Kivimaki et al. **Progeronic cytokines as drivers of ageing phenotypes and dementia development**

**Statistical code**

R version 4.4.1 (2024-06-14)

**Cytokines and incident dementia**

# Load libraries  
library(dplyr)

library(readr)  
library(ggplot2)  
library(survival)  
library(writexl)  
library(tidyr)  
library(Matrix)

library(haven)  
library(ggrepel)  
  
# Read data files  
age\_data <- read.csv("Age.csv")  
covariates <- read.csv("Covariates\_p5.csv")  
panel\_data <- read.csv("Cytokine\_panelSL.csv")  
data <- read.csv("dementia.csv")  
data <- data[, c("ID\_OUT", setdiff(names(data), "ID\_OUT"))]  
V40\_data <- read.csv("V40\_anmlSMP.csv")  
V41\_data <- read.csv("V41\_anmlSMP.csv")  
  
# Rank-inverse normal transformation   
invnormal <- function(x) {  
 qnorm((rank(x, na.last = "keep") - 0.5) / sum(!is.na(x)))  
}  
  
V40\_data[-1] <- apply(V40\_data[-1], 2, invnormal)  
V41\_data[-1] <- apply(V41\_data[-1], 2, invnormal)  
V40\_data <- setNames(V40\_data, gsub("X", "", gsub("\\.", "-", names(V40\_data))))  
V41\_data <- setNames(V41\_data, gsub("X", "", gsub("\\.", "-", names(V41\_data))))  
  
# Filter proteins of interest and prepare dataset  
proteins\_of\_interest <- panel\_data$Protein\_ID   
V40\_filtered <- V40\_data[, c("ID\_OUT", intersect(names(V40\_data)[-1], proteins\_of\_interest))]  
V41\_filtered <- V41\_data[, c("ID\_OUT", intersect(names(V41\_data)[-1], proteins\_of\_interest))]  
  
common\_proteins <- intersect(names(V40\_filtered)[-1], names(V41\_filtered)[-1])  
V41\_filtered\_common <- V41\_filtered[, c("ID\_OUT", common\_proteins)]  
appended\_data <- bind\_rows(V40\_filtered, V41\_filtered\_common)  
unique\_appended\_data <- distinct(appended\_data, ID\_OUT, .keep\_all = TRUE)  
  
merged\_data <- merge(unique\_appended\_data, covariates, by = "ID\_OUT", all.x = TRUE)  
merged\_data <- merge(merged\_data, age\_data, by = "ID\_OUT", all.x = TRUE)  
merged\_data <- merged\_data %>%  
 distinct(ID\_OUT, .keep\_all = TRUE)  
data <- merge(data, merged\_data, by = "ID\_OUT", all.x = TRUE)  
  
data\_filtered <- data %>% select(ID\_OUT, all\_of(intersect(names(data), proteins\_of\_interest)), dementia, sex, age, ethn\_ds, futimePH5\_dem)  
data\_filtered <- data\_filtered %>%   
 filter(complete.cases(.)) %>%   
 filter(futimePH5\_dem >= 0)  
  
data\_filtered$sex <- factor(data\_filtered$sex, levels = c(0, 1), labels = c("male", "female"))  
data\_filtered$ethn\_ds <- factor(data\_filtered$ethn\_ds, levels = c(0, 1), labels = c("white", "non-white"))  
  
names(data\_filtered) <- make.names(names(data\_filtered), unique = TRUE)  
  
# Cox proportional hazards regression analysis   
results\_list <- list()  
for (protein in setdiff(names(data\_filtered)[4:length(names(data\_filtered)) - 2], c("dementia", "futimePH5\_dem", "sex", "age", "ethn\_ds"))) {  
 formula <- as.formula(paste("Surv(futimePH5\_dem, dementia) ~", protein, "+ sex + age + ethn\_ds"))  
   
 print(formula)  
   
 model <- tryCatch({  
 coxph(formula, data = data\_filtered)  
 }, error = function(e) {  
 print(paste("Error", protein))  
 print(e)  
 return(NULL)  
 })  
   
 if (!is.null(model)) {  
 coefs <- coef(summary(model))  
 ci\_lower <- exp(coefs[protein, "coef"] - 1.96 \* coefs[protein, "se(coef)"])  
 ci\_upper <- exp(coefs[protein, "coef"] + 1.96 \* coefs[protein, "se(coef)"])  
 hazard\_ratio <- exp(coefs[protein, "coef"])  
 p\_value <- coefs[protein, "Pr(>|z|)"]  
 results\_list[[protein]] <- c(hazard\_ratio, ci\_lower, ci\_upper, p\_value)  
 } else {  
 results\_list[[protein]] <- rep(NA, 4)  
 }  
}

## Surv(futimePH5\_dem, dementia) ~ X10044.12 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X10365.132 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X10367.62 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X10455.196 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X10851.77 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X11129.66 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X13093.6 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X13113.7 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X13236.25 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X13268.45 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X13676.46 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X13733.5 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X13738.8 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X14026.24 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X14061.48 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X14063.17 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X14093.10 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X14127.240 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X14129.1 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X14149.9 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X14150.7 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X14587.16 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X15346.31 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X15404.3 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X15405.23 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X15446.25 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X15503.20 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X15666.21 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X15667.39 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X15668.19 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X15686.49 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X15692.300 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X16746.12 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X16755.195 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X16756.30 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X16760.2 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X17356.34 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X18375.28 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X18389.11 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X18814.21 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X18878.15 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X19437.61 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X19568.17 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X19622.7 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X2333.72 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X2524.56 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X2597.8 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X2599.51 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X2607.54 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X2708.54 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X2714.78 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X2723.9 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X2752.62 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X2765.4 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X2773.50 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X2774.10 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X2778.10 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X2829.19 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X2906.55 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X2968.61 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X2970.60 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X2972.57 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X3010.53 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X3025.50 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X3035.80 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X3037.62 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X3052.8 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X3059.50 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X3067.67 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X3070.1 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X3321.2 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X3348.49 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X3421.54 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X3489.9 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X3497.13 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X3505.6 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X3506.49 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X3520.58 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X3534.14 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X3554.24 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X3587.53 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X3738.54 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X4136.40 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X4137.57 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X4138.25 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X4140.3 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X4156.74 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X4272.46 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X4374.45 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X4396.54 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X4397.26 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X4493.92 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X4556.10 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X4673.13 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X4697.59 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X4703.87 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X4717.55 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X4851.25 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X4867.15 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X4880.21 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X4889.82 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X4992.49 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X5011.11 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X5353.89 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X5355.69 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X5598.3 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X5609.92 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X5661.15 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X5713.9 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X5714.88 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X5807.77 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X5813.58 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X5834.18 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X5939.42 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X6210.100 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X6214.84 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X6230.56 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X6252.62 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X6421.52 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X6495.14 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X6517.14 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X6624.94 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X7124.18 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X7180.114 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X7196.21 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X7790.21 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X8059.1 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X8080.24 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X8221.19 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X8304.50 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X8467.9 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X8833.20 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X9051.13 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X9117.4 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X9177.6 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X9178.30 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X9188.119 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X9255.5 + sex + age + ethn\_ds  
## Surv(futimePH5\_dem, dementia) ~ X9377.25 + sex + age + ethn\_ds

results\_df <- do.call(rbind, results\_list)  
colnames(results\_df) <- c("Hazard Ratio", "Lower CI", "Upper CI", "P-value")  
rownames(results\_df) <- names(results\_list)  
results\_df <- data.frame(Protein = rownames(results\_df), results\_df, row.names = NULL)  
  
results\_df$Protein <- gsub("X", "", results\_df$Protein)  
results\_df$Protein <- gsub("\\.", "-", results\_df$Protein)  
  
id\_to\_name <- setNames(panel\_data$Protein\_name, panel\_data$Protein\_ID)  
results\_df$Protein <- id\_to\_name[results\_df$Protein]  
View(results\_df)  
  
  
## Create plot  
significance\_threshold <- 0.01  
  
volcano\_df <- results\_df %>%  
 mutate(Estimate = `Hazard.Ratio`,  
 significance = case\_when(  
 as.numeric(P.value) < significance\_threshold & `Hazard.Ratio` > 1 ~ "Increased Risk",  
 as.numeric(P.value) < significance\_threshold & `Hazard.Ratio` < 1 ~ "Decreased Risk",  
 TRUE ~ "Non-significant"  
 ))  
  
volcano\_df$significance <- factor(volcano\_df$significance, levels = c("Non-significant", "Decreased Risk", "Increased Risk"))  
  
  
volcano\_plot <- ggplot(volcano\_df, aes(x = Estimate, y = -log10(P.value), color = significance)) +  
 geom\_point(alpha = 0.9, size = 8) +  
 scale\_color\_manual(  
 values = c("Non-significant" = "darkgray", "Decreased Risk" = "darkblue", "Increased Risk" = "red2"),  
 labels = c("Non-significant", "Decreased Risk", "Increased Risk")  
 ) +  
 geom\_label\_repel(aes(label = ifelse(as.numeric(P.value) < significance\_threshold & Estimate < 1, Protein, "")),  
 size = 12, #   
 box.padding = unit(0.5, "lines"),   
 point.padding = unit(0.5, "lines"),  
 segment.color = 'grey50',   
 max.overlaps = Inf,  
 color = 'black',  
 fill = 'white') +  
 geom\_label\_repel(aes(label = ifelse(as.numeric(P.value) < significance\_threshold & Estimate >= 1, Protein, "")),  
 size = 12,   
 box.padding = unit(0.5, "lines"),   
 point.padding = unit(0.5, "lines"),  
 segment.color = 'grey50',   
 max.overlaps = Inf,  
 color = 'black',  
 fill = 'white') +  
 theme\_minimal() +  
 labs(title = "",  
 x = "Hazard Ratio",  
 y = "-log10(p-value)") +  
 theme(legend.position = "bottom",  
 legend.title = element\_blank(),  
 legend.text = element\_text(size = 14),  
 panel.grid = element\_blank(),   
 plot.title = element\_text(hjust = 0.5, size = 16, face = "bold"),  
 axis.title = element\_text(size = 14, colour = "black"),  
 axis.text = element\_text(size = 14, colour = "black")) +  
 geom\_hline(yintercept = -log10(significance\_threshold), linetype = "dashed", color = "red") +  
 geom\_vline(xintercept = c(1), linetype = "dashed", color = "darkgrey")  
  
print(volcano\_plot)

A graph with text and numbers

Description automatically generated with medium confidence

**Cytokines and cognitive decline**

# Load libraries  
library(dplyr)

library(readr)  
library(ggplot2)  
library(lme4)

library(lmerTest)

library(broom.mixed)  
library(writexl)  
library(tidyr)

library(Matrix)

library(ggrepel)  
  
V40\_data <- read.csv("V40\_anmlSMP.csv")  
V41\_data <- read.csv("V41\_anmlSMP.csv")  
panel\_data <- read.csv("Cytokine\_panelSL.csv")  
apoe4\_data <- read.csv("apoe4.csv")  
  
# Rank-inverse normal transformation   
invnormal <- function(x) {  
 qnorm((rank(x, na.last = "keep") - 0.5) / sum(!is.na(x)))  
}  
  
V40\_data[-1] <- apply(V40\_data[-1], 2, invnormal)  
V41\_data[-1] <- apply(V41\_data[-1], 2, invnormal)  
V40\_data <- setNames(V40\_data, gsub("X", "", gsub("\\.", "-", names(V40\_data))))  
V41\_data <- setNames(V41\_data, gsub("X", "", gsub("\\.", "-", names(V41\_data))))  
  
# Filter proteins and prepare datasets  
proteins\_of\_interest <- panel\_data$Protein\_ID   
V40\_filtered <- V40\_data[, c("ID\_OUT", intersect(names(V40\_data)[-1], proteins\_of\_interest))]  
V41\_filtered <- V41\_data[, c("ID\_OUT", intersect(names(V41\_data)[-1], proteins\_of\_interest))]  
common\_proteins <- intersect(names(V40\_filtered)[-1], names(V41\_filtered)[-1])  
V41\_filtered\_common <- V41\_filtered[, c("ID\_OUT", common\_proteins)]  
  
appended\_data <- bind\_rows(V40\_filtered, V41\_filtered\_common)  
unique\_appended\_data <- distinct(appended\_data, ID\_OUT, .keep\_all = TRUE)  
  
covariate\_data <- read.csv("Covariates\_p5.csv")  
age\_data <- read.csv("Age.csv")  
merged\_data <- merge(unique\_appended\_data, covariate\_data, by = "ID\_OUT", all.x = TRUE)  
merged\_data <- merge(merged\_data, age\_data, by = "ID\_OUT", all.x = TRUE)  
merged\_data <- merge(merged\_data, apoe4\_data, by = "ID\_OUT", all.x = TRUE)  
covariates <- c("age", "sex", "ethn\_ds", "apoe4", "w5\_PhyAct", "w5\_smoking", "w5\_education", "w5\_alcohol", "time\_centered")  
data <- read.csv("Final\_dataset\_cog\_decline.csv")  
data <- data %>%  
 distinct(ID\_OUT, .keep\_all = TRUE)  
  
  
# Cognition  
Fluency1.sd <- sd(data$TANIMALS, na.rm = TRUE)  
Fluency1.mean <- mean(data$TANIMALS, na.rm = TRUE)  
Fluency2.sd <- sd(data$TSWORDS, na.rm = TRUE)  
Fluency2.mean <- mean(data$TSWORDS, na.rm = TRUE)  
Memory.sd <- sd(data$TMEM, na.rm = TRUE)  
Memory.mean <- mean(data$TMEM, na.rm = TRUE)  
Reasoning.sd <- sd(data$TAH4, na.rm = TRUE)  
Reasoning.mean <- mean(data$TAH4, na.rm = TRUE)  
  
data$fluency1.1 <- (data$TANIMALS - Fluency1.mean) / Fluency1.sd  
data$fluency1.2 <- (data$MANIMALS - Fluency1.mean) / Fluency1.sd  
data$fluency1.3 <- (data$JANIMALS - Fluency1.mean) / Fluency1.sd  
data$fluency1.4 <- (data$FANIMALS - Fluency1.mean) / Fluency1.sd  
data$fluency1.5 <- (data$DANIMALS - Fluency1.mean) / Fluency1.sd  
data$fluency2.1 <- (data$TSWORDS - Fluency2.mean) / Fluency2.sd  
data$fluency2.2 <- (data$MSWORDS - Fluency2.mean) / Fluency2.sd  
data$fluency2.3 <- (data$JSWORDS - Fluency2.mean) / Fluency2.sd  
data$fluency2.4 <- (data$FSWORDS - Fluency2.mean) / Fluency2.sd  
data$fluency2.5 <- (data$DSWORDS - Fluency2.mean) / Fluency2.sd  
data$memory.1 <- (data$TMEM - Memory.mean) / Memory.sd  
data$memory.2 <- (data$MMEM - Memory.mean) / Memory.sd  
data$memory.3 <- (data$JMEM - Memory.mean) / Memory.sd  
data$memory.4 <- (data$FMEM - Memory.mean) / Memory.sd  
data$memory.5 <- (data$DMEM - Memory.mean) / Memory.sd  
data$reasoning.1 <- (data$TAH4 - Reasoning.mean) / Reasoning.sd  
data$reasoning.2 <- (data$MAH4 - Reasoning.mean) / Reasoning.sd  
data$reasoning.3 <- (data$JAH4 - Reasoning.mean) / Reasoning.sd  
data$reasoning.4 <- (data$FAH4 - Reasoning.mean) / Reasoning.sd  
data$reasoning.5 <- (data$DAH4 - Reasoning.mean) / Reasoning.sd  
  
data$gscore.1 <- rowSums(data[, c('fluency1.1', 'fluency2.1', 'memory.1', 'reasoning.1')])  
data$gscore.2 <- rowSums(data[, c('fluency1.2', 'fluency2.2', 'memory.2', 'reasoning.2')])  
data$gscore.3 <- rowSums(data[, c('fluency1.3', 'fluency2.3', 'memory.3', 'reasoning.3')])  
data$gscore.4 <- rowSums(data[, c('fluency1.4', 'fluency2.4', 'memory.4', 'reasoning.4')])  
data$gscore.5 <- rowSums(data[, c('fluency1.5', 'fluency2.5', 'memory.5', 'reasoning.5')])  
gscore.mean <- mean(data$gscore.1, na.rm = TRUE)  
gscore.sd <- sd(data$gscore.1, na.rm = TRUE)  
data$gscore.z1 <- (data$gscore.1 - gscore.mean) / gscore.sd  
data$gscore.z2 <- (data$gscore.2 - gscore.mean) / gscore.sd  
data$gscore.z3 <- (data$gscore.3 - gscore.mean) / gscore.sd  
data$gscore.z4 <- (data$gscore.4 - gscore.mean) / gscore.sd  
data$gscore.z5 <- (data$gscore.5 - gscore.mean) / gscore.sd  
  
data <- data %>%  
 rename(  
 Date1 = TDATSCRN,  
 Date2 = MDATSCRN,  
 Date3 = JDATSCRN,  
 Date4 = FDATSCRN,  
 Date5 = DDATSCRN  
 )  
  
data\_new <- data %>% select(ID\_OUT, gscore.z1, gscore.z2, gscore.z3, gscore.z4, gscore.z5, Date1, Date2, Date3, Date4, Date5)  
  
  
data\_long\_scores <- pivot\_longer(data\_new,  
 cols = starts\_with("gscore.z"),  
 names\_to = "timepoint",  
 names\_prefix = "gscore.z",  
 values\_to = "cog\_decline")  
  
data\_long\_dates <- pivot\_longer(data\_new,  
 cols = starts\_with("Date"),  
 names\_to = "date\_timepoint",  
 names\_prefix = "Date",  
 values\_to = "date")  
  
  
data\_long\_dates$date <- as.Date(data\_long\_dates$date, format = "%d-%b-%y")  
  
data\_long <- data\_long\_scores %>%  
 left\_join(data\_long\_dates, by = c("ID\_OUT", "timepoint" = "date\_timepoint")) %>%  
 select(ID\_OUT, cog\_decline, timepoint, date)  
  
  
data\_long <- data\_long %>%  
 group\_by(ID\_OUT) %>%  
 mutate(time = as.numeric(difftime(date, first(date), units = "days")) / 365.25) %>%  
 ungroup()  
  
data\_filtered <- data\_long %>%  
 group\_by(ID\_OUT) %>%  
 filter(sum(!is.na(cog\_decline)) >= 2) %>%  
 ungroup()  
  
merged\_data\_filtered <- merge(data\_filtered, merged\_data, by = "ID\_OUT", all.x = TRUE)  
  
merged\_data\_filtered <- merged\_data\_filtered %>%  
 mutate(  
 age\_centered = scale(age, center = TRUE, scale = TRUE),  
 time\_centered = scale(time, center = TRUE, scale = TRUE)  
 )  
  
merged\_data\_filtered <- merged\_data\_filtered %>%  
 select(-w5\_PhyAct, -w5\_smoking, -w5\_education, -w5\_alcohol, -date, -time, -age, -apoe4)   
  
  
original\_protein\_names <- names(merged\_data\_filtered)[4:(length(names(merged\_data\_filtered)))]  
valid\_protein\_names <- make.names(original\_protein\_names, unique = TRUE)  
  
name\_mapping <- data.frame(  
 original = original\_protein\_names,  
 valid = valid\_protein\_names,  
 stringsAsFactors = FALSE  
)  
  
names(merged\_data\_filtered)[4:(length(names(merged\_data\_filtered)))] <- valid\_protein\_names  
  
covariates <- c("age\_centered", "sex", "ethn\_ds")  
time\_centered <- c("time\_centered")  
  
merged\_data\_filtered <- merged\_data\_filtered %>%  
 filter(  
 !is.na(cog\_decline) &  
 !is.na(age\_centered) &  
 !is.na(sex) &  
 !is.na(ethn\_ds) &  
 !is.na(time\_centered)  
 )

results\_list <- list()  
  
# Linear mixed models  
for (protein in setdiff(valid\_protein\_names, covariates)) {  
 formula <- as.formula(paste("cog\_decline ~", protein, "\*", time\_centered, "+", paste(covariates, collapse = " + "), "+ (", time\_centered, "| ID\_OUT)"))   
 model <- tryCatch({  
 lmer(formula, data = merged\_data\_filtered, control = lmerControl(optimizer = "bobyqa"))  
 }, error = function(e) {  
 print(paste("Error", protein))  
 print(e)  
 return(NULL)  
 })  
   
 if (!is.null(model)) {  
 coefs <- summary(model)$coefficients  
 if (protein %in% rownames(coefs)) {  
 protein\_main\_effect <- coefs[protein, ]  
 estimate <- protein\_main\_effect["Estimate"]  
 ci\_lower <- estimate - 1.96 \* protein\_main\_effect["Std. Error"]  
 ci\_upper <- estimate + 1.96 \* protein\_main\_effect["Std. Error"]  
 p\_value <- protein\_main\_effect["Pr(>|t|)"]  
 results\_list[[protein]] <- c(estimate, ci\_lower, ci\_upper, p\_value)  
 } else {  
 results\_list[[protein]] <- rep(NA, 4)  
 }  
 } else {  
 results\_list[[protein]] <- rep(NA, 4)  
 }  
}  
  
results\_df <- do.call(rbind, results\_list)  
colnames(results\_df) <- c("Estimate", "Lower CI", "Upper CI", "P-value")  
rownames(results\_df) <- names(results\_list)  
results\_df <- data.frame(Protein = rownames(results\_df), results\_df, row.names = NULL)  
results\_df <- results\_df[!results\_df$Protein %in% c("time\_centered", covariates), ]  
results\_df$Protein <- name\_mapping$original[match(results\_df$Protein, name\_mapping$valid)]  
results\_df$Protein <- gsub("X", "", results\_df$Protein)  
results\_df$Protein <- gsub("\\.", "-", results\_df$Protein)  
  
id\_to\_name <- setNames(panel\_data$Protein\_name, panel\_data$Protein\_ID)  
results\_df$Protein <- id\_to\_name[results\_df$Protein]  
  
  
significance\_threshold <- 0.01  
  
# Create plot  
volcano\_df <- results\_df %>%  
 mutate(label = ifelse(as.numeric(P.value) < significance\_threshold, Protein, NA),  
 significance = case\_when(  
 as.numeric(P.value) < significance\_threshold & Estimate > 0 ~ "Decreased Risk",  
 as.numeric(P.value) < significance\_threshold & Estimate < 0 ~ "Increased Risk",  
 TRUE ~ "Non-significant"  
 ))  
  
volcano\_df$significance <- factor(volcano\_df$significance, levels = c("Non-significant", "Decreased Risk", "Increased Risk"))  
  
volcano\_plot2 <- ggplot(volcano\_df, aes(x = -Estimate, y = -log10(P.value), color = significance)) +  
 geom\_point(alpha = 0.9, size = 8) +  
 scale\_color\_manual(  
 values = c("Non-significant" = "darkgray", "Decreased Risk" = "darkblue", "Increased Risk" = "red2"),  
 labels = c("Non-significant", "Decreased Risk", "Increased Risk")  
 ) +  
 geom\_label\_repel(aes(label = ifelse(as.numeric(P.value) < significance\_threshold & Estimate < 0, Protein, "")),  
 size = 12,   
 box.padding = unit(0.5, "lines"),   
 point.padding = unit(0.5, "lines"),  
 segment.color = 'grey50',   
 max.overlaps = Inf,  
 color = 'black',  
 fill = 'white') +  
 geom\_label\_repel(aes(label = ifelse(as.numeric(P.value) < significance\_threshold & Estimate >= 0, Protein, "")),  
 size = 12,   
 box.padding = unit(0.5, "lines"),   
 point.padding = unit(0.5, "lines"),  
 segment.color = 'grey50',   
 max.overlaps = Inf,  
 color = 'black',  
 fill = 'white') +  
 theme\_minimal() +  
 labs(title = "",  
 x = "Beta coefficient",  
 y = "-log10(p-value)") +  
 theme(legend.position = "bottom",  
 legend.title = element\_blank(),  
 legend.text = element\_text(size = 14,colour = "black"),  
 panel.grid = element\_blank(),   
 plot.title = element\_text(hjust = 0.5, size = 16, face = "bold"),  
 axis.title = element\_text(size = 14, colour = "black"),  
 axis.text = element\_text(size = 14, colour = "black")) +  
 scale\_x\_continuous(labels = function(x) -x) +   
 geom\_hline(yintercept = -log10(significance\_threshold), linetype = "dashed", color = "red") +  
 geom\_vline(xintercept = c(0), linetype = "dashed", color = "darkgrey")  
  
print(volcano\_plot2)

A diagram of numbers and letters

Description automatically generated with medium confidence

SAS (version 9.4)

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*;

\*\*\* AGE-RELATED DISEASES \*\*\*;

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*;

\*\*\*\*\*\*\*\* Cox model (incident diseases) \*\*\*\*\*\*\*\*;

**%macro** hallmark (dg,prot,resfile);

data hm;

set diseases;

IF dgnro=&dg;

keep STNO slaalkupvm;

data wh2;

merge wh1(in=i) hm;

by STNO;

if i;

data wh3;

set wh2;

futime\_hall=(min(slaalkupvm,enddate)-startdate)/**365.25**;

if **.**<slaalkupvm<=enddate then status\_hall=**1**; else status\_hall=**0**;

IF futime\_hall>**0**;

age\_end=age+futime\_hall;

run;

proc means data=wh3 noprint; where status\_hall=**1**;

var age\_end; output out=ma mean= std=std;

proc phreg data=wh3;

model futime\_hall\*status\_hall(**0**) = &prot sex age ETHN\_DS / rl;

ods output ParameterEstimates=pe CensoredSummary=cs;

data res; merge pe(obs=**1**) cs ma; dgnro=&dg; format exposure $20.; Exposure=Parameter;

keep dgnro Total Event Exposure HazardRatio HRLowerCL HRUpperCL ProbChiSq age\_end std;

data res; merge res(in=i) Dgnro\_labels; by dgnro; if i;

keep Exposure dgnro label category Total Event HazardRatio HRLowerCL HRUpperCL ProbChiSq age\_end std;

run;

proc append base=&resfile data=res;

run;

**%mend**;

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*;

\*\* all diseases \*\*;

**proc** **datasets** lib=work memtype=data nolist; delete results\_all; **quit**;

**%MACRO** ***hallmarkall***;

%DO I = **1** %TO **83**;

%***hallmark***(&I,mic,results\_all);

%END;

**%MEND** hallmarkall;

%***hallmarkall***;

**proc** **print** data=results\_all;

id dgnro;

var label category Total Event HazardRatio HRLowerCL HRUpperCL ProbChiSq;

**run**;

\*\* Disease groups \*\*;

**proc** **datasets** lib=work memtype=data nolist; delete results\_all; **quit**;

%***hallmark***(**201**,mic,results\_all);

%***hallmark***(**202**,mic,results\_all);

%***hallmark***(**203**,mic,results\_all);

%***hallmark***(**204**,mic,results\_all);

%***hallmark***(**205**,mic,results\_all);

%***hallmark***(**206**,mic,results\_all);

%***hallmark***(**207**,mic,results\_all);

%***hallmark***(**208**,mic,results\_all);

**proc** **print** data=results\_all;

id dgnro;

var label category Total Event HazardRatio HRLowerCL HRUpperCL ProbChiSq;

**run**;

\*\* Dementia \*\*;

**proc** **datasets** lib=work memtype=data nolist; delete results\_all; **quit**;

%***hallmark***(**42**,mic,results\_all);

%***hallmark***(**421**,mic,results\_all); \*Alzheimer's\*;

%***hallmark***(**422**,mic,results\_all); \*Vascular\*;

**proc** **print** data=results\_all;

id dgnro;

var label category Total Event HazardRatio HRLowerCL HRUpperCL ProbChiSq;

**run**;

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*;

\*\*\* MEDIATION \*\*\*;

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*;

**data** diseases2;

set diseases;

IF dgnro IN (**42**,**1**,**2**,**3**,**4**,**8**,**11**,**30**,**37**,**38**,**39**,**41**,**47**,**48**,**50**,**52**,**53**,**54**,**55**,**60**,

**61**,**63**,**64**,**65**,**67**,**68**,**69**,**70**,**72**,**79**,**81**,**82**);

**run**;

\* any disease \*;

**proc** **means** data=diseases2 nway noprint;

var slaalkupvm;

class STNO;

output out=anydis min=slaalkupvm1;

**run**;

**data** wh2;

merge wh1(in=i) anydis;

by STNO;

if i;

**run**;

**data** wh3\_nodis;

set wh2;

IF **.**<slaalkupvm1<=startdate then delete;

**run**;

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*;

\* Dementia(42), Vascular dementia(422) \*;

**data** demen;

set diseases;

IF dgnro=**42**;

rename slaalkupvm=demen\_date;

**run**;

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*;

\*\*\*\*\*\*\*\* Cox model (dementia) \*\*\*\*\*\*\*\*;

**%macro** mediation (dg,resfile);

data med;

set diseases;

IF dgnro=&dg;

rename slaalkupvm=med\_date;

data wh4;

merge wh3\_nodis(in=i) demen med;

by STNO;

if i;

futime\_demen=(min(demen\_date,enddate)-startdate)/**365.25**;

if **.**<demen\_date<=enddate then status\_demen=**1**; else status\_demen=**0**;

\*mediator\*;

if med\_date>**.** then med\_disease=**1**; else med\_disease=**0**;

if **.**<demen\_date<med\_date then med\_disease=**0**;

run;

proc phreg data=wh4;

model futime\_demen\*status\_demen(**0**) = mic med\_disease sex age ETHN\_DS / rl;

ods output ParameterEstimates=pe CensoredSummary=cs;

run;

data res; merge pe(obs=**1**) cs; dgnro=&dg;

keep dgnro Total Event HazardRatio HRLowerCL HRUpperCL;

data res; merge res(in=i) Dgnro\_labels; by dgnro; if i;

keep dgnro label Total Event HazardRatio HRLowerCL HRUpperCL;

run;

proc append base=&resfile data=res;

run;

**%mend**;

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*;

\*\* Mediators \*\*;

**proc** **datasets** lib=work memtype=data nolist; delete results\_all; **quit**;

%***mediation***(**8**,results\_all);

%***mediation***(**11**,results\_all);

%***mediation***(**52**,results\_all);

%***mediation***(**53**,results\_all);

%***mediation***(**54**,results\_all);

%***mediation***(**55**,results\_all);

%***mediation***(**60**,results\_all);

%***mediation***(**61**,results\_all);

%***mediation***(**63**,results\_all);

%***mediation***(**67**,results\_all);

%***mediation***(**68**,results\_all);

%***mediation***(**69**,results\_all);

%***mediation***(**70**,results\_all);

%***mediation***(**37**,results\_all);

%***mediation***(**38**,results\_all);

%***mediation***(**39**,results\_all);

%***mediation***(**41**,results\_all);

%***mediation***(**81**,results\_all);

%***mediation***(**82**,results\_all);

%***mediation***(**1**,results\_all);

%***mediation***(**2**,results\_all);

%***mediation***(**3**,results\_all);

%***mediation***(**4**,results\_all);

%***mediation***(**64**,results\_all);

%***mediation***(**65**,results\_all);

%***mediation***(**48**,results\_all);

%***mediation***(**50**,results\_all);

%***mediation***(**30**,results\_all);

%***mediation***(**47**,results\_all);

%***mediation***(**72**,results\_all);

%***mediation***(**79**,results\_all);

%***mediation***(**201**,results\_all);

%***mediation***(**202**,results\_all);

%***mediation***(**203**,results\_all);

%***mediation***(**204**,results\_all);

%***mediation***(**205**,results\_all);

%***mediation***(**206**,results\_all);

%***mediation***(**207**,results\_all);

%***mediation***(**208**,results\_all);

**proc** **print** data=results\_all;

id dgnro;

var label Total Event HazardRatio HRLowerCL HRUpperCL;

**run**;

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*;

\*\*\*\*\*\*\*\* Mediators = all groups \*\*\*\*\*\*\*\*;

**data** med201; set diseases; IF dgnro=**201**; rename slaalkupvm=med\_date201;

**data** med202; set diseases; IF dgnro=**202**; rename slaalkupvm=med\_date202;

**data** med203; set diseases; IF dgnro=**203**; rename slaalkupvm=med\_date203;

**data** med204; set diseases; IF dgnro=**204**; rename slaalkupvm=med\_date204;

**data** med205; set diseases; IF dgnro=**205**; rename slaalkupvm=med\_date205;

**data** med206; set diseases; IF dgnro=**206**; rename slaalkupvm=med\_date206;

**data** med207; set diseases; IF dgnro=**207**; rename slaalkupvm=med\_date207;

**data** med208; set diseases; IF dgnro=**208**; rename slaalkupvm=med\_date208;

**data** med30; set diseases; IF dgnro=**30**; rename slaalkupvm=med\_date30;

**data** med47; set diseases; IF dgnro=**47**; rename slaalkupvm=med\_date47;

**data** med72; set diseases; IF dgnro=**72**; rename slaalkupvm=med\_date72;

**data** med79; set diseases; IF dgnro=**79**; rename slaalkupvm=med\_date79;

**run**;

**data** wh4;

merge wh3\_nodis(in=i) demen med201 med202 med203 med204 med205 med206

med207 med208 med30 med47 med72 med79;

by STNO;

if i;

futime\_demen=(min(demen\_date,enddate)-startdate)/**365.25**;

if **.**<demen\_date<=enddate then status\_demen=**1**; else status\_demen=**0**;

\*mediators\*;

if med\_date201>**.** then med\_disease201=**1**; else med\_disease201=**0**; if **.**<demen\_date<med\_date201 then med\_disease201=**0**;

if med\_date202>**.** then med\_disease202=**1**; else med\_disease202=**0**; if **.**<demen\_date<med\_date202 then med\_disease202=**0**;

if med\_date203>**.** then med\_disease203=**1**; else med\_disease203=**0**; if **.**<demen\_date<med\_date203 then med\_disease203=**0**;

if med\_date204>**.** then med\_disease204=**1**; else med\_disease204=**0**; if **.**<demen\_date<med\_date204 then med\_disease204=**0**;

if med\_date205>**.** then med\_disease205=**1**; else med\_disease205=**0**; if **.**<demen\_date<med\_date205 then med\_disease205=**0**;

if med\_date206>**.** then med\_disease206=**1**; else med\_disease206=**0**; if **.**<demen\_date<med\_date206 then med\_disease206=**0**;

if med\_date207>**.** then med\_disease207=**1**; else med\_disease207=**0**; if **.**<demen\_date<med\_date207 then med\_disease207=**0**;

if med\_date208>**.** then med\_disease208=**1**; else med\_disease208=**0**; if **.**<demen\_date<med\_date208 then med\_disease208=**0**;

if med\_date30>**.** then med\_disease30=**1**; else med\_disease30=**0**; if **.**<demen\_date<med\_date30 then med\_disease30=**0**;

if med\_date47>**.** then med\_disease47=**1**; else med\_disease47=**0**; if **.**<demen\_date<med\_date47 then med\_disease47=**0**;

if med\_date72>**.** then med\_disease72=**1**; else med\_disease72=**0**; if **.**<demen\_date<med\_date72 then med\_disease72=**0**;

if med\_date79>**.** then med\_disease79=**1**; else med\_disease79=**0**; if **.**<demen\_date<med\_date79 then med\_disease79=**0**;

med\_disease\_sum=sum(of med\_disease201-med\_disease208);

med\_disease\_sum2=sum(of med\_disease201-med\_disease205 med\_disease207-med\_disease208);

**run**;

**proc** **phreg** data=wh4;

model futime\_demen\*status\_demen(**0**) = mic med\_disease201 med\_disease202 med\_disease203 med\_disease204

med\_disease205 med\_disease206 med\_disease207 med\_disease208

sex age ETHN\_DS / rl;

**run**;

**proc** **phreg** data=wh4;

model futime\_demen\*status\_demen(**0**) = mic med\_disease\_sum

sex age ETHN\_DS / rl;

**run**;

**proc** **phreg** data=wh4;

model futime\_demen\*status\_demen(**0**) = mic med\_disease\_sum2

sex age ETHN\_DS / rl;

**run**;

**proc** **phreg** data=wh4;

model futime\_demen\*status\_demen(**0**) = mic med\_disease206 med\_disease\_sum2

sex age ETHN\_DS / rl;

**run**;

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*;

\*\*\* TRAJECTORIES (BMI, walking speed, grip strength, physical activity) \*\*\*;

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*;

\*\*\* Unadjusted means \*\*\*;

**proc** **means** data=wh5 fw=**5** n mean stderr;

var TBMI MBMI JBMI FBMI

mwalkspd jwalkspd fwalkspd

jmaxgrip fmaxgrip

jmetwk fmetwk;

class mic;

**run**;

\*\*\* Adjusted means \*\*\*;

**data** wh6;

set wh5;

year=**0**; BMI=TBMI; output;

year=**5**; BMI=MBMI; walkspd=Mwalkspd; output;

year=**10**; BMI=JBMI; walkspd=Jwalkspd; maxgrip=Jmaxgrip; metwk=Jmetwk; output;

year=**15**; BMI=FBMI; walkspd=Fwalkspd; maxgrip=Fmaxgrip; metwk=Fmetwk; output;

keep ID\_OUT mic year SEX eth age5 BMI walkspd maxgrip metwk;

**run**;

\* BMI, walkspd, maxgrip, metwk \*;

**proc** **genmod** data=wh6;

class ID\_OUT mic year;

model BMI = SEX eth age5 mic year mic\*year / type3;

lsmeans mic\*time / cl;

repeated subject=ID\_OUT / within=year type=exch;

**run**;