικιακή ομάδα 85+.

Όταν μία γυναίκα ανήκει στην ηλικιακή ομάδα 40-44, η συνάρτηση κινδύνου πολλαπλασιάζεται με τον εκθετικό συντελεστή $e^{2,364}=10,635$ σε σχέση με την ηλικιακή ομάδα 85+.

Όταν μία γυναίκα ανήκει στην ηλικιακή ομάδα 45-49, η συνάρτηση κινδύνου πολλαπλασιάζεται με τον εκθετικό συντελεστή $e^{2,036}=7,657$ σε σχέση με την ηλικιακή ομάδα 85+.

Όταν μία γυναίκα ανήκει στην ηλικιακή ομάδα 50-54, η συνάρτηση κινδύνου πολλαπλασιάζεται με τον εκθετικό συντελεστή $e^{1,975}=7,207$ σε σχέση με την ηλικιακή ομάδα 85+.

Όταν μία γυναίκα ανήκει στην ηλικιακή ομάδα 55-59, η συνάρτηση κινδύνου πολλαπλασιάζεται με τον εκθετικό συντελεστή $e^{1,993}=7,340$ σε σχέση με την ηλικιακή ομάδα 85+.

Όταν μία γυναίκα ανήκει στην ηλικιακή ομάδα 60-64, η συνάρτηση κινδύνου πολλαπλασιάζεται με τον εκθετικό συντελεστή $e^{1,651}=5,213$ σε σχέση με την ηλικιακή ομάδα 85+.

Όταν μία γυναίκα ανήκει στην ηλικιακή ομάδα 65-69, η συνάρτηση κινδύνου πολλαπλασιάζεται με τον εκθετικό συντελεστή $e^{1,858}=6,410$ σε σχέση με την ηλικιακή ομάδα 85+.

Όταν μία γυναίκα ανήκει στην ηλικιακή ομάδα 70-74, η συνάρτηση κινδύνου πολλαπλασιάζεται με τον εκθετικό συντελεστή $e^{1,387}=4,005$ σε σχέση με την ηλικιακή ομάδα 85+.

Όταν μία γυναίκα ανήκει στην ηλικιακή ομάδα 75-79, η συνάρτηση κινδύνου πολλαπλασιάζεται με τον εκθετικό συντελεστή $e^{0,571}=1,771$ σε σχέση με την ηλικιακή ομάδα 85+.

Όταν μία γυναίκα ανήκει στην ηλικιακή ομάδα 80-84, η συνάρτηση κινδύνου πολλαπλασιάζεται με τον εκθετικό συντελεστή $e^{0,672}=1,958$ σε σχέση με την ηλικιακή ομάδα 85+.

Όταν τα καρκινικά κύτταρα είναι βαθμού 1, δηλαδή μοιάζουν λίγο διαφορετικά από τα φυσιολογικά κύτταρα, η συνάρτηση κινδύνου πολλαπλασιάζεται με τον εκθετικό συντελεστή $e^{0,480}=1,616$ σε σχέση με το βαθμό III.

Όταν τα καρκινικά κύτταρα είναι βαθμού 2, δηλαδή δεν μοιάζουν με τα φυσιολογικά κύτταρα, η συνάρτηση κινδύνου πολλαπλασιάζεται με τον εκθετικό συντελεστή $e^{0,592}=1,808$ σε σχέση με το βαθμό III.

Όταν οι υποδοχείς οιστρογόνου είναι θετικοί, η συνάρτηση κινδύνου πολλαπλασιάζεται με τον εκθετικό συντελεστή $e^{0,972}=2,644$.

Όταν οι υποδοχείς προγεστερόνης είναι θετικοί, η συνάρτηση κινδύνου πολλαπλασιάζεται με τον εκθετικό συντελεστή $e^{0,447}=1,564$.

Όταν ο δευτερεύον τύπος μαστού είναι τριπλός αρνητικός, η συνάρτηση κινδύνου πολλαπλασιάζεται με τον εκθετικό συντελεστή $e^{1,201}=3,322$ σε σχέση με το HER2 - enriched.

Όταν το Στάδιο T είναι T1a, η συνάρτηση κινδύνου πολλαπλασιάζεται με τον εκθετικό συντελεστή $e^{1,697}=5,458$ σε σχέση με το T4d.

Όταν το Στάδιο T είναι T1b, η συνάρτηση κινδύνου πολλαπλασιάζεται με τον εκθετικό συντελεστή $e^{1,431}=4,184$ σε σχέση με το T4d.

Όταν το Στάδιο T είναι T1c, η συνάρτηση κινδύνου πολλαπλασιάζεται με τον εκθετικό συντελεστή $e^{1,535}=4,641$ σε σχέση με το T4d.

Όταν το Στάδιο T είναι T2, η συνάρτηση κινδύνου πολλαπλασιάζεται με τον εκθετικό συντελεστή $e^{1,000}=2,719$ σε σχέση με το T4d.

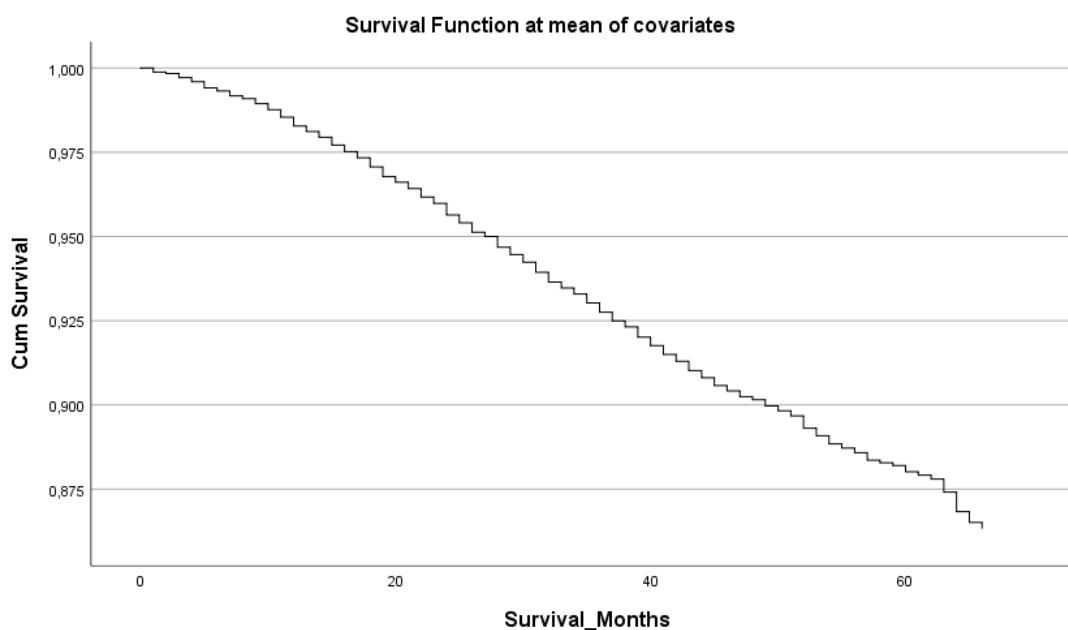
Όταν το Στάδιο T είναι T3, η συνάρτηση κινδύνου πολλαπλασιάζεται με τον εκθετικό συντελεστή $e^{0,740}=2,097$ σε σχέση με το T4d.

Όταν το Στάδιο N είναι N0, η συνάρτηση κινδύνου πολλαπλασιάζεται με τον εκθετικό συντελεστή $e^{0,538}=1,713$ σε σχέση με το N3.

Όταν το Στάδιο N είναι N1, η συνάρτηση κινδύνου πολλαπλασιάζεται με τον εκθετικό συντελεστή $e^{0,668}=1,951$ σε σχέση με το N3.

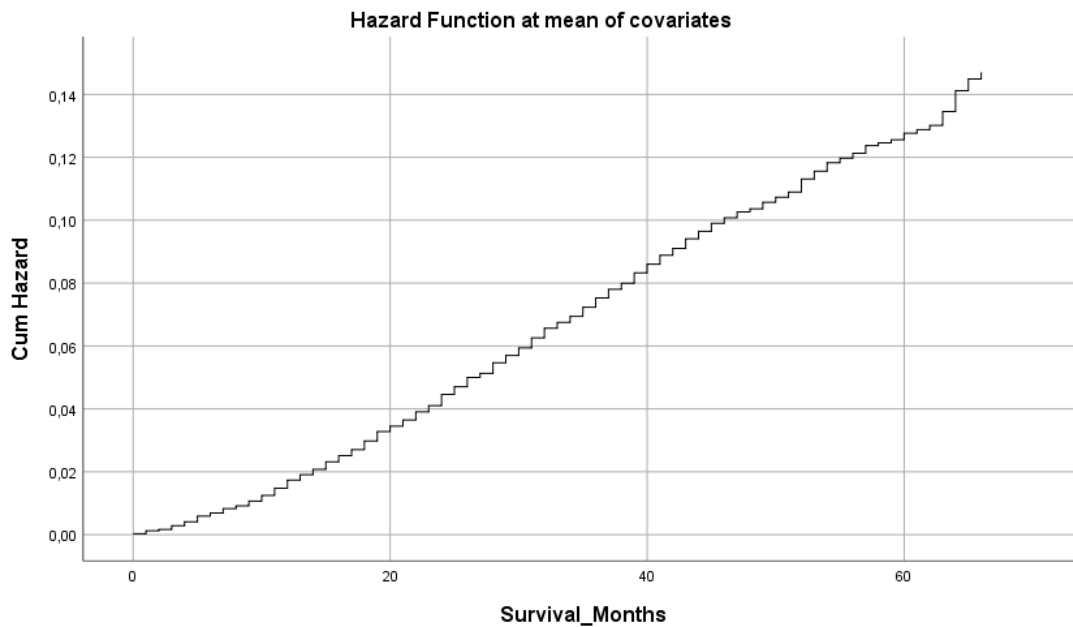
Όταν το Στάδιο N είναι N2, η συνάρτηση κινδύνου πολλαπλασιάζεται με τον εκθετικό συντελεστή $e^{0,417}=1,518$ σε σχέση με το N3.

Για αύξηση μίας μονάδας των περιφερειακών όγκων που εξετάστηκαν, η συνάρτηση κινδύνου πολλαπλασιάζεται με τον εκθετικό συντελεστή $e^{0,012}=1,012$.



Διάγραμμα 6.1 : Συνάρτηση Επιβίωσης.

Από το παραπάνω διάγραμμα παρατηρούμε ότι η συνάρτηση επιβίωσης είναι φθίνουσα συνάρτηση. Αυτό είναι λογικό, καθώς όσο αυξάνουν οι μήνες που κάποιος παραμένει ζωντανός, η πιθανότητα επιβίωσης μειώνεται. Συγκεκριμένα, καθώς ξεκινάει οι μήνες ζωής ενός ανθρώπου, η πιθανότητα επιβίωσης από τον καρκίνο του μαστού ισούται με τη μονάδα. Αντίστοιχα, ύστερα από περίπου πέντε χρόνια ζωής, η πιθανότητα επιβίωσης από τον καρκίνο του μαστού πέφτει κάτω από το 0,875.

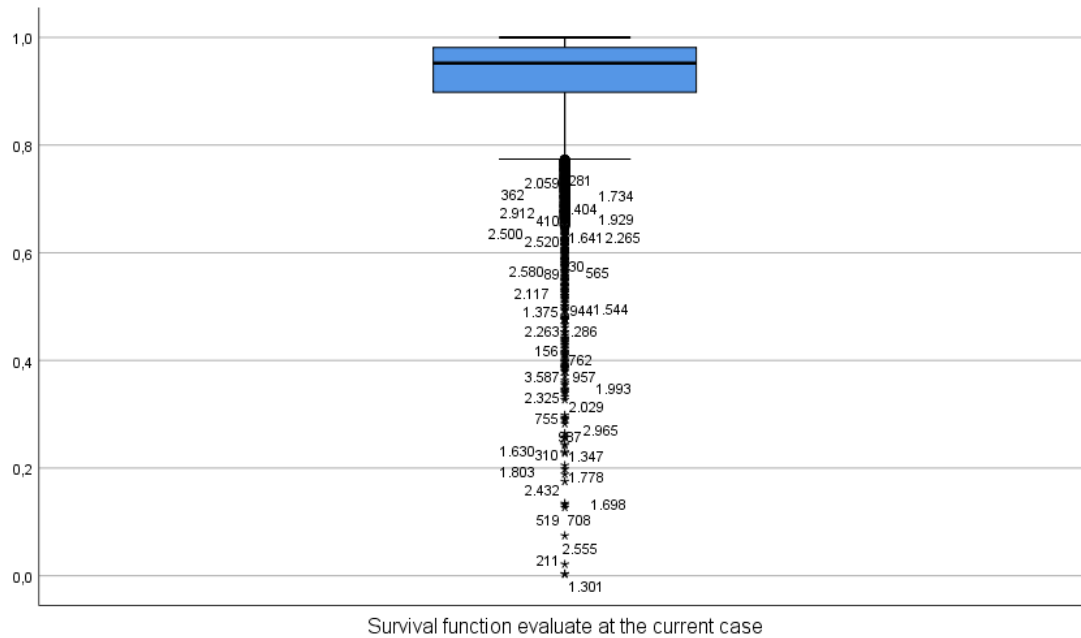


Διάγραμμα 6.2 : Συνάρτηση Κινδύνου.

Από το παραπάνω διάγραμμα παρατηρούμε ότι η συνάρτηση κινδύνου είναι αύξουσα συνάρτηση. Αυτό είναι λογικό, καθώς όσο αυξάνουν οι μήνες που κάποιος παραμένει ζωντανός, η πιθανότητα κινδύνου αυξάνεται. Συγκεκριμένα, καθώς ξεκινάει οι μήνες ζωής ενός ανθρώπου, η πιθανότητα κινδύνου από τον καρκίνο του μαστού ισούται με το μηδέν. Αντίστοιχα, ύστερα από περίπου πέντε χρόνια ζωής, η πιθανότητα κινδύνου από τον καρκίνο του μαστού ανεβαίνει περίπου στο 0,15.

- Περιγραφική Στατιστική ανάλυση για την συνάρτηση επιβίωσης και την συνάρτηση κινδύνου

- Συνάρτηση Επιβίωσης



Διάγραμμα 6.3 : Θηκόγραμμα Συνάρτηση Επιβίωσης.

Descriptives

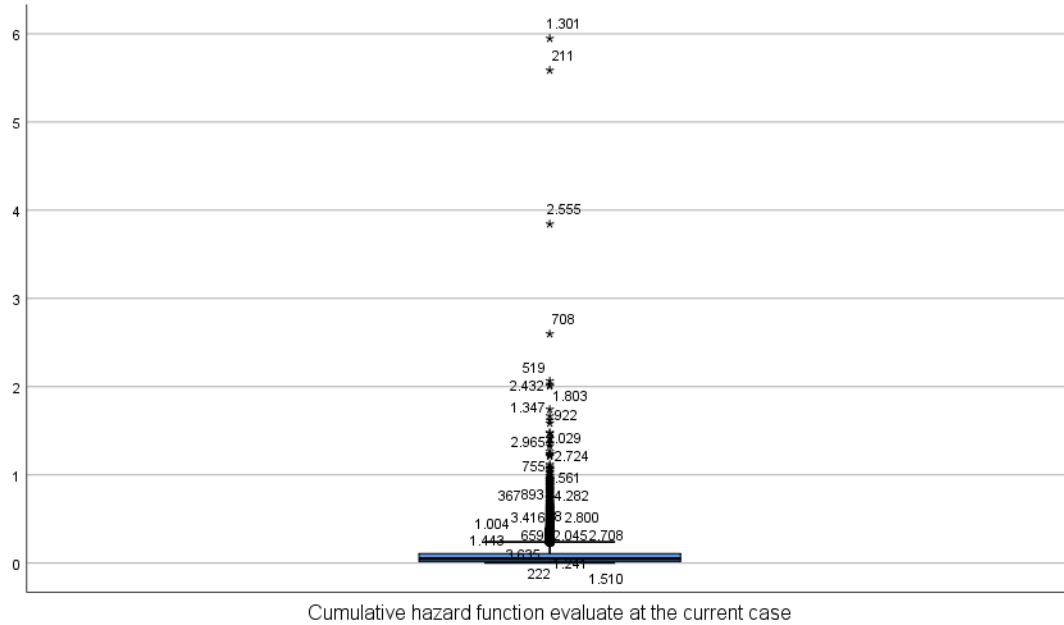
		Statistic	Std. Error	
Survival function evaluate at the current case	Mean	,9153557	,00148075	
	95% Confidence Interval for Mean	Lower Bound	,9124528	
		Upper Bound	,9182585	
	5% Trimmed Mean	,9312893		
	Median	,9521119		
	Variance	,012		
	Std. Deviation	,11033302		
	Minimum	,00261		
	Maximum	,99994		
	Range	,99733		
	Interquartile Range	,08305		
	Skewness	-3,025	,033	
	Kurtosis	12,424	,066	

Πίνακας 6.4 : Περιγραφικά μέτρα μεταβλητών.

Από το θηκόγραμμα και τον πίνακα περιγραφικών της συνάρτησης επιβίωσης παρατηρούμε ότι, το μεγαλύτερο ποσοστό των τιμών της είναι συγκεντρωμένο ανάμεσα στις τιμές 0.80 με 1.00 και συγκεκριμένα πιο κοντά στο 1.00. Επίσης, παρατηρούμε πως η διάμεσος έχει τιμή ανάμεσα στις τιμές 0.80 με 1.00 και

συγκεκριμένα είναι 0.95. Αυτό σημαίνει πως όσες παρατηρήσεις βρίσκονται πάνω από την τιμή 0.95 άλλες τόσες βρίσκονται κάτω από αυτήν. Τέλος, παρατηρούμε πως έχει κάποιες ακραίες τιμές οι οποίες παίρνουν χαμηλότερες τιμές.

ο **Συνάρτηση Κινδύνου**



Διάγραμμα 6.4 : Θηκόγραμμα Συνάρτηση Κινδύνου.

Descriptives

		Statistic	Std. Error	
Cumulative hazard function evaluate at the current case	Mean	,1008441	,00268037	
	95% Confidence Interval for Mean	Lower Bound	,0955895	
		Upper Bound	,1060987	
	5% Trimmed Mean	,0737340		
	Median	,0490727		
	Variance	,040		
	Std. Deviation	,19971933		
	Minimum	,00006		
	Maximum	5,94750		
	Range	5,94744		
	Interquartile Range	,08845		
	Skewness	12,079	,033	
	Kurtosis	272,711	,066	

Πίνακας 6.5 : Περιγραφικά μέτρα μεταβλητών.

Από το θηκόγραμμα και τον πίνακα περιγραφικών μέτρων της συνάρτησης κινδύνου παρατηρούμε ότι, το μεγαλύτερο ποσοστό των τιμών της είναι συγκεντρωμένο κοντά στο 0.00. Επίσης, παρατηρούμε πως η διάμεσος έχει τιμή κοντά στο 0.00 και συγκεκριμένα είναι 0.05. Αυτό σημαίνει πως όσες παρατηρήσεις

βρίσκονται πάνω από την τιμή 0.05 άλλες τόσες βρίσκονται κάτω από αυτήν. Τέλος, παρατηρούμε πως έχει κάποιες ακραίες τιμές οι οποίες παίρνουν υψηλότερες τιμές.

- Συμπερασματικά, από την παραπάνω ανάλυση, παρατηρούμε ότι το μεγαλύτερο ποσοστό των τιμών της συνάρτησης επιβίωσης είναι κοντά στην μονάδα ενώ το μεγαλύτερο ποσοστό των τιμών της συνάρτησης κινδύνου είναι κοντά στο μηδέν. Αυτό σημαίνει ότι η πιθανότητα επιβίωσης είναι μεγαλύτερη από την πιθανότητα αποβίωσης για το μεγαλύτερο μέρος του δείγματος μας. Δηλαδή, ένα άτομο που πάσχει από καρκίνο του μαστού, είναι πιο πιθανό να επιβιώσει από ότι να αποβιώσει.

ΚΕΦΑΛΑΙΟ 7

ΣΥΜΠΕΡΑΣΜΑΤΑ

7.1 Εισαγωγή

Στο κεφάλαιο αυτό θα παρουσιάσουμε τα αποτελέσματα και τα συμπεράσματα που προκύπτουν από την ανάλυση που προηγήθηκε στα παραπάνω κεφάλαια, για την επιβίωση των ασθενών με καρκίνο του μαστού με βάση κάποια κλινικά, διαγνωστικά και δημογραφικά χαρακτηριστικά.

7.2 Συμπεράσματα από την περιγραφική ανάλυση των μεταβλητών

- Σύμφωνα με την περιγραφική ανάλυση των ποσοτικών μεταβλητών, προκύπτει πως ο μέσος χρόνος ζωής ενός ατόμου που πάσχει από καρκίνο του μαστού είναι 32,72 μήνες, δηλαδή 2.67 χρόνια. Επίσης, το μέσο μέγεθος όγκου ενός πάσχοντα είναι 30,59 mm. Τέλος, το μέσο πλήθος των περιφερειακών όγκων που εξετάστηκαν σε ένα άτομο είναι 12,80 ενώ βγήκαν θετικοί κατά μέσο όρο 9,25 περιφερειακοί όγκοι σε ένα άτομο.
- Σύμφωνα με την περιγραφική ανάλυση των ποιοτικών μεταβλητών, προκύπτει πως το 0,1% του δείγματος ανήκει στην ηλικιακή ομάδα 20 με 24, το 0,8% του δείγματος ανήκει στην ηλικιακή ομάδα 25-29, το 2,3% του δείγματος ανήκει στην ηλικιακή ομάδα 30-34, το 4,3% του δείγματος ανήκει στην ηλικιακή ομάδα 35-39, το 8,4% του δείγματος ανήκει στην ηλικιακή ομάδα 40-44, το 11,7% του δείγματος ανήκει στην ηλικιακή ομάδα 45-49, το 13,1% του δείγματος ανήκει στην ηλικιακή ομάδα 50-54, το 12% του δείγματος ανήκει στην ηλικιακή ομάδα 55-59, το 13,9% του δείγματος ανήκει στην ηλικιακή ομάδα 60-64, το 11,7% του δείγματος ανήκει στην ηλικιακή ομάδα 65-69, το 8,1% του δείγματος ανήκει στην ηλικιακή ομάδα 70-74, το 5,8% του δείγματος ανήκει στην ηλικιακή ομάδα 75-79, το 4% του δείγματος ανήκει στην ηλικιακή ομάδα 80-84 και 3,7% του δείγματος ανήκει στην ηλικιακή ομάδα 85+.

Το 4,1% του δείγματος ήταν χρώματος μαύρου, το 85,6% του δείγματος ήταν χρώματος λευκού, το 0,3% του δείγματος ήταν άγνωστης φυλής και το 10,1% του δείγματος ανήκαν σε άλλη φυλή (American Indian/AK Native, Asian/Pacific Islander).

Το 12,3% των ατόμων του δείγματος ήταν διαζευγμένο, το 60,9% του δείγματος ήταν παντρεμένο, το 1,4% του δείγματος ήταν σε διάσταση, το 12,9% του δείγματος ήταν ανύπαντρο, το 1% του δείγματος ήταν με συγγάτοικο αλλά χωρίς να έχουν παντρευτεί και το 11,4% του δείγματος είχαν χηρέψει.

Το 13,1% του δείγματος ήταν είχαν καρκινικά κύτταρα βαθμού I, δηλαδή κύτταρα που μοιάζουν λίγο διαφορετικά από τα φυσιολογικά κύτταρα και αναπτύσσονται αργά, το 45% του δείγματος είχαν καρκινικά κύτταρα βαθμού

II, δηλαδή κύτταρα που δεν μοιάζουν με τα φυσιολογικά/κανονικά κύτταρα και αναπτύσσονται λίγο πιο γρήγορα από τα κανονικά και το 41,9% του δείγματος είχαν καρκινικά κύτταρα βαθμού III, δηλαδή κύτταρα που φαίνονται πολύ διαφορετικά από τα φυσιολογικά/κανονικά κύτταρα και αναπτύσσονται γρήγορα σε μη οργανωμένα σχήματα, με πολλά διαιρούμενα να δημιουργούν νέα καρκινικά κύτταρα.

Το 51,2% του δείγματος είχαν προέλευση πρωτογενούς όγκου στο αριστερό στήθος ενώ το 48,8% του δείγματος είχαν προέλευση πρωτογενούς όγκου στο δεξιό στήθος.

Το 4,8% του δείγματος είχαν έκταση της νόσου μακρινή ενώ το 95,2% του δείγματος είχαν έκταση της νόσου τοπική.

Το 16,3%, δηλαδή 904 από τις 5552 γυναίκες του δείγματος είχαν υποδοχείς οιστρογόνου αρνητικούς ενώ το 83,7% δηλαδή 4648 από τις 5552 είχαν υποδοχείς οιστρογόνου θετικούς.

Το 25% του δείγματος είχαν υποδοχείς προγεστερόνης αρνητικούς ενώ το 75% είχαν υποδοχείς προγεστερόνης θετικούς.

Το 73% του δείγματος είχαν δευτερεύον τύπου μαστό θετικό στους υποδοχείς ορμονών (οιστρογόνου ή προγεστερόνης) και αρνητικό στο Her2 (χρησιμοποιείται για την πρόβλεψη της ανταπόκρισης στην ορμονοθεραπεία και τη χημειοθεραπεία - παρακολούθηση της θεραπείας), το 5% του δείγματος είχαν δευτερεύον τύπου μαστό αρνητικό στους υποδοχείς ορμονών και θετικό στο Her2, το 12,1% του δείγματος είχαν δευτερεύον τύπου μαστό θετικό στους υποδοχείς ορμονών και θετικό στο Her2 και το 9,9% του δείγματος είχαν δευτερεύον τύπου μαστό αρνητικό στους υποδοχείς ορμονών και στο Her2.

Το 1,1% του δείγματος είχαν καρκίνο σε στάδιο 1 (αυτό το στάδιο είναι συνήθως ένας μικρός καρκίνος ή όγκος που δεν έχει αναπτυχθεί βαθιά σε κοντινούς ιστούς κι επίσης δεν έχει εξαπλωθεί στους λεμφαδένες ή σε άλλα μέρη του σώματος - ονομάζεται συχνά καρκίνος πρώιμου σταδίου), το 30,5% του δείγματος είχαν καρκίνο σε στάδιο 2A, το 33,7% του δείγματος είχαν καρκίνο σε στάδιο 2B, το 20,8% του δείγματος είχαν καρκίνο σε στάδιο 3A, το 5,2% του δείγματος είχαν καρκίνο σε στάδιο 3B και το 8,8% του δείγματος είχαν καρκίνο σε στάδιο 3C. (Γενικά, αυτά τα 2 στάδια -2^ο και 3^ο- υποδεικνύουν μεγαλύτερους καρκίνους ή όγκους που έχουν αναπτυχθεί βαθύτερα σε κοντινούς ιστούς. Μπορεί επίσης, να έχουν εξαπλωθεί σε λεμφαδένες αλλά όχι σε άλλα μέρη του σώματος.)

Το 0,9% του δείγματος είχαν καρκίνο σε στάδιο T1a, το 5,6% του δείγματος είχαν καρκίνο σε στάδιο T1b, το 26,4% του δείγματος είχαν καρκίνο σε στάδιο T1c, το 46,7% του δείγματος είχαν καρκίνο σε στάδιο T2, το 14,2% του δείγματος είχαν καρκίνο σε στάδιο T3 και το 0,9% του δείγματος είχαν καρκίνο σε στάδιο T4a. (Το γράμμα «T» συν έναν αριθμό από 0 έως 4 περιγράφει το μέγεθος και τη θέση του όγκου, συμπεριλαμβανομένου του κατά πόσο ο όγκος έχει αναπτυχθεί σε κοντινούς ιστούς. Το μέγεθος του όγκου μετριέται σε εκατοστά. Ένας μεγαλύτερος όγκος ή ένας όγκος που έχει αναπτυχθεί βαθύτερα στον περιβάλλοντα ιστό λαμβάνει μεγαλύτερο αριθμό.)

Το 5,4% του δείγματος είχαν καρκίνο σε στάδιο N0, το 71,8% του δείγματος είχαν καρκίνο σε στάδιο N1, το 14% του δείγματος είχαν καρκίνο σε στάδιο N2, το 8,8% του δείγματος είχαν καρκίνο σε στάδιο N3. (Το γράμμα «N» συν έναρ αριθμός από 0 έως 3 σημαίνει λεμφαδένες. Τις περισσότερες φορές, όσο είναι περισσότεροι οι λεμφαδένες με καρκίνο, τόσο μεγαλύτερος είναι ο αριθμός που αποδίδεται.)

Τέλος, το 90% του δείγματος δεν πέθαναν από καρκίνο του μαστού ενώ το 10% πέθαναν από αυτή τη νόσο.

7.3 Συμπεράσματα από τις συσχετίσεις των μεταβλητών

Σύμφωνα με την μελέτη μας, προκύπτει πως η ηλικιακή ομάδα στην οποία ανήκει ένα άτομο (*Age*), η οικογενειακή κατάσταση (*Marital_status*), η μέτρηση που δείχνει το κατά πόσο τα καρκινικά κύτταρα μοιάζουν με κανονικά (*Grade*), η έκταση της νόσου (*Stage_A*), το αν ο υποδοχέας οιστρογόνου είναι θετικός ή αρνητικός (*ER_Status*), το αν ο υποδοχέας προγεστερόνης είναι θετικός ή αρνητικός (*PR_Status*), ο δευτερεύον τύπος μαστού (*BreastSubtype*), το κλινικό ή παθολογικό στάδιο (*6th_Stage*), το μέγεθος και η θέση του όγκου (*Stage_T*), η μέτρηση που δείχνει τον αριθμό των λεμφαδένων ενός ατόμου που έχουν καρκίνο (*Stage_N*), το μέγεθος του όγκου (*Tumor_Size*), οι περιφερειακοί όγκοι που εξετάστηκαν (*Regional Nodes Examined*) καθώς και οι περιφερειακοί όγκοι που είναι θετικοί (*RegionalNodesPositive*) παίζουν ρόλο στο εάν κάποιος θα επιβιώσει ή όχι από τον καρκίνο του μαστού (**Status**).

Επίσης, προκύπτει πως το μέγεθος του όγκου (*Tumor_Size*), οι περιφερειακοί όγκοι που εξετάστηκαν (*Regional Nodes Examined*), η ηλικιακή ομάδα στην οποία ανήκει ένα άτομο (*Age*), η οικογενειακή κατάσταση (*Marital_status*), η μέτρηση που δείχνει το κατά πόσο τα καρκινικά κύτταρα μοιάζουν με κανονικά (*Grade*), η έκταση της νόσου (*Stage_A*), το αν ο υποδοχέας οιστρογόνου είναι θετικός ή αρνητικός (*ER_Status*), το αν ο υποδοχέας προγεστερόνης είναι θετικός ή αρνητικός (*PR_Status*), ο δευτερεύον τύπος μαστού (*BreastSubtype*), το κλινικό ή παθολογικό στάδιο (*6th_Stage*), το μέγεθος και η θέση του όγκου (*Stage_T*), η μέτρηση που δείχνει τον αριθμό των λεμφαδένων ενός ατόμου που έχουν καρκίνο (*Stage_N*) παίζουν ρόλο στους μήνες επιβίωσης ενός ατόμου που πάσχει από καρκίνο του μαστού (*Survival_months*).

7.4 Συμπεράσματα από την προσαρμογή των μοντέλων

Προσαρμογή του μοντέλου της Πολλαπλής γραμμικής παλινδρόμησης.

Εφαρμόζοντας το μοντέλο της Πολλαπλής γραμμικής παλινδρόμησης με την μέθοδο Stepwise, είδαμε πως η ηλικία ενός ατόμου Age_11, Age_12, Age_13, Age_14, η θέση του όγκου Stage T3, Stage_T7, Stage_T9, η οικογενειακή κατάσταση MaritalStatus_6, ο δευτερεύον τύπος μαστού BreastSubtype_4 και ο αριθμός των λεμφαδένων με καρκίνο Stage_N3 παίζουν ρόλο στην διάρκεια των μηνών που θα επιβιώσει κάποιος που πάσχει από καρκίνο του μαστού.

Στην συνέχεια, κάνοντας τους ελέγχους προϋποθέσεων για την προσαρμογή του μοντέλου, είδαμε πως τα σφάλματα δεν ακολουθούν κανονική κατανομή. Επομένως, καταλήγουμε στο συμπέρασμα πως η πολλαπλή γραμμική παλινδρόμηση δεν είναι κατάλληλη μέθοδος για την σωστή περιγραφή του δείγματος μας.

Προσαρμογή του μοντέλου της Λογιστικής παλινδρόμησης.

Εφαρμόζοντας το μοντέλο της Λογιστικής παλινδρόμησης, είδαμε πως η ηλικία (Age) εκτός της ηλικιακής ομάδας 20-24, η μέτρηση που δείχνει κατά πόσο τα καρκινικά κύτταρα μοιάζουν με κανονικά κύτταρα (Grade), ο δευτερεύον τύπος μαστού 3, ο υποδοχέας προγεστερόνης (PR Status), ο υποδοχέας οιστρογόνου (ER Status), το μέγεθος και η θέση του όγκου (Stage T2, T3), η μέτρηση που δείχνει τον αριθμό των λεμφαδένων ενός ατόμου που έχουν καρκίνο (Stage N), οι περιφερειακοί όγκοι που εξετάστηκαν (Regional Nodes Examined) καθώς και οι περιφερειακοί όγκοι που είναι θετικοί (Regional Nodes Positive) παίζουν ρόλο στην επιβίωση ή μη κάποιου που πάσχει από καρκίνο του μαστού.

Προσαρμογή του μοντέλου του μοντέλου του Cox.

Εφαρμόζοντας το μοντέλο του Cox, είδαμε πως οι περιφερειακοί όγκοι που εξετάστηκαν (Regional Nodes Examined), η ηλικία (Age) εκτός της ηλικιακής ομάδας 20-24, η μέτρηση που δείχνει κατά πόσο τα καρκινικά κύτταρα μοιάζουν με κανονικά κύτταρα (Grade), το αν ο υποδοχέας προγεστερόνης είναι θετικός ή αρνητικός (PR Status), το αν ο υποδοχέας οιστρογόνου είναι θετικός ή αρνητικός (ER Status), ο δευτερεύον τύπος μαστού 3, το μέγεθος και η θέση του όγκου (Stage T1, T2, T3, T4, T5) καθώς και πόσοι λεμφαδένες έχουν προσβληθεί από καρκίνο (Stage N1, N2, N3) παίζουν ρόλο στην επιβίωση ή μη των ατόμων που πάσχουν από καρκίνο του μαστού.

- Από την γραφική παράσταση της συνάρτησης επιβίωσης προκύπτει ότι όσο αυξάνουν οι μήνες που κάποιος παραμένει ζωντανός, η πιθανότητα επιβίωσης μειώνεται. Συγκεκριμένα, καθώς ξεκινάνε οι μήνες ζωής ενός ανθρώπου, η πιθανότητα επιβίωσης από τον καρκίνο του μαστού ισούται με τη μονάδα. Αντίστοιχα, ύστερα από περίπου πέντε χρόνια ζωής, η πιθανότητα επιβίωσης από τον καρκίνο του μαστού πέφτει κάτω από το 0,86.

- Από την γραφική παράσταση της συνάρτησης κινδύνου προκύπτει ότι όσο αυξάνουν οι μήνες που κάποιος παραμένει ζωντανός, η πιθανότητα κινδύνου αυξάνεται. Συγκεκριμένα, καθώς ξεκινάνε οι μήνες ζωής ενός ανθρώπου, η πιθανότητα κινδύνου από τον καρκίνο του μαστού ισούται με το μηδέν. Αντίστοιχα, ύστερα από περίπου πέντε χρόνια ζωής, η πιθανότητα κινδύνου από τον καρκίνο του μαστού ανεβαίνει περίπου στο 0,15.

Τέλος, από την περιγραφική ανάλυση των τιμών των συναρτήσεων επιβίωσης και κινδύνου βγαίνει το συμπέρασμα ότι η πιθανότητα επιβίωσης ενός ατόμου που πάσχει από καρκίνο του μαστού είναι μεγαλύτερη από την πιθανότητα αποβίωσης για το μεγαλύτερο ποσοστό του δείγματος.

ΒΙΒΛΙΟΓΡΑΦΙΑ

Ελληνική

1. Athens Breast Clinic, *Διαγνωστική διερεύνηση του καρκίνου του μαστού*
2. mastology-senology.gr Ιατρείο μαστού, *Ορμόνες και Καρκίνος του μαστού*
3. NOVARTIS, *Καρκίνος του Μαστού*
4. Άλμα ζωής Πανελλήνιος Σύλλογος Γυναικών με Καρκίνο του Μαστού, *Στάδια του καρκίνου του μαστού*
5. Ανάλυση Παλινδρόμησης, Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης, Πειραιάς 2011. [Slide 1 \(unipi.gr\)](#)
6. ΒΙΚΙΠΑΙΔΕΙΑ, *Συντελεστής Συσχέτισης Spearman*
7. Όμιλος Ιατρικού Αθηνών, *Καρκίνος του Μαστού*
8. Σύγκριση Συχνοτήτων Κατηγοριών: Το στατιστικό κριτήριο χ^2 [ROUSSOS 01-02 001-047 me eisagogi:Layout 5 \(toposbooks.gr\)](#)
9. Ανθιμίδης Γ., *Ανατομία του μαστού – Καρκίνος του μαστού*
10. ΔΕΣΠΟΥΔΗΣ Θ. – ΦΥΤΙΛΗΣ Χ. (2009), *Ανάλυση Επιβίωσης*. Αριστοτέλειο Πανεπιστήμιο Τμήμα Πληροφορικής
11. ΕΜΒΑΛΩΤΗΣ Α. ΚΑΤΣΗΣ Α. ΣΙΔΕΡΙΔΗΣ Γ. (2006), *Στατιστική Μεθοδολογία Εκπαίδευσης Έρευνας* [Microsoft Word - 2-Emvalwtis-Statist. Metodologia.doc \(aegean.gr\)](#)
12. Καλατζής Ι. (2020), *Επαγωγική Στατιστική: Έλεγχος Στατιστικών Υποθέσεων*
13. Καλόγηρος Γ. Β., *Καρκίνος του Μαστού – Θεραπεία*
14. Κρομμύδας Α. Χ., *HY – SPSS Statistical Package for Social Sciences*,
file:///C:/Users/stofp/Downloads/20220925_181613.pdf
15. Μάρκο Ε. Μ. (2011), *Ανάλυση του χρόνου επιβίωσης ασθενών με καρκίνο του μαστού μετά από χειρουργική επέμβαση*. Διπλωματική Εργασία Πανεπιστήμιο Πειραιώς
16. Μπούτσικας Μ. (2004), *“Πολλαπλή Γραμμική Παλινδρόμηση (Multiple Linear Regression)”*. Τμήμα Στατιστικής & Ασφαλιστικής Επιστήμης, Πανεπιστήμιο Πειραιώς,

17. Πατρίκου Κ. Δ. (2019), *Μελέτη των συμπτωμάτων που βιώνουν οι ασθενείς με καρκίνο πνευμόνων τελικού σταδίου και η επίπτωσή τους στην ποιότητα ζωής τους*. Διπλωματική Εργασία Πανεπιστήμιο Πειραιώς
18. Ράντης Ι. (2021), *Πρόβλεψη πιθανότητας χρεοκοπίας: Σύγκριση του μοντέλου του Altman και της λογιστικής παλινδρόμησης*, Διπλωματική Εργασία Πανεπιστήμιο Πειραιώς
19. Σπινέλη Α. Δ. (2010), *Ανάλυση γονιδιακών εκφράσεων για τη μελέτη του χρόνου επιβίωσης γυναικών με καρκίνο του μαστού*. Διπλωματική Εργασία Πανεπιστήμιο Πειραιώς
20. Φύσσας Γ. Π., *Καρκίνος του Μαστού: Αιτιολογία – Κληρονομικότητα – Παράγοντες Κινδύνου, Είδη καρκίνου του μαστού, Συμπτώματα Καρκίνου του Μαστού*

Ξένη

1. SEER Data and Software, SEER Data & Software (cancer.gov)
