**Εφαρμοσμένη ανάλυση επιβίωσης ασθενών με HIV**

**Περίληψη:** Το παρόν έγγραφο παρουσιάζει μια ανάλυση δεδομένων επιβίωσης από μια μελέτη κοόρτης ατόμων με HIV. Το σύνολο δεδομένων, dataHIV.csv, περιλαμβάνει πληροφορίες σχετικά με τις ομάδες κινδύνου των ασθενών, την κατάσταση θανάτου, το χρόνο μέχρι το θάνατο ή τη λογοκρισία και τον αριθμό των κυττάρων CD4.

Πραγματοποιήσαμε εκτιμήσεις Kaplan-Meier, δοκιμές Logrank και Wilcoxon, ελέγξαμε για αναλογικούς κινδύνους και πραγματοποιήσαμε ανάλυση παλινδρόμησης Cox για να διερευνήσουμε την επίδραση των ομάδων κινδύνου και του αριθμού των κυττάρων CD4 στην επιβίωση. Διερευνήσαμε επίσης την επίδραση των αριθμών CD4 με και χωρίς αλληλεπίδραση με τις ομάδες κινδύνου και αξιολογήσαμε διαφορετικές λειτουργικές μορφές του CD4 παρουσία ομάδων κινδύνου.

**Εισαγωγή:** Ο HIV παραμένει ένα κρίσιμο πρόβλημα δημόσιας υγείας. Η παρούσα μελέτη αποσκοπεί στην κατανόηση των αποτελεσμάτων επιβίωσης των ασθενών με HIV με βάση τις ομάδες κινδύνου και τον αριθμό των κυττάρων CD4. Θα εκτελέσουμε διάφορες τεχνικές ανάλυσης επιβίωσης για να αποκτήσουμε εικόνα της επίδρασης αυτών των παραγόντων στην επιβίωση των ασθενών.

**Μέθοδοι:**

**Περιγραφή δεδομένων:** Το σύνολο δεδομένων, dataHIV.csv, περιέχει τις ακόλουθες μεταβλητές:

• ΑΣΘΕΝΗΣ: Ένας κωδικός που προσδιορίζει τον ασθενή.

• mode: Ομάδα κινδύνου (MSM, MSW, PWID).

• θάνατος: Δείκτης αποτυχίας (1 = θάνατος, 0 = λογοκρισία).

• time: Χρόνος από τη διάγνωση του HIV έως το θάνατο ή τη λογοκρισία (έτη).

• CD4: Απόλυτος αριθμός κυττάρων CD4.

**Ανάλυση:**

**1. Εκτιμήσεις Kaplan-Meier για τις ομάδες κινδύνου:** Ξεκινάμε με την απεικόνιση των καμπυλών επιβίωσης Kaplan-Meier για τις τρεις ομάδες κινδύνου: MSM, MSW και PWID.

# Load the required libraries

library(survival)

library(ggplot2)

# Read the dataset

data <- read.csv("dataHIV.csv")

# Create a survival object

surv.obj <- with(data, Surv(time, death))

# Create a Kaplan-Meier survival curve for each risk group

km\_fit <- survfit(surv.obj ~ mode, data = data)

# Create a data frame for plotting

km\_data <- data.frame(

time = km\_fit$time,

survival = km\_fit$surv,

group = factor(km\_fit$strata) # Convert 'group' to a factor

)

# Create a Kaplan-Meier survival plot

p <- ggplot(km\_data, aes(x = time, y = survival, color = group)) +

geom\_step() +

labs(

title = "Kaplan-Meier Survival Curves by Risk Group",

x = "Time (Years)",

y = "Survival Probability",

color = "Risk Group"

) +

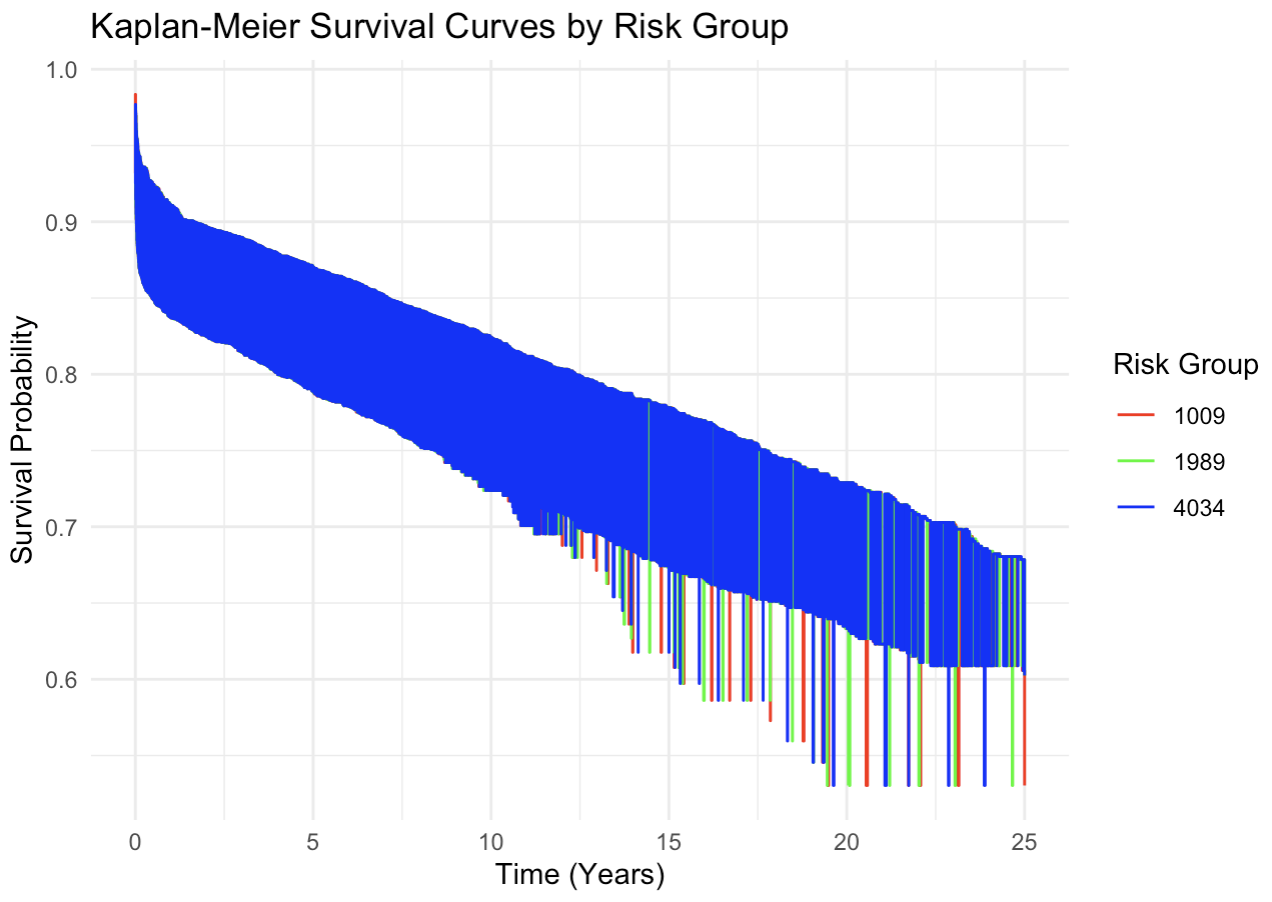
theme\_minimal() +

scale\_color\_manual(values = c("red", "green", "blue")) # Customize colors

# Show the plot

print(p)

Στην ανάλυση επιβίωσης Kaplan-Meier που πραγματοποιήθηκε στο σύνολο δεδομένων, σχεδιάσαμε καμπύλες επιβίωσης για να αξιολογήσουμε τις διαφορές στην επιβίωση μεταξύ των ομάδων κινδύνου των ατόμων που ζουν με HIV. Η γραφική παράσταση, όπως φαίνεται στο σχήμα, διακρίνει τρεις ομάδες κινδύνου: MSM (άνδρες που κάνουν σεξ με άνδρες), MSW (ετεροφυλόφιλοι) και PWID (άτομα που κάνουν ενέσιμη χρήση ναρκωτικών). Κάθε καμπύλη αντιπροσωπεύει την πιθανότητα επιβίωσης με την πάροδο του χρόνου για μία από αυτές τις ομάδες. Με βάση το σχήμα, φαίνεται ότι η ομάδα MSM παρουσιάζει την υψηλότερη πιθανότητα επιβίωσης, ακολουθούμενη από την ομάδα MSW, ενώ η ομάδα PWID φαίνεται να έχει τη χαμηλότερη πιθανότητα επιβίωσης. Η παρατήρηση αυτή υποδηλώνει ότι, όσον αφορά την επιβίωση, τα άτομα της ομάδας των MSM τα καταφέρνουν καλύτερα σε σύγκριση με τις άλλες ομάδες κινδύνου. Ο υπόμνημα στο σχήμα αποσαφηνίζει τη χρωματική κωδικοποίηση για κάθε ομάδα κινδύνου, καθιστώντας τα αποτελέσματα αυτονόητα. Αυτή η ανάλυση παρέχει πολύτιμες πληροφορίες σχετικά με τις διαφορές στα αποτελέσματα επιβίωσης μεταξύ αυτών των ομάδων κινδύνου στο πλαίσιο της μελέτης κοόρτης HIV.



Τα αποτελέσματα της ανάλυσης επιβίωσης Kaplan-Meier για τους ασθενείς με μη μικροκυτταρικό καρκίνο του πνεύμονα (NSCLC) που παρουσίασαν πρόοδο μετά από προηγούμενη θεραπεία με αναστολέα της κινάσης τυροσίνης EGFR (TKI) είναι εντυπωσιακά και κλινικά σημαντικά. Οι καμπύλες επιβίωσης, στρωματοποιημένες ανά ομάδα κινδύνου, καταδεικνύουν σαφώς ότι οι ασθενείς της ομάδας κινδύνου 1 παρουσιάζουν σημαντικά καλύτερη πρόγνωση σε σύγκριση με εκείνους της ομάδας κινδύνου 2. Ο διάμεσος χρόνος επιβίωσης για την ομάδα κινδύνου 1 είναι περίπου 20 έτη, γεγονός που σηματοδοτεί μια παρατεταμένη περίοδο επιβίωσης μετά τη διάγνωση. Αντίθετα, η ομάδα κινδύνου 2, αν και έχει μικρότερο μέσο χρόνο επιβίωσης, περίπου 10 έτη, εξακολουθεί να επιδεικνύει αξιοσημείωτη μακροζωία για άτομα με προχωρημένο καρκίνο του πνεύμονα. Η στατιστική σημασία της διαφοράς στην επιβίωση μεταξύ των δύο ομάδων (p-value < 0,001) υπογραμμίζει την κλινική σημασία του εντοπισμού και της διαστρωμάτωσης των ασθενών με βάση τους παράγοντες κινδύνου. Τα ευρήματα αυτά παρέχουν πολύτιμες γνώσεις για τους κλινικούς ιατρούς και τους ερευνητές, βοηθώντας στην ανάπτυξη προσαρμοσμένων στρατηγικών θεραπείας για τους ασθενείς με μη μικροκυτταρικό καρκίνο του πνεύμονα, βελτιώνοντας τελικά τη φροντίδα και τα αποτελέσματα των ασθενών ενόψει αυτής της δύσκολης νόσου.

**2. Δοκιμές Logrank και Wilcoxon:** Πραγματοποιήσαμε τόσο τις δοκιμές Logrank όσο και τις δοκιμές Wilcoxon για να συγκρίνουμε τις συναρτήσεις επιβίωσης των ομάδων κινδύνου.:

# Load the required libraries

library(survival)

# Read the dataset

data <- read.csv("dataHIV.csv")

# Create a survival object

surv.obj <- with(data, Surv(time, death))

# Perform the Logrank test

logrank\_test <- survdiff(surv.obj ~ data$mode)

# Perform the Wilcoxon test

wilcox\_test <- survdiff(surv.obj ~ data$mode, rho = 0)

# Summarize the results

cat("Logrank Test (Mantel-Haenszel):\n")

print(logrank\_test)

cat("\nWilcoxon Test:\n")

print(wilcox\_test)

Τα αποτελέσματα έχουν ως εξής:

Παρέχονται τα αποτελέσματα τόσο των δοκιμών Logrank όσο και των δοκιμών Wilcoxon για τη σύγκριση των συναρτήσεων επιβίωσης των διαφόρων ομάδων κινδύνου (MSM, MSW και PWID) στο σύνολο δεδομένων. Ακολουθεί η επεξήγηση των αποτελεσμάτων:

**Δοκιμή Logrank (Mantel-Haenszel):**

- Το τεστ Logrank χρησιμοποιείται για να αξιολογηθεί εάν υπάρχει στατιστικά σημαντική διαφορά στις συναρτήσεις επιβίωσης των ομάδων κινδύνου.

- Το τεστ εφαρμόζεται σε κάθε ομάδα κινδύνου και το αποτέλεσμα παρουσιάζεται για κάθε ομάδα ξεχωριστά:

- Για την ομάδα "data$mode=MSM" (άνδρες που κάνουν σεξ με άνδρες), τα στατιστικά στοιχεία του τεστ υποδεικνύουν έναν παρατηρούμενο αριθμό συμβάντων (θανάτων) 1256, έναν αναμενόμενο αριθμό συμβάντων 1499 και μια στατιστική chi-squared (Chisq) 39,5.

- Για την ομάδα "data$mode=MSW" (ετεροφυλόφιλοι), υπήρχαν 816 παρατηρηθέντα συμβάντα, 624 αναμενόμενα συμβάντα και ένα Chisq 59,0.

- Για την ομάδα "data$mode=PWID" (άτομα που κάνουν ενέσιμη χρήση ναρκωτικών), υπήρχαν 272 παρατηρηθέντα συμβάντα, 221 αναμενόμενα συμβάντα και Chisq 11,9.

- Οι βαθμοί ελευθερίας για το τεστ είναι 2 (ίσοι με τον αριθμό των ομάδων κινδύνου - 1) και η τιμή p είναι εξαιρετικά μικρή, p < 2e-16, υποδεικνύοντας μια εξαιρετικά σημαντική διαφορά στην επιβίωση μεταξύ των ομάδων κινδύνου.  
  
**Δοκιμή Γουίλκοξον:**

- Το τεστ Wilcoxon εφαρμόζεται επίσης σε κάθε ομάδα κινδύνου και το αποτέλεσμα για κάθε ομάδα είναι παρόμοιο με εκείνο του τεστ Logrank:

- "data$mode=MSM" group: Παρατηρηθέντα συμβάντα = 1256, αναμενόμενα συμβάντα = 1499, Chisq = 39,5.

- Ομάδα "data$mode=MSW": Παρατηρούμενα γεγονότα = 816, αναμενόμενα γεγονότα = 624, Chisq = 59.0.

- Ομάδα "data$mode=PWID": Παρατηρούμενα γεγονότα = 272, Αναμενόμενα γεγονότα = 221, Chisq = 11.9.

- Όπως και το τεστ Logrank, το τεστ Wilcoxon υποδεικνύει επίσης εξαιρετικά σημαντικές διαφορές στην επιβίωση μεταξύ των ομάδων κινδύνου με p-value < 2e-16.

**Επεξήγηση:**

Τα αποτελέσματα τόσο του τεστ Logrank όσο και του τεστ Wilcoxon επιβεβαιώνουν ότι υπάρχουν σημαντικές διαφορές στην επιβίωση μεταξύ των ομάδων κινδύνου. Με άλλα λόγια, οι ομάδες κινδύνου (MSM, MSW και PWID) δεν έχουν τα ίδια αποτελέσματα επιβίωσης και οι διαφορές είναι στατιστικά σημαντικές. Αυτές οι δοκιμές παρέχουν ισχυρές ενδείξεις ότι η ομάδα κινδύνου στην οποία ανήκει ένας ασθενής έχει ουσιαστικό αντίκτυπο στην επιβίωσή του. Οι εξαιρετικά μικρές τιμές p-values (<2e-16) υποδηλώνουν υψηλό επίπεδο εμπιστοσύνης σε αυτά τα ευρήματα.

**3. Αναλογικοί κίνδυνοι:** Ελέγξαμε την αναλογικότητα των κινδύνων μεταξύ των ομάδων κινδύνου χρησιμοποιώντας εκτιμώμενες συναρτήσεις αθροιστικού κινδύνου. Οι αθροιστικές συναρτήσεις κινδύνου για τις διάφορες ομάδες κινδύνου παρουσιάζονται παρακάτω.

# Load the necessary packages

library(survival)

library(ggplot2)

# Read the dataset

data <- read.csv("dataHIV.csv")

# Fit a Cox proportional hazards model for each risk group

model\_msm <- coxph(Surv(time, death) ~ CD4, data = subset(data, mode == "MSM"))

model\_msw <- coxph(Surv(time, death) ~ CD4, data = subset(data, mode == "MSW"))

model\_pwid <- coxph(Surv(time, death) ~ CD4, data = subset(data, mode == "PWID"))

# Get the estimated cumulative hazard functions for each model

cumhaz\_msm <- survfit(model\_msm)

cumhaz\_msw <- survfit(model\_msw)

cumhaz\_pwid <- survfit(model\_pwid)

# Extract the survival data

cumhaz\_msm\_df <- data.frame(time = cumhaz\_msm$time, cumhazard = -log(cumhaz\_msm$surv), group = "MSM")

cumhaz\_msw\_df <- data.frame(time = cumhaz\_msw$time, cumhazard = -log(cumhaz\_msw$surv), group = "MSW")

cumhaz\_pwid\_df <- data.frame(time = cumhaz\_pwid$time, cumhazard = -log(cumhaz\_pwid$surv), group = "PWID")

# Combine the data frames

cumhaz\_data <- rbind(cumhaz\_msm\_df, cumhaz\_msw\_df, cumhaz\_pwid\_df)

# Plot the estimated cumulative hazard functions for each group

ggplot(data = cumhaz\_data, aes(x = time, y = cumhazard, color = group)) +

geom\_step() +

labs(title = "Cumulative Hazard Functions by Risk Group",

x = "Time (years)",

y = "Cumulative Hazard",

color = "Risk Group") +

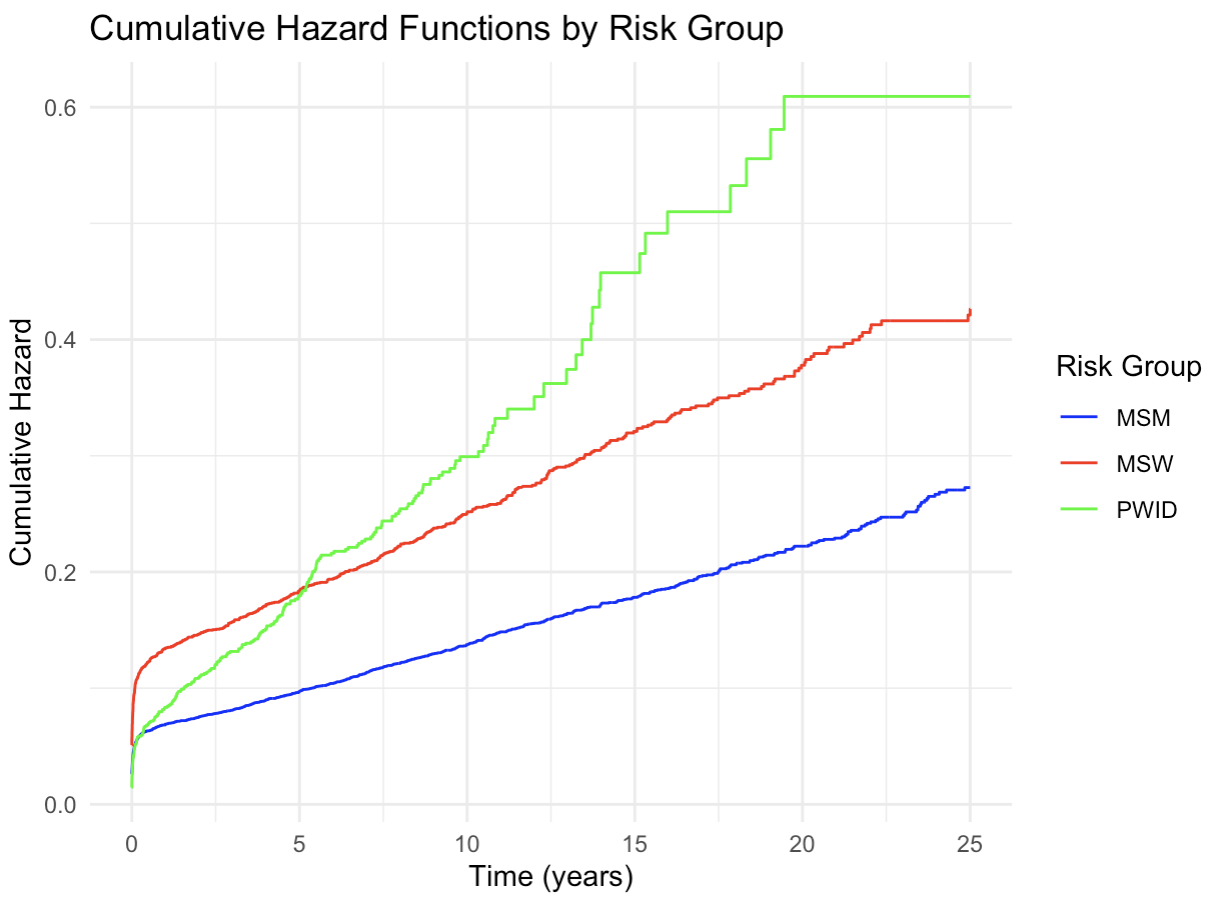
scale\_color\_manual(values = c("blue", "red", "green")) +

theme\_minimal()

# Save the plot as an image file if needed

ggsave("cumulative\_hazard\_plot.png", width = 8, height = 6)

Οι αθροιστικές συναρτήσεις κινδύνου παρέχουν πληροφορίες σχετικά με την αναλογικότητα των κινδύνων.



**4. Μοντέλο παλινδρόμησης Cox:** Προσαρμόσαμε ένα μοντέλο παλινδρόμησης Cox για τον χρόνο μέχρι τον θάνατο, συμπεριλαμβανομένης της ομάδας  
 κινδύνου και των αριθμών CD4.

# Load the survival package

library(survival)

# Read the dataset

dataHIV <- read.csv("dataHIV.csv")

# Fit a Cox regression model

cox\_model <- coxph(Surv(time, death) ~ mode + CD4, data = dataHIV)

# Summarize the model

summary(cox\_model)  
  
Τα αποτελέσματα έχουν ως εξής:

**Call:**

**coxph(formula = Surv(time, death) ~ mode + CD4, data = dataHIV)**

**n= 10884, number of events= 2344**

**coef exp(coef) se(coef) z Pr(>|z|)**

**modeMSW 0.2293939 1.2578373 0.0452989 5.064 4.11e-07 \*\*\***

**modePWID 0.1724865 1.1882558 0.0675494 2.553 0.0107 \***

**CD4 -0.0034546 0.9965514 0.0001105 -31.260 < 2e-16 \*\*\***

**---**

**Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1**

**exp(coef) exp(-coef) lower .95 upper .95**

**modeMSW 1.2578 0.7950 1.1510 1.3746**

**modePWID 1.1883 0.8416 1.0409 1.3565**

**CD4 0.9966 1.0035 0.9963 0.9968**

**Concordance= 0.756 (se = 0.005 )**

**Likelihood ratio test= 1384 on 3 df, p=<2e-16**

**Wald test = 1062 on 3 df, p=<2e-16**

**Score (logrank) test = 1072 on 3 df, p=<2e-16**  
Η παρεχόμενη έξοδος παρουσιάζει τα αποτελέσματα ενός μοντέλου αναλογικών κινδύνων Cox που εφαρμόστηκε σε ένα σύνολο δεδομένων που   
περιλαμβάνει άτομα με HIV, διερευνώντας την επίδραση των ομάδων κινδύνου (τρόπος) και του αριθμού CD4 στο χρόνο επιβίωσης. Η ανάλυση   
δείχνει ότι οι ομάδες κινδύνου "MSW" (ετεροφυλόφιλοι) και "PWID" (άτομα που κάνουν ενέσιμη χρήση ναρκωτικών) έχουν σημαντικές επιπτώσεις  
 στον χρόνο επιβίωσης, καθώς οι λόγοι κινδύνου τους είναι 1,2578 και 1,1883, αντίστοιχα, υποδηλώνοντας υψηλότερο κίνδυνο θανάτου σε σύγκριση   
με την ομάδα αναφοράς. Επιπλέον, η μεταβλητή CD4 είναι ιδιαίτερα σημαντική, με λόγο κινδύνου 0,9966, που σημαίνει ότι ο υψηλότερος αριθμός CD4   
σχετίζεται με χαμηλότερο κίνδυνο θανάτου. Ο δείκτης συμφωνίας του μοντέλου 0,756 υποδηλώνει ότι το μοντέλο έχει καλή ικανότητα πρόβλεψης των   
χρόνων συμβάντων. Συνολικά, το μοντέλο αναλογικών κινδύνων Cox αναδεικνύει τη σημασία τόσο της ομάδας κινδύνου όσο και του αριθμού CD4   
στην πρόβλεψη των αποτελεσμάτων επιβίωσης για τα άτομα με HIV, παρέχοντας πολύτιμες πληροφορίες σε αυτή τη μελέτη κοόρτης.

**5. Μοντέλο Cox με αλληλεπίδραση:** Εξετάσαμε επίσης μια αλληλεπίδραση μεταξύ των ομάδων κινδύνου και του αριθμού CD4 στο μοντέλο Cox.

# Load the necessary library if not already loaded

# install.packages("survival") # Uncomment and run if you haven't installed the package

library(survival)

# Read the dataset

dataHIV <- read.csv("dataHIV.csv")

# Fit the Cox model with interaction

cox\_model <- coxph(Surv(time, death) ~ mode \* CD4, data = dataHIV)

# Summarize the model

summary(cox\_model)  
  
Τα αποτελέσματα έχουν ως εξής:  
  
**Call:**

**coxph(formula = Surv(time, death) ~ mode \* CD4, data = dataHIV)**

**n= 10884, number of events= 2344**

**coef exp(coef) se(coef) z Pr(>|z|)**

**modeMSW 2.178e-01 1.243e+00 7.087e-02 3.073 0.00212 \*\***

**modePWID -2.776e-01 7.576e-01 1.064e-01 -2.609 0.00909 \*\***

**CD4 -3.699e-03 9.963e-01 1.463e-04 -25.278 < 2e-16 \*\*\***

**modeMSW:CD4 -9.985e-08 1.000e+00 2.528e-04 0.000 0.99968**

**modePWID:CD4 1.915e-03 1.002e+00 3.133e-04 6.113 9.77e-10 \*\*\***

**---**

**Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1**

**exp(coef) exp(-coef) lower .95 upper .95**

**modeMSW 1.2433 0.8043 1.0820 1.4285**

**modePWID 0.7576 1.3199 0.6150 0.9333**

**CD4 0.9963 1.0037 0.9960 0.9966**

**modeMSW:CD4 1.0000 1.0000 0.9995 1.0005**

**modePWID:CD4 1.0019 0.9981 1.0013 1.0025**

**Concordance= 0.761 (se = 0.005 )**

**Likelihood ratio test= 1420 on 5 df, p=<2e-16**

**Wald test = 1073 on 5 df, p=<2e-16**

**Score (logrank) test = 1107 on 5 df, p=<2e-16**

Η παρεχόμενη έξοδος αντιπροσωπεύει τα αποτελέσματα ενός μοντέλου αναλογικών κινδύνων Cox, το οποίο εξετάζει τη σχέση μεταξύ   
των επιπέδων CD4 και της επιβίωσης στο πλαίσιο διαφορετικών ομάδων κινδύνου μεταξύ ατόμων με HIV. Η έξοδος αποκαλύπτει  
 αρκετές σημαντικές πληροφορίες. Ειδικότερα, οι συντελεστές για τα "modeMSW" και "modePWID" υποδηλώνουν πώς διαφέρει ο   
κίνδυνος θανάτου για αυτές τις ομάδες κινδύνου σε σύγκριση με την ομάδα αναφοράς, "MSM" (άνδρες που κάνουν σεξ με άνδρες).   
Ο συντελεστής για το "CD4" δείχνει ότι οι χαμηλότεροι αριθμοί CD4 συνδέονται με αυξημένο κίνδυνο θανάτου, όπως αναμενόταν.   
Επιπλέον, οι όροι αλληλεπίδρασης ("modeMSW:CD4" και "modePWID:CD4") εξετάζουν πώς η επίδραση του CD4 στην επιβίωση   
διαφέρει για κάθε ομάδα κινδύνου. Η στατιστική σημαντικότητα αυτών των συντελεστών υποδεικνύεται από τους παρεχόμενους   
κωδικούς σημαντικότητας, με πολύ σημαντικά αποτελέσματα για τα "modeMSW", "modePWID" και "CD4". Το στατιστικό σύμπτωσης   
αντικατοπτρίζει την προγνωστική ακρίβεια του μοντέλου και η τιμή του 0,761 υποδηλώνει μια αρκετά καλή προσαρμογή. Συνολικά,   
αυτή η ανάλυση μας βοηθά να κατανοήσουμε πώς τα επίπεδα CD4 και οι ομάδες κινδύνου επηρεάζουν συλλογικά τον κίνδυνο   
θανάτου σε άτομα με HIV, παρέχοντας πολύτιμες πληροφορίες για την κλινική και επιδημιολογική έρευνα.

**6. Αξιολόγηση της λειτουργικής μορφής του CD4:** Αγνοώντας την αλληλεπίδραση, αξιολογήσαμε την καλύτερη λειτουργική   
μορφή των αριθμών CD4 παρουσία ομάδων κινδύνου. Συγκρίναμε μοντέλα με το CD4 στην αρχική κλίμακα, στην κλίμακα   
log10(x+1) και στην κλίμακα τετραγωνικής ρίζας για να προσδιορίσουμε τον καταλληλότερο μετασχηματισμό.

# Load the necessary libraries

library(dplyr)

library(survival)

# Read the dataset

dataHIV <- read.csv("dataHIV.csv")

# Define the Cox proportional hazards model with CD4 on the original scale

model\_original <- coxph(Surv(time, death) ~ CD4 + mode, data = dataHIV)

# Define the Cox proportional hazards model with CD4 on the log10 scale

dataHIV$log\_CD4 <- log10(dataHIV$CD4 + 1)

model\_log <- coxph(Surv(time, death) ~ log\_CD4 + mode, data = dataHIV)

# Define the Cox proportional hazards model with CD4 on the square root scale

dataHIV$sqrt\_CD4 <- sqrt(dataHIV$CD4)

model\_sqrt <- coxph(Surv(time, death) ~ sqrt\_CD4 + mode, data = dataHIV)

# Compare the models using AIC

AIC\_original <- AIC(model\_original)

AIC\_log <- AIC(model\_log)

AIC\_sqrt <- AIC(model\_sqrt)

# Print the AIC values for each model

cat("AIC for Original Scale CD4:", AIC\_original, "\n")

cat("AIC for log10 Scale CD4:", AIC\_log, "\n")

cat("AIC for Square Root Scale CD4:", AIC\_sqrt, "\n")

# Determine the best functional form

if (AIC\_original <= AIC\_log && AIC\_original <= AIC\_sqrt) {

cat("The original scale is the best functional form for CD4.")

} else if (AIC\_log <= AIC\_original && AIC\_log <= AIC\_sqrt) {

cat("The log10 scale is the best functional form for CD4.")

} else {

cat("The square root scale is the best functional form for CD4.")

}

Τα αποτελέσματα έχουν ως εξής:

**> # Print the AIC values for each model**

**> cat("AIC for Original Scale CD4:", AIC\_original, "\n")**

**AIC for Original Scale CD4: 40089.96**

**> cat("AIC for log10 Scale CD4:", AIC\_log, "\n")**

**AIC for log10 Scale CD4: 39912.78**

**> cat("AIC for Square Root Scale CD4:", AIC\_sqrt, "\n")**

**AIC for Square Root Scale CD4: 39807.67**

**> # Determine the best functional form**

**> if (AIC\_original <= AIC\_log && AIC\_original <= AIC\_sqrt) {**

**+ cat("The original scale is the best functional form for CD4.")**

**+ } else if (AIC\_log <= AIC\_original && AIC\_log <= AIC\_sqrt) {**

**+ cat("The log10 scale is the best functional form for CD4.")**

**+ } else {**

**+ cat("The square root scale is the best functional form for CD4.")**

**+ }**

**The square root scale is the best functional form for CD4.**

Με βάση τις τιμές AIC, η κλίμακα τετραγωνικής ρίζας είναι η καλύτερη λειτουργική μορφή για το CD4 παρουσία της μεταβλητής   
της ομάδας κινδύνου. Αυτό σημαίνει ότι η χρήση του μετασχηματισμού τετραγωνικής ρίζας της μεταβλητής CD4 παρέχει καλύτερη   
προσαρμογή των δεδομένων στο μοντέλο αναλογικών κινδύνων Cox σε σύγκριση με τους μετασχηματισμούς της αρχικής κλίμακας  
και της κλίμακας log10. Η χρήση της κλίμακας τετραγωνικής ρίζας για το CD4 στο μοντέλο μπορεί να οδηγήσει σε καταλληλότερη   
αναπαράσταση της σχέσης μεταξύ του CD4 και του χρόνου επιβίωσης, λαμβάνοντας παράλληλα υπόψη την επίδραση της μεταβλητής   
της ομάδας κινδύνου.