rm(list=ls())

library(dplyr)

library(moonBook)

library(readxl)

library(survival)

library(car)

#library(maxstat)

library(caret)

library(ggplot2)

rm(list=ls())

data\_org = read\_excel("real\_id31\_table4.xlsx") %>% data.frame()

####################################################

#전체 데이터 전처리

####################################################

data = data\_org

dim(data)

data$age\_g = factor(data$age\_g)

#data[data$dth\_fu == 0,]

data$Cancer\_meta = ifelse(data$Cancer == 1 | data$metastatic == 1,1,0)

names(data)

data$nocd = data$Arthritis + data$Asthma + data$spectrum + data$Cancer\_meta + data$cardiac\_arrhythmias + data$kidney +

data$pulmonary + data$hf + data$Coronary + data$Dementia + data$Depression + data$Diabetes + data$Hepatitis + data$HIV +

data$Dyslipidemia + data$HTN + data$Osteoporosis + data$Schizophrenia +

data$Stroke + data$SAD

# nocd\_g (number of chronic disease grouping)

data$nocd\_g = ifelse(data$nocd >= 5,5,data$nocd) %>% factor()

# smk\_g (smk grouping)

data$smk\_g = ifelse(is.na(data$Q\_SMK\_YN),NA,

ifelse(data$Q\_SMK\_YN==3,1,0))

# Surger & Chemotherapy 移댄뀒怨좊━

data$surgChemo = data$Surgery + data$chemotherapy

data$surgChemo\_g = ifelse(data$surgChemo >= 2,1,0)

# Surger & Radiation 移댄뀒怨좊━

data$surgRtx = data$Surgery + data$radiation

data$surgRtx\_g = ifelse(data$surgRtx >= 2,1,0)

# chemotherapy & Radiation 移댄뀒怨좊━

data$chemoRtx = data$chemotherapy + data$radiation

data$chemoRtx\_g = ifelse(data$chemoRtx >= 2,1,0)

# Surger & Chemotherapy & Radiation 移댄뀒怨좊━

data$surgChemoRtx = data$Surgery + data$chemotherapy + data$radiation

data$surgChemoRtx\_g = ifelse(data$surgChemoRtx >= 3,1,0)

# None Treatment grouping

data$nonetreatment\_g = ifelse(data$surgChemoRtx == 0,1,0)

names(data)

data$treatment = ifelse(data$surgChemoRtx\_g == 1,7,

ifelse(data$chemoRtx\_g == 1,6,

ifelse(data$surgRtx\_g == 1,5,

ifelse(data$surgChemo\_g == 1,4,

ifelse(data$radiation == 1,3,

ifelse(data$Surgery == 1,2,

ifelse(data$chemotherapy == 1,1,0))))))) %>% factor()

names(data)

data = data %>% select(dth\_fu, dth, SEX\_TYPE, age\_g, Arthritis, Asthma, spectrum,

Cancer,metastatic, cardiac\_arrhythmias, kidney,

pulmonary, hf, Coronary, Dementia,

Depression,Diabetes, Hepatitis, Dyslipidemia, HIV,

HTN, Osteoporosis, Schizophrenia,

Stroke, SAD, nocd\_g, smk\_g, drk\_g, pa\_g, low\_incm, treatment,

bmi\_g, mx\_ER,admin\_day,

age, G1E\_BMI, G1E\_BP\_SYS, G1E\_BP\_DIA, G1E\_TOT\_CHOL, mx\_Seoul,

Cancer\_meta) %>% na.omit()

names(data)

data$mx\_Seoul %>% table()

data$W = 1

data$W1 = data$W

data$admin\_day\_g = ifelse(data$admin\_day == 0,0,

ifelse(data$admin\_day < 7,1,

ifelse(data$admin\_day < 14,2,3)))

#######################################################################################################################################################################################

# Table1

#######################################################################################################################################################################################

#############################################################

# Total -Table1

#############################################################

Table1\_all\_mean =

mytable(W1~SEX\_TYPE + age\_g + Arthritis + Asthma + spectrum +

Cancer + metastatic + cardiac\_arrhythmias + kidney +

pulmonary + hf + Coronary + Dementia +

Depression + Diabetes + Hepatitis + Dyslipidemia + HIV +

HTN + Osteoporosis + Schizophrenia +

Stroke + SAD + nocd\_g + smk\_g + drk\_g + pa\_g + low\_incm + treatment+

bmi\_g + mx\_ER + admin\_day\_g +

age + G1E\_BMI + G1E\_BP\_SYS + G1E\_BP\_DIA + G1E\_TOT\_CHOL+admin\_day +

dth + dth\_fu,data = data, method = 1,

digits = 2)

mycsv(Table1\_all\_mean, file = "Table1\_all\_mean.csv")

Table1\_all\_median =

mytable(W1~age + G1E\_BMI + G1E\_BP\_SYS + G1E\_BP\_DIA + G1E\_TOT\_CHOL+admin\_day +

dth + dth\_fu,data = data, method = 2,

digits = 2)

mycsv(Table1\_all\_median, file = "Table1\_all\_median.csv")

Table1\_all\_mean\_sex\_class =

mytable(SEX\_TYPE~ age\_g + Arthritis + Asthma + spectrum +

Cancer + metastatic + cardiac\_arrhythmias + kidney +

pulmonary + hf + Coronary + Dementia +

Depression + Diabetes + Hepatitis + Dyslipidemia + HIV +

HTN + Osteoporosis + Schizophrenia +

Stroke + SAD + nocd\_g + smk\_g + drk\_g + pa\_g + low\_incm + treatment+

bmi\_g + mx\_ER + admin\_day\_g +

age + G1E\_BMI + G1E\_BP\_SYS + G1E\_BP\_DIA + G1E\_TOT\_CHOL+admin\_day +

dth + dth\_fu,data = data, method = 1,

digits = 2)

mycsv(Table1\_all\_mean\_sex\_class, file = "Table1\_all\_mean\_sex\_class.csv")

Table1\_all\_median\_sex\_class =

mytable(SEX\_TYPE ~ age + G1E\_BMI + G1E\_BP\_SYS + G1E\_BP\_DIA + G1E\_TOT\_CHOL+admin\_day +

dth + dth\_fu,data = data, method = 2,

digits = 2)

mycsv(Table1\_all\_median\_sex\_class, file = "Table1\_all\_median\_sex\_class.csv")

#############################################################

# Seoul -Table1

#############################################################

Seoul = data %>% filter(mx\_Seoul == 1)

Table1\_Seoul\_mean =

mytable(W1~SEX\_TYPE + age\_g + Arthritis + Asthma + spectrum +

Cancer + metastatic + cardiac\_arrhythmias + kidney +

pulmonary + hf + Coronary + Dementia +

Depression + Diabetes + Hepatitis + Dyslipidemia + HIV +

HTN + Osteoporosis + Schizophrenia +

Stroke + SAD + nocd\_g + smk\_g + drk\_g + pa\_g + low\_incm + treatment+

bmi\_g + mx\_ER + admin\_day\_g +

age + G1E\_BMI + G1E\_BP\_SYS + G1E\_BP\_DIA + G1E\_TOT\_CHOL+admin\_day +

dth + dth\_fu,data = Seoul, method = 1,

digits = 2)

mycsv(Table1\_Seoul\_mean, file = "Table1\_Seoul\_mean.csv")

Table1\_Seoul\_median =

mytable(W1~age + G1E\_BMI + G1E\_BP\_SYS + G1E\_BP\_DIA + G1E\_TOT\_CHOL+admin\_day +

dth + dth\_fu,data = data, method = 2,

digits = 2)

mycsv(Table1\_Seoul\_median, file = "Table1\_Seoul\_median.csv")

Table1\_Seoul\_mean\_sex\_class =

mytable(SEX\_TYPE~ age\_g + Arthritis + Asthma + spectrum +

Cancer + metastatic + cardiac\_arrhythmias + kidney +

pulmonary + hf + Coronary + Dementia +

Depression + Diabetes + Hepatitis + Dyslipidemia + HIV +

HTN + Osteoporosis + Schizophrenia +

Stroke + SAD + nocd\_g + smk\_g + drk\_g + pa\_g + low\_incm + treatment+

bmi\_g + mx\_ER + admin\_day\_g +

age + G1E\_BMI + G1E\_BP\_SYS + G1E\_BP\_DIA + G1E\_TOT\_CHOL+admin\_day +

dth + dth\_fu,data = Seoul, method = 1,

digits = 2)

mycsv(Table1\_Seoul\_mean\_sex\_class, file = "Table1\_Seoul\_mean\_sex\_class.csv")

Table1\_Seoul\_median\_sex\_class =

mytable(SEX\_TYPE ~ age + G1E\_BMI + G1E\_BP\_SYS + G1E\_BP\_DIA + G1E\_TOT\_CHOL+admin\_day +

dth + dth\_fu,data = Seoul, method = 2,

digits = 2)

mycsv(Table1\_Seoul\_median\_sex\_class, file = "Table1\_Seoul\_median\_sex\_class.csv")

#############################################################

# Busan -Table1

#############################################################

Busan = data %>% filter(mx\_Seoul == 0)

Table1\_Busan\_mean =

mytable(W1~SEX\_TYPE + age\_g + Arthritis + Asthma + spectrum +

Cancer + metastatic + cardiac\_arrhythmias + kidney +

pulmonary + hf + Coronary + Dementia +

Depression + Diabetes + Hepatitis + Dyslipidemia + HIV +

HTN + Osteoporosis + Schizophrenia +

Stroke + SAD + nocd\_g + smk\_g + drk\_g + pa\_g + low\_incm + treatment+

bmi\_g + mx\_ER + admin\_day\_g +

age + G1E\_BMI + G1E\_BP\_SYS + G1E\_BP\_DIA + G1E\_TOT\_CHOL+admin\_day +

dth + dth\_fu,data = Busan, method = 1,

digits = 2)

mycsv(Table1\_Busan\_mean, file = "Table1\_Busan\_mean.csv")

Table1\_Busan\_median =

mytable(W1~age + G1E\_BMI + G1E\_BP\_SYS + G1E\_BP\_DIA + G1E\_TOT\_CHOL+admin\_day +

dth + dth\_fu,data = Busan, method = 2,

digits = 2)

mycsv(Table1\_Busan\_median, file = "Table1\_Busan\_median.csv")

Table1\_Busan\_mean\_sex\_class =

mytable(SEX\_TYPE~ age\_g + Arthritis + Asthma + spectrum +

Cancer + metastatic + cardiac\_arrhythmias + kidney +

pulmonary + hf + Coronary + Dementia +

Depression + Diabetes + Hepatitis + Dyslipidemia + HIV +

HTN + Osteoporosis + Schizophrenia +

Stroke + SAD + nocd\_g + smk\_g + drk\_g + pa\_g + low\_incm + treatment+

bmi\_g + mx\_ER + admin\_day\_g +

age + G1E\_BMI + G1E\_BP\_SYS + G1E\_BP\_DIA + G1E\_TOT\_CHOL+admin\_day +

dth + dth\_fu,data = Busan, method = 1,

digits = 2)

mycsv(Table1\_Busan\_mean\_sex\_class, file = "Table1\_Busan\_mean\_sex\_class.csv")

Table1\_Busan\_median\_sex\_class =

mytable(SEX\_TYPE ~ age + G1E\_BMI + G1E\_BP\_SYS + G1E\_BP\_DIA + G1E\_TOT\_CHOL+admin\_day +

dth + dth\_fu,data = Busan, method = 2,

digits = 2)

mycsv(Table1\_Busan\_median\_sex\_class, file = "Table1\_Busan\_median\_sex\_class.csv")

#############################################################

# Total -Table1\_1

#############################################################

Table1\_1\_all\_mean\_nocd\_g =

mytable(nocd\_g~ age\_g + SEX\_TYPE + Arthritis + Asthma + spectrum +

Cancer + metastatic + cardiac\_arrhythmias + kidney +

pulmonary + hf + Coronary + Dementia +

Depression + Diabetes + Hepatitis + Dyslipidemia + HIV +

HTN + Osteoporosis + Schizophrenia +

Stroke + SAD + smk\_g + drk\_g + pa\_g + low\_incm + treatment+

bmi\_g + mx\_ER + admin\_day\_g +

age + G1E\_BMI + G1E\_BP\_SYS + G1E\_BP\_DIA + G1E\_TOT\_CHOL+admin\_day +

dth + dth\_fu,data = data, method = 1,

digits = 2)

mycsv(Table1\_1\_all\_mean\_nocd\_g, file = "Table1\_1\_all\_mean\_nocd\_g.csv")

Table1\_1\_all\_median\_nocd\_g =

mytable(nocd\_g ~ age + G1E\_BMI + G1E\_BP\_SYS + G1E\_BP\_DIA + G1E\_TOT\_CHOL+admin\_day +

dth + dth\_fu,data = data, method = 2,

digits = 2)

mycsv(Table1\_1\_all\_median\_nocd\_g, file = "Table1\_1\_all\_median\_nocd\_g.csv")

#############################################################

# Seoul -Table1\_1

#############################################################

Seoul = data %>% filter(mx\_Seoul == 1)

Table1\_1\_Seoul\_mean\_nocd\_g =

mytable(nocd\_g~ age\_g + SEX\_TYPE + Arthritis + Asthma + spectrum +

Cancer + metastatic + cardiac\_arrhythmias + kidney +

pulmonary + hf + Coronary + Dementia +

Depression + Diabetes + Hepatitis + Dyslipidemia + HIV +

HTN + Osteoporosis + Schizophrenia +

Stroke + SAD + smk\_g + drk\_g + pa\_g + low\_incm + treatment+

bmi\_g + mx\_ER + admin\_day\_g +

age + G1E\_BMI + G1E\_BP\_SYS + G1E\_BP\_DIA + G1E\_TOT\_CHOL+admin\_day +

dth + dth\_fu,data = Seoul, method = 1,

digits = 2)

mycsv(Table1\_1\_Seoul\_mean\_nocd\_g, file = "Table1\_1\_Seoul\_mean\_nocd\_g.csv")

Table1\_1\_Seoul\_median\_nocd\_g =

mytable(nocd\_g ~ age + G1E\_BMI + G1E\_BP\_SYS + G1E\_BP\_DIA + G1E\_TOT\_CHOL+admin\_day +

dth + dth\_fu,data = Seoul, method = 2,

digits = 2)

mycsv(Table1\_1\_Seoul\_median\_nocd\_g, file = "Table1\_1\_Seoul\_median\_nocd\_g.csv")

#############################################################

# Busan -Table1\_1

#############################################################

Busan = data %>% filter(mx\_Seoul == 0)

Table1\_1\_Busan\_mean\_nocd\_g =

mytable(nocd\_g~ age\_g + SEX\_TYPE + Arthritis + Asthma + spectrum +

Cancer + metastatic + cardiac\_arrhythmias + kidney +

pulmonary + hf + Coronary + Dementia +

Depression + Diabetes + Hepatitis + Dyslipidemia + HIV +

HTN + Osteoporosis + Schizophrenia +

Stroke + SAD + smk\_g + drk\_g + pa\_g + low\_incm + treatment+

bmi\_g + mx\_ER + admin\_day\_g +

age + G1E\_BMI + G1E\_BP\_SYS + G1E\_BP\_DIA + G1E\_TOT\_CHOL+admin\_day +

dth + dth\_fu,data = Busan, method = 1,

digits = 2)

mycsv(Table1\_1\_Busan\_mean\_nocd\_g, file = "Table1\_1\_Busan\_mean\_nocd\_g.csv")

Table1\_1\_Busan\_median\_nocd\_g =

mytable(nocd\_g ~ age + G1E\_BMI + G1E\_BP\_SYS + G1E\_BP\_DIA + G1E\_TOT\_CHOL+admin\_day +

dth + dth\_fu,data = Busan, method = 2,

digits = 2)

mycsv(Table1\_1\_Busan\_median\_nocd\_g, file = "Table1\_1\_Busan\_median\_nocd\_g.csv")

###################################################################################################

# Table2 Seoul로만 해야함

###################################################################################################

mk\_h1\_digits = function(x){

return (ifelse(!is.na(x),sprintf("%.2f",x),x))

}

mk\_h1\_digits\_p = function(x){

return (ifelse(!is.na(x),sprintf("%.3f",x),x))

}

data$W = 1

mk\_HR\_data = function(value,time,event,data){

Entire = data

Entire$group = factor(Entire[,value])

Entire$time = Entire[,time]

Entire$event = Entire[,event]

e1 = aggregate(W~group,data = Entire ,FUN = "sum")

e2 = aggregate(event~group,data = Entire ,FUN = "sum")

e3 = aggregate(time~group,data = Entire ,FUN = "sum")

e\_f = merge(merge(e1,e2),e3)

e\_f$PY = e\_f$time / 365

e\_f$rate = mk\_h1\_digits((e\_f$event / e\_f$W)\*100)

e\_f$IRR100 = mk\_h1\_digits((e\_f$event / e\_f$PY)\*100)

e\_f$event = paste0(e\_f$event, " (", e\_f$rate, "%)")

e\_f$group = paste0(value,e\_f$group)

e\_f = e\_f[,-c(4,6)]

model\_cox = coxph(Surv(time,event == 1) ~ group, data = Entire)

k = summary(model\_cox)

ggg = unique(Entire$group[!is.na(Entire$group)])

if(length(ggg)==2){

tmp2 = k$conf.int[,-2] %>% t() %>% round(2) %>% data.frame()

}else{

tmp2 = k$conf.int[,-2]%>% round(2)%>% data.frame()

}

tmp2$exp.coef. = mk\_h1\_digits(tmp2$exp.coef.)

tmp2$lower..95 = mk\_h1\_digits(tmp2$lower..95)

tmp2$upper..95 = mk\_h1\_digits(tmp2$upper..95)

tmp2$group = sort(ggg)[2:length(ggg)]

tmp2$group = paste0(value,tmp2$group)

tmp2$HR = paste0(tmp2$exp.coef.," (",tmp2$lower..95,"-",tmp2$upper..95,")")

tmp2$p = k$coefficients[,ncol(k$coefficients)] %>% round(3)

tmp2$p = ifelse(tmp2$p <0.001,"<0.001", mk\_h1\_digits\_p(tmp2$p))

tmp3 = tmp2[c("group","HR","p")]

final = merge(e\_f,tmp3,by = "group", all.x = T)

#final$HR[1] = "1(Ref.)"

#final$p[1] = ""

return(final)

}

######################################

# Table2 Total

######################################

Seoul = data %>% filter(mx\_Seoul == 1)

Seoul$bmi\_g2 = factor(Seoul$bmi\_g, levels = c(2,1,3))

dim(Seoul)

var\_list = c("Arthritis", "Asthma" , "Cancer", "metastatic", "cardiac\_arrhythmias", "kidney",

"pulmonary", "hf", "Coronary", "Dementia", "Depression", "Diabetes", "Hepatitis", "Dyslipidemia",

"HIV" , "HTN", "Osteoporosis", "Schizophrenia", "Stroke", "SAD",

"SEX\_TYPE","smk\_g", "drk\_g","pa\_g","low\_incm","bmi\_g2", "mx\_ER", "admin\_day\_g","treatment",

"age\_g")

all = data.frame()

for(i in var\_list){

tmp = mk\_HR\_data(data = Seoul, value = i, time = "dth\_fu", event = "dth")

all = rbind(all,tmp)

}

all

write.csv(all,"all.csv",row.names = F)

######################################

# Table2 Male

######################################

var\_list = c("Arthritis", "Asthma" , "Cancer", "metastatic", "cardiac\_arrhythmias", "kidney",

"pulmonary", "hf", "Coronary", "Dementia", "Depression", "Diabetes", "Hepatitis", "Dyslipidemia",

"HIV" , "HTN", "Osteoporosis", "Schizophrenia", "Stroke", "SAD",

"smk\_g", "drk\_g","pa\_g","low\_incm","bmi\_g2", "mx\_ER", "admin\_day\_g","treatment",

"age\_g")

male = Seoul %>% filter(SEX\_TYPE == "1")

all\_male = data.frame()

for(i in var\_list){

tmp = mk\_HR\_data(data = male, value = i, time = "dth\_fu", event = "dth")

all\_male = rbind(all\_male,tmp)

}

all\_male

write.csv(all\_male,"all\_male.csv",row.names = F)

######################################

# Table2 Female

######################################

female = Seoul %>% filter(SEX\_TYPE == "2")

all\_female = data.frame()

var\_list = c("Arthritis", "Asthma" , "Cancer", "metastatic", "cardiac\_arrhythmias", "kidney",

"pulmonary", "hf", "Coronary", "Dementia", "Depression", "Diabetes", "Hepatitis", "Dyslipidemia",

"HIV" , "HTN", "Osteoporosis", "Schizophrenia", "Stroke", "SAD",

"smk\_g", "drk\_g","pa\_g","low\_incm","bmi\_g2", "mx\_ER", "admin\_day\_g","treatment",

"age\_g")

for(i in var\_list){

tmp = mk\_HR\_data(data = female, value = i, time = "dth\_fu", event = "dth")

all\_female = rbind(all\_female,tmp)

#cat(i,"\n")

}

all\_female

write.csv(all\_female,"all\_female.csv",row.names = F)

###########################################

# Model1\_12 - Total

###########################################

Entire = data %>% filter(mx\_Seoul == 1); time = "dth\_fu"; event = "dth"

Entire$bmi\_g2 = factor(Entire$bmi\_g, levels = c(2,1,3))

Entire$time = Entire[,time]

Entire$event = Entire[,event]

Entire$admin\_day\_g = factor(Entire$admin\_day\_g)

Entire <- Entire %>%

select(time,event,Arthritis, Asthma, Cancer, metastatic,

cardiac\_arrhythmias, kidney, pulmonary, hf, Coronary, Dementia, Depression,

Diabetes, Hepatitis, HTN, Schizophrenia,

Stroke, SAD, SEX\_TYPE, smk\_g,

drk\_g,pa\_g,low\_incm,bmi\_g2,mx\_ER,admin\_day\_g,treatment,age\_g)

model\_cox = coxph(Surv(time,event == 1) ~ Arthritis + Asthma + Cancer + metastatic +

cardiac\_arrhythmias + kidney + pulmonary + hf + Coronary + Dementia + Depression +

Diabetes + Hepatitis + HTN + Schizophrenia + Stroke + SAD + SEX\_TYPE + smk\_g +

drk\_g +pa\_g +low\_incm +bmi\_g2 +mx\_ER +admin\_day\_g +treatment +age\_g,

data = Entire)

k = summary(model\_cox)

tmp2 = k$conf.int[,-2]%>% round(2)%>% data.frame()

tmp2$exp.coef. = mk\_h1\_digits(tmp2$exp.coef.)

tmp2$lower..95 = mk\_h1\_digits(tmp2$lower..95)

tmp2$upper..95 = mk\_h1\_digits(tmp2$upper..95)

tmp2$HR = paste0(tmp2$exp.coef.," (",tmp2$lower..95,"-",tmp2$upper..95,")")

tmp2$p = k$coefficients[,ncol(k$coefficients)] %>% round(3)

tmp2$p = ifelse(tmp2$p <0.001,"<0.001", mk\_h1\_digits\_p(tmp2$p))

tmp3 = tmp2[c("HR","p")]

write.csv(tmp3, "Model1\_1\_total.csv")

model\_cox = coxph(Surv(time,event == 1) ~ Arthritis + Asthma + Cancer + metastatic +

cardiac\_arrhythmias + kidney + pulmonary + hf + Coronary + Dementia + Depression +

Diabetes + Hepatitis + HTN + Schizophrenia + Stroke + SAD + SEX\_TYPE + smk\_g +

drk\_g +pa\_g +low\_incm +bmi\_g2 +mx\_ER +treatment +age\_g,

data = Entire)

k = summary(model\_cox)

tmp2 = k$conf.int[,-2]%>% round(2)%>% data.frame()

tmp2$exp.coef. = mk\_h1\_digits(tmp2$exp.coef.)

tmp2$lower..95 = mk\_h1\_digits(tmp2$lower..95)

tmp2$upper..95 = mk\_h1\_digits(tmp2$upper..95)

tmp2$HR = paste0(tmp2$exp.coef.," (",tmp2$lower..95,"-",tmp2$upper..95,")")

tmp2$p = k$coefficients[,ncol(k$coefficients)] %>% round(3)

tmp2$p = ifelse(tmp2$p <0.001,"<0.001", mk\_h1\_digits\_p(tmp2$p))

tmp3 = tmp2[c("HR","p")]

write.csv(tmp3, "Model1\_2\_total.csv")

model\_cox = coxph(Surv(time,event == 1) ~ Arthritis + Asthma + Cancer + metastatic +

cardiac\_arrhythmias + kidney + hf + Coronary + Dementia + Depression +

Diabetes + Hepatitis + HTN + Stroke + SAD + SEX\_TYPE + smk\_g +

drk\_g +pa\_g +low\_incm +bmi\_g2 +mx\_ER +treatment +age\_g,

data = Entire)

k = summary(model\_cox)

tmp2 = k$conf.int[,-2]%>% round(2)%>% data.frame()

tmp2$exp.coef. = mk\_h1\_digits(tmp2$exp.coef.)

tmp2$lower..95 = mk\_h1\_digits(tmp2$lower..95)

tmp2$upper..95 = mk\_h1\_digits(tmp2$upper..95)

tmp2$HR = paste0(tmp2$exp.coef.," (",tmp2$lower..95,"-",tmp2$upper..95,")")

tmp2$p = k$coefficients[,ncol(k$coefficients)] %>% round(3)

tmp2$p = ifelse(tmp2$p <0.001,"<0.001", mk\_h1\_digits\_p(tmp2$p))

tmp3 = tmp2[c("HR","p")]

write.csv(tmp3, "Model1\_3\_total.csv")

###########################################

# Model1\_12 -Male

###########################################

Entire = data %>% filter(mx\_Seoul == 1 & SEX\_TYPE == "1"); time = "dth\_fu"; event = "dth"

Entire$bmi\_g2 = factor(Entire$bmi\_g, levels = c(2,1,3))

Entire$time = Entire[,time]

Entire$event = Entire[,event]

Entire$admin\_day\_g = factor(Entire$admin\_day\_g)

names(Entire)

Entire <- Entire %>%

select(time,event,Arthritis, Asthma, Cancer, metastatic,

cardiac\_arrhythmias, kidney, pulmonary, hf, Coronary, Dementia, Depression,

Diabetes, Hepatitis, HTN, Osteoporosis, Schizophrenia,

Stroke, SAD, SEX\_TYPE, smk\_g,

drk\_g,pa\_g,low\_incm,bmi\_g2,mx\_ER,admin\_day\_g,treatment,age\_g)

model\_cox = coxph(Surv(time,event == 1) ~ Arthritis + Asthma + Cancer + metastatic +

cardiac\_arrhythmias + kidney + pulmonary + hf + Coronary + Dementia + Depression +

Diabetes + Hepatitis + HTN + Osteoporosis + Schizophrenia + Stroke + SAD + smk\_g +

drk\_g +pa\_g +low\_incm +bmi\_g2 +mx\_ER +admin\_day\_g +treatment +age\_g,

data = Entire)

k = summary(model\_cox)

tmp2 = k$conf.int[,-2]%>% round(2)%>% data.frame()

tmp2$exp.coef. = mk\_h1\_digits(tmp2$exp.coef.)

tmp2$lower..95 = mk\_h1\_digits(tmp2$lower..95)

tmp2$upper..95 = mk\_h1\_digits(tmp2$upper..95)

tmp2$HR = paste0(tmp2$exp.coef.," (",tmp2$lower..95,"-",tmp2$upper..95,")")

tmp2$p = k$coefficients[,ncol(k$coefficients)] %>% round(3)

tmp2$p = ifelse(tmp2$p <0.001,"<0.001", mk\_h1\_digits\_p(tmp2$p))

tmp3 = tmp2[c("HR","p")]

write.csv(tmp3, "Model1\_1\_male.csv")

model\_cox = coxph(Surv(time,event == 1) ~ Arthritis + Asthma + Cancer + metastatic +

cardiac\_arrhythmias + kidney + pulmonary + hf + Coronary + Dementia + Depression +

Diabetes + Hepatitis + HTN + Osteoporosis + Schizophrenia + Stroke + SAD + smk\_g +

drk\_g +pa\_g +low\_incm +bmi\_g2 +mx\_ER +treatment +age\_g,

data = Entire)

k = summary(model\_cox)

tmp2 = k$conf.int[,-2]%>% round(2)%>% data.frame()

tmp2$exp.coef. = mk\_h1\_digits(tmp2$exp.coef.)

tmp2$lower..95 = mk\_h1\_digits(tmp2$lower..95)

tmp2$upper..95 = mk\_h1\_digits(tmp2$upper..95)

tmp2$HR = paste0(tmp2$exp.coef.," (",tmp2$lower..95,"-",tmp2$upper..95,")")

tmp2$p = k$coefficients[,ncol(k$coefficients)] %>% round(3)

tmp2$p = ifelse(tmp2$p <0.001,"<0.001", mk\_h1\_digits\_p(tmp2$p))

tmp3 = tmp2[c("HR","p")]

write.csv(tmp3, "Model1\_2\_male.csv")

model\_cox = coxph(Surv(time,event == 1) ~ Arthritis + Asthma + Cancer + metastatic +

cardiac\_arrhythmias + kidney + hf + Coronary + Dementia + Depression +

Hepatitis + HTN + Osteoporosis + Stroke + SAD + smk\_g +

drk\_g +pa\_g +low\_incm +bmi\_g2 +mx\_ER +treatment +age\_g,

data = Entire)

k = summary(model\_cox)

tmp2 = k$conf.int[,-2]%>% round(2)%>% data.frame()

tmp2$exp.coef. = mk\_h1\_digits(tmp2$exp.coef.)

tmp2$lower..95 = mk\_h1\_digits(tmp2$lower..95)

tmp2$upper..95 = mk\_h1\_digits(tmp2$upper..95)

tmp2$HR = paste0(tmp2$exp.coef.," (",tmp2$lower..95,"-",tmp2$upper..95,")")

tmp2$p = k$coefficients[,ncol(k$coefficients)] %>% round(3)

tmp2$p = ifelse(tmp2$p <0.001,"<0.001", mk\_h1\_digits\_p(tmp2$p))

tmp3 = tmp2[c("HR","p")]

write.csv(tmp3, "Model1\_3\_male.csv")

###########################################

# Model1\_1 -Female

###########################################

Entire = data %>% filter(mx\_Seoul == 1 & SEX\_TYPE == "2"); time = "dth\_fu"; event = "dth"

Entire$bmi\_g2 = factor(Entire$bmi\_g, levels = c(2,1,3))

Entire$time = Entire[,time]

Entire$event = Entire[,event]

Entire$admin\_day\_g = factor(Entire$admin\_day\_g)

names(Entire)

Entire <- Entire %>%

select(time,event,Arthritis, Asthma, Cancer, metastatic,

cardiac\_arrhythmias, kidney, pulmonary, hf, Coronary, Dementia, Depression,

Diabetes, Hepatitis, HTN, Osteoporosis, Schizophrenia,

Stroke, SEX\_TYPE, smk\_g,

drk\_g,pa\_g,low\_incm,bmi\_g2,mx\_ER,admin\_day\_g,treatment,age\_g)

model\_cox = coxph(Surv(time,event == 1) ~ Arthritis + Asthma + Cancer + metastatic +

cardiac\_arrhythmias + kidney + pulmonary + hf + Coronary + Dementia + Depression +

Diabetes + Hepatitis + HTN + Osteoporosis + Schizophrenia + Stroke + smk\_g +

drk\_g +pa\_g +low\_incm +bmi\_g2 +mx\_ER +admin\_day\_g +treatment +age\_g,

data = Entire)

k = summary(model\_cox)

tmp2 = k$conf.int[,-2]%>% round(2)%>% data.frame()

tmp2$exp.coef. = mk\_h1\_digits(tmp2$exp.coef.)

tmp2$lower..95 = mk\_h1\_digits(tmp2$lower..95)

tmp2$upper..95 = mk\_h1\_digits(tmp2$upper..95)

tmp2$HR = paste0(tmp2$exp.coef.," (",tmp2$lower..95,"-",tmp2$upper..95,")")

tmp2$p = k$coefficients[,ncol(k$coefficients)] %>% round(3)

tmp2$p = ifelse(tmp2$p <0.001,"<0.001", mk\_h1\_digits\_p(tmp2$p))

tmp3 = tmp2[c("HR","p")]

write.csv(tmp3, "Model1\_1\_Female.csv")

model\_cox = coxph(Surv(time,event == 1) ~ Arthritis + Asthma + Cancer + metastatic +

cardiac\_arrhythmias + kidney + pulmonary + hf + Coronary + Dementia + Depression +

Diabetes + Hepatitis + HTN + Osteoporosis + Schizophrenia + Stroke + smk\_g +

drk\_g +pa\_g +low\_incm +bmi\_g2 +mx\_ER +treatment +age\_g,

data = Entire)

k = summary(model\_cox)

tmp2 = k$conf.int[,-2]%>% round(2)%>% data.frame()

tmp2$exp.coef. = mk\_h1\_digits(tmp2$exp.coef.)

tmp2$lower..95 = mk\_h1\_digits(tmp2$lower..95)

tmp2$upper..95 = mk\_h1\_digits(tmp2$upper..95)

tmp2$HR = paste0(tmp2$exp.coef.," (",tmp2$lower..95,"-",tmp2$upper..95,")")

tmp2$p = k$coefficients[,ncol(k$coefficients)] %>% round(3)

tmp2$p = ifelse(tmp2$p <0.001,"<0.001", mk\_h1\_digits\_p(tmp2$p))

tmp3 = tmp2[c("HR","p")]

write.csv(tmp3, "Model1\_2\_Female.csv")

model\_cox = coxph(Surv(time,event == 1) ~ Arthritis + Cancer + metastatic +

kidney + hf + Coronary + Dementia +

Diabetes + Hepatitis + HTN + Stroke + smk\_g +

pa\_g +low\_incm +bmi\_g2 +mx\_ER +treatment +age\_g,

data = Entire)

k = summary(model\_cox)

tmp2 = k$conf.int[,-2]%>% round(2)%>% data.frame()

tmp2$exp.coef. = mk\_h1\_digits(tmp2$exp.coef.)

tmp2$lower..95 = mk\_h1\_digits(tmp2$lower..95)

tmp2$upper..95 = mk\_h1\_digits(tmp2$upper..95)

tmp2$HR = paste0(tmp2$exp.coef.," (",tmp2$lower..95,"-",tmp2$upper..95,")")

tmp2$p = k$coefficients[,ncol(k$coefficients)] %>% round(3)

tmp2$p = ifelse(tmp2$p <0.001,"<0.001", mk\_h1\_digits\_p(tmp2$p))

tmp3 = tmp2[c("HR","p")]

write.csv(tmp3, "Model1\_3\_Female.csv")

#################################################################################################################################

# Table3

#################################################################################################################################

###########################################

# 만성질환만 따로하기

###########################################

Seoul = data %>% filter(mx\_Seoul == 1)

male = Seoul %>% filter(SEX\_TYPE == "1")

female = Seoul %>% filter(SEX\_TYPE == "2")

Total\_nocd = mk\_HR\_data(data = Seoul, value = "nocd\_g", time = "dth\_fu", event = "dth")

write.csv(Total\_nocd, "Total\_nocd.csv")

female\_nocd = mk\_HR\_data(data = female, value = "nocd\_g", time = "dth\_fu", event = "dth")

write.csv(female\_nocd, "female\_nocd.csv")

male\_nocd = mk\_HR\_data(data = male, value = "nocd\_g", time = "dth\_fu", event = "dth")

write.csv(male\_nocd, "male\_nocd.csv")

############################

#P for trend -Model1

############################

Entire = Seoul

Entire$time = Entire$dth\_fu

Entire$event = Entire$dth

Entire$group = Entire$nocd\_g %>% as.numeric()

model\_cox = coxph(Surv(time,event == 1) ~ group, data = Entire)

k = summary(model\_cox)

k

Entire = male

Entire$time = Entire$dth\_fu

Entire$event = Entire$dth

Entire$group = Entire$nocd\_g %>% as.numeric()

model\_cox = coxph(Surv(time,event == 1) ~ group, data = Entire)

k = summary(model\_cox)

k

Entire = female

Entire$time = Entire$dth\_fu

Entire$event = Entire$dth

Entire$group = Entire$nocd\_g %>% as.numeric()

model\_cox = coxph(Surv(time,event == 1) ~ group, data = Entire)

k = summary(model\_cox)

k

###########################################

# Model2\_12 -Total

###########################################

Entire = data %>% filter(mx\_Seoul == 1); time = "dth\_fu"; event = "dth"

Entire$bmi\_g2 = factor(Entire$bmi\_g, levels = c(2,1,3))

Entire$time = Entire[,time]

Entire$event = Entire[,event]

Entire$admin\_day\_g = factor(Entire$admin\_day\_g)

names(Entire)

Entire <- Entire %>%

select(time,event,nocd\_g,SEX\_TYPE, smk\_g,

drk\_g,pa\_g,low\_incm,bmi\_g2,mx\_ER,admin\_day\_g,treatment,age\_g)

###############

#Model2\_1 P for trend

###############

Entire$group = Entire$nocd\_g %>% as.numeric()

model\_cox = coxph(Surv(time,event == 1) ~ group + SEX\_TYPE + smk\_g +

drk\_g +pa\_g +low\_incm +bmi\_g2 +mx\_ER +admin\_day\_g +treatment +age\_g,

data = Entire)

k = summary(model\_cox)

k

model\_cox = coxph(Surv(time,event == 1) ~ nocd\_g + SEX\_TYPE + smk\_g +

drk\_g +pa\_g +low\_incm +bmi\_g2 +mx\_ER +admin\_day\_g +treatment +age\_g,

data = Entire)

k = summary(model\_cox)

tmp2 = k$conf.int[,-2]%>% round(2)%>% data.frame()

tmp2$exp.coef. = mk\_h1\_digits(tmp2$exp.coef.)

tmp2$lower..95 = mk\_h1\_digits(tmp2$lower..95)

tmp2$upper..95 = mk\_h1\_digits(tmp2$upper..95)

tmp2$HR = paste0(tmp2$exp.coef.," (",tmp2$lower..95,"-",tmp2$upper..95,")")

tmp2$p = k$coefficients[,ncol(k$coefficients)] %>% round(3)

tmp2$p = ifelse(tmp2$p <0.001,"<0.001", mk\_h1\_digits\_p(tmp2$p))

tmp3 = tmp2[c("HR","p")]

write.csv(tmp3, "Model2\_1\_Total.csv")

###############

#Model2\_2 P for trend

###############

model\_cox = coxph(Surv(time,event == 1) ~ group + SEX\_TYPE + smk\_g +

drk\_g +pa\_g +low\_incm +bmi\_g2 +mx\_ER +treatment +age\_g,

data = Entire)

k = summary(model\_cox)

k

model\_cox = coxph(Surv(time,event == 1) ~ nocd\_g + SEX\_TYPE + smk\_g +

drk\_g +pa\_g +low\_incm +bmi\_g2 +mx\_ER +treatment +age\_g,

data = Entire)

k = summary(model\_cox)

tmp2 = k$conf.int[,-2]%>% round(2)%>% data.frame()

tmp2$exp.coef. = mk\_h1\_digits(tmp2$exp.coef.)

tmp2$lower..95 = mk\_h1\_digits(tmp2$lower..95)

tmp2$upper..95 = mk\_h1\_digits(tmp2$upper..95)

tmp2$HR = paste0(tmp2$exp.coef.," (",tmp2$lower..95,"-",tmp2$upper..95,")")

tmp2$p = k$coefficients[,ncol(k$coefficients)] %>% round(3)

tmp2$p = ifelse(tmp2$p <0.001,"<0.001", mk\_h1\_digits\_p(tmp2$p))

tmp3 = tmp2[c("HR","p")]

write.csv(tmp3, "Model2\_2\_Total.csv")

###########################################

# Model2\_12 -Male

###########################################

Entire = data %>% filter(mx\_Seoul == 1 & SEX\_TYPE == 1); time = "dth\_fu"; event = "dth"

Entire$bmi\_g2 = factor(Entire$bmi\_g, levels = c(2,1,3))

Entire$time = Entire[,time]

Entire$event = Entire[,event]

Entire$admin\_day\_g = factor(Entire$admin\_day\_g)

names(Entire)

Entire <- Entire %>%

select(time,event,nocd\_g,SEX\_TYPE, smk\_g,

drk\_g,pa\_g,low\_incm,bmi\_g2,mx\_ER,admin\_day\_g,treatment,age\_g)

###############

#Male Model2-1 P for trend

###############

Entire$group = Entire$nocd\_g %>% as.numeric()

model\_cox = coxph(Surv(time,event == 1) ~ group + smk\_g +

drk\_g +pa\_g +low\_incm +bmi\_g2 +mx\_ER +admin\_day\_g +treatment +age\_g,

data = Entire)

k = summary(model\_cox)

k

model\_cox = coxph(Surv(time,event == 1) ~ nocd\_g + smk\_g +

drk\_g +pa\_g +low\_incm +bmi\_g2 +mx\_ER +admin\_day\_g +treatment +age\_g,

data = Entire)

k = summary(model\_cox)

tmp2 = k$conf.int[,-2]%>% round(2)%>% data.frame()

tmp2$exp.coef. = mk\_h1\_digits(tmp2$exp.coef.)

tmp2$lower..95 = mk\_h1\_digits(tmp2$lower..95)

tmp2$upper..95 = mk\_h1\_digits(tmp2$upper..95)

tmp2$HR = paste0(tmp2$exp.coef.," (",tmp2$lower..95,"-",tmp2$upper..95,")")

tmp2$p = k$coefficients[,ncol(k$coefficients)] %>% round(3)

tmp2$p = ifelse(tmp2$p <0.001,"<0.001", mk\_h1\_digits\_p(tmp2$p))

tmp3 = tmp2[c("HR","p")]

write.csv(tmp3, "Model2\_1\_Male.csv")

###############

#Male Model2-2 P for trend

###############

Entire$group = Entire$nocd\_g %>% as.numeric()

model\_cox = coxph(Surv(time,event == 1) ~ group + smk\_g +

drk\_g +pa\_g +low\_incm +bmi\_g2 +mx\_ER +treatment +age\_g,

data = Entire)

k = summary(model\_cox)

k

model\_cox = coxph(Surv(time,event == 1) ~ nocd\_g + smk\_g +

drk\_g +pa\_g +low\_incm +bmi\_g2 +mx\_ER +treatment +age\_g,

data = Entire)

k = summary(model\_cox)

tmp2 = k$conf.int[,-2]%>% round(2)%>% data.frame()

tmp2$exp.coef. = mk\_h1\_digits(tmp2$exp.coef.)

tmp2$lower..95 = mk\_h1\_digits(tmp2$lower..95)

tmp2$upper..95 = mk\_h1\_digits(tmp2$upper..95)

tmp2$HR = paste0(tmp2$exp.coef.," (",tmp2$lower..95,"-",tmp2$upper..95,")")

tmp2$p = k$coefficients[,ncol(k$coefficients)] %>% round(3)

tmp2$p = ifelse(tmp2$p <0.001,"<0.001", mk\_h1\_digits\_p(tmp2$p))

tmp3 = tmp2[c("HR","p")]

write.csv(tmp3, "Model2\_2\_Male.csv")

###########################################

# Model2\_12 -Female

###########################################

Entire = data %>% filter(mx\_Seoul == 1 & SEX\_TYPE == 2); time = "dth\_fu"; event = "dth"

Entire$bmi\_g2 = factor(Entire$bmi\_g, levels = c(2,1,3))

Entire$time = Entire[,time]

Entire$event = Entire[,event]

Entire$admin\_day\_g = factor(Entire$admin\_day\_g)

names(Entire)

Entire <- Entire %>%

select(time,event,nocd\_g,SEX\_TYPE, smk\_g,

drk\_g,pa\_g,low\_incm,bmi\_g2,mx\_ER,admin\_day\_g,treatment,age\_g)

###############

#Female Model2-1 P for trend

###############

Entire$group = Entire$nocd\_g %>% as.numeric()

model\_cox = coxph(Surv(time,event == 1) ~ group + smk\_g +

drk\_g +pa\_g +low\_incm +bmi\_g2 +mx\_ER +admin\_day\_g +treatment +age\_g,

data = Entire)

k = summary(model\_cox)

k

model\_cox = coxph(Surv(time,event == 1) ~ nocd\_g + smk\_g +

drk\_g +pa\_g +low\_incm +bmi\_g2 +mx\_ER +admin\_day\_g +treatment +age\_g,

data = Entire)

k = summary(model\_cox)

tmp2 = k$conf.int[,-2]%>% round(2)%>% data.frame()

tmp2$exp.coef. = mk\_h1\_digits(tmp2$exp.coef.)

tmp2$lower..95 = mk\_h1\_digits(tmp2$lower..95)

tmp2$upper..95 = mk\_h1\_digits(tmp2$upper..95)

tmp2$HR = paste0(tmp2$exp.coef.," (",tmp2$lower..95,"-",tmp2$upper..95,")")

tmp2$p = k$coefficients[,ncol(k$coefficients)] %>% round(3)

tmp2$p = ifelse(tmp2$p <0.001,"<0.001", mk\_h1\_digits\_p(tmp2$p))

tmp3 = tmp2[c("HR","p")]

write.csv(tmp3, "Model2\_1\_Female.csv")

###############

#Female Model2-2 P for trend

###############

Entire$group = Entire$nocd\_g %>% as.numeric()

model\_cox = coxph(Surv(time,event == 1) ~ group + smk\_g +

drk\_g +pa\_g +low\_incm +bmi\_g2 +mx\_ER +treatment +age\_g,

data = Entire)

k = summary(model\_cox)

k

model\_cox = coxph(Surv(time,event == 1) ~ nocd\_g + smk\_g +

drk\_g +pa\_g +low\_incm +bmi\_g2 +mx\_ER +treatment +age\_g,

data = Entire)

k = summary(model\_cox)

tmp2 = k$conf.int[,-2]%>% round(2)%>% data.frame()

tmp2$exp.coef. = mk\_h1\_digits(tmp2$exp.coef.)

tmp2$lower..95 = mk\_h1\_digits(tmp2$lower..95)

tmp2$upper..95 = mk\_h1\_digits(tmp2$upper..95)

tmp2$HR = paste0(tmp2$exp.coef.," (",tmp2$lower..95,"-",tmp2$upper..95,")")

tmp2$p = k$coefficients[,ncol(k$coefficients)] %>% round(3)

tmp2$p = ifelse(tmp2$p <0.001,"<0.001", mk\_h1\_digits\_p(tmp2$p))

tmp3 = tmp2[c("HR","p")]

write.csv(tmp3, "Model2\_2\_Female.csv")

#############################################################################

#Beta값구하기

##############################################################################

#############################################################

# Table2 beta값

#############################################################

Entire = data %>% filter(mx\_Seoul == 1)

Entire$bmi\_g2 = factor(Entire$bmi\_g, levels = c(2,1,3))

Entire$admin\_day\_g = factor(Entire$admin\_day\_g)

model\_glm = glm(dth ~ Arthritis + Asthma + Cancer + metastatic +

cardiac\_arrhythmias + kidney + pulmonary + hf + Coronary + Dementia + Depression +

Diabetes + Hepatitis + HTN + Schizophrenia + Stroke + SAD + SEX\_TYPE + smk\_g +

drk\_g +pa\_g +low\_incm +bmi\_g2 +mx\_ER +admin\_day\_g +treatment +age\_g,

data = Entire)

k = summary(model\_glm)

df = k$coefficients[-1,1] %>% data.frame()

names(df)[1] = "beta"

df$point = round(df$beta / min(abs(df$beta)))

df$name\_before = lag(rownames(df))

ttt = data.frame(beta = "1(Ref.)", point = "", name\_before = "")

for(i in 1:nrow(df)){

tmp = df[i,]

tmp$beta = round(tmp$beta,4)

if(i == 1) all = tmp

else if(gsub("[0-9]","",df$name\_before[i]) != gsub("[0-9]","",rownames(df)[i])){

all = rbind(all,ttt)

all = rbind(all,tmp)

}else{

all = rbind(all,tmp)

}

}

write.csv(all,"Table2\_beta\_model1\_1.csv")

model\_glm = glm(dth ~ Arthritis + Asthma + Cancer + metastatic +

cardiac\_arrhythmias + kidney + pulmonary + hf + Coronary + Dementia + Depression +

Diabetes + Hepatitis + HTN + Schizophrenia + Stroke + SAD + SEX\_TYPE + smk\_g +

drk\_g +pa\_g +low\_incm +bmi\_g2 +mx\_ER +treatment +age\_g,

data = Entire)

k = summary(model\_glm)

df = k$coefficients[-1,1] %>% data.frame()

names(df)[1] = "beta"

df$point = round(df$beta / min(abs(df$beta)))

df$name\_before = lag(rownames(df))

ttt = data.frame(beta = "1(Ref.)", point = "", name\_before = "")

for(i in 1:nrow(df)){

tmp = df[i,]

tmp$beta = round(tmp$beta,4)

if(i == 1) all = tmp

else if(gsub("[0-9]","",df$name\_before[i]) != gsub("[0-9]","",rownames(df)[i])){

all = rbind(all,ttt)

all = rbind(all,tmp)

}else{

all = rbind(all,tmp)

}

}

write.csv(all,"Table2\_beta\_model1\_2.csv")

model\_glm = glm(dth ~ Arthritis + Asthma + Cancer + metastatic +

cardiac\_arrhythmias + kidney + pulmonary + hf + Dementia + Depression +

Diabetes + Hepatitis + HTN + Stroke + SAD + SEX\_TYPE + smk\_g +

drk\_g +pa\_g +low\_incm +bmi\_g2 +mx\_ER +treatment +age\_g,

data = Entire)

k = summary(model\_glm)

df = k$coefficients[-1,1] %>% data.frame()

names(df)[1] = "beta"

df$point = round(df$beta / min(abs(df$beta)))

df$name\_before = lag(rownames(df))

ttt = data.frame(beta = "1(Ref.)", point = "", name\_before = "")

for(i in 1:nrow(df)){

tmp = df[i,]

tmp$beta = round(tmp$beta,4)

if(i == 1) all = tmp

else if(gsub("[0-9]","",df$name\_before[i]) != gsub("[0-9]","",rownames(df)[i])){

all = rbind(all,ttt)

all = rbind(all,tmp)

}else{

all = rbind(all,tmp)

}

}

write.csv(all,"Table2\_beta\_model1\_3.csv")

#############################################################

# Male beta값

#############################################################

Entire = data %>% filter(mx\_Seoul == 1 & SEX\_TYPE == 1)

Entire$bmi\_g2 = factor(Entire$bmi\_g, levels = c(2,1,3))

Entire$admin\_day\_g = factor(Entire$admin\_day\_g)

model\_glm = glm(dth ~ Arthritis + Asthma + Cancer + metastatic +

cardiac\_arrhythmias + kidney + pulmonary + hf + Coronary + Dementia + Depression +

Diabetes + Hepatitis + HTN + Osteoporosis + Schizophrenia + Stroke + SAD + smk\_g +

drk\_g +pa\_g +low\_incm +bmi\_g2 +mx\_ER +admin\_day\_g +treatment +age\_g,

data = Entire)

k = summary(model\_glm)

df = k$coefficients[-1,1] %>% data.frame()

names(df)[1] = "beta"

df$point = round(df$beta / min(abs(df$beta)))

df$name\_before = lag(rownames(df))

ttt = data.frame(beta = "1(Ref.)", point = "", name\_before = "")

for(i in 1:nrow(df)){

tmp = df[i,]

tmp$beta = round(tmp$beta,4)

if(i == 1) all = tmp

else if(gsub("[0-9]","",df$name\_before[i]) != gsub("[0-9]","",rownames(df)[i])){

all = rbind(all,ttt)

all = rbind(all,tmp)

}else{

all = rbind(all,tmp)

}

}

write.csv(all,"Male\_beta\_model1\_1.csv")

model\_glm = glm(dth ~ Arthritis + Asthma + Cancer + metastatic +

cardiac\_arrhythmias + kidney + pulmonary + hf + Coronary + Dementia + Depression +

Diabetes + Hepatitis + HTN + Osteoporosis + Schizophrenia + Stroke + SAD + smk\_g +

drk\_g +pa\_g +low\_incm +bmi\_g2 +mx\_ER +treatment +age\_g,

data = Entire)

k = summary(model\_glm)

df = k$coefficients[-1,1] %>% data.frame()

names(df)[1] = "beta"

df$point = round(df$beta / min(abs(df$beta)))

df$name\_before = lag(rownames(df))

ttt = data.frame(beta = "1(Ref.)", point = "", name\_before = "")

for(i in 1:nrow(df)){

tmp = df[i,]

tmp$beta = round(tmp$beta,4)

if(i == 1) all = tmp

else if(gsub("[0-9]","",df$name\_before[i]) != gsub("[0-9]","",rownames(df)[i])){

all = rbind(all,ttt)

all = rbind(all,tmp)

}else{

all = rbind(all,tmp)

}

}

write.csv(all,"Male\_beta\_model1\_2.csv")

model\_glm = glm(dth ~ Arthritis + Asthma + Cancer + metastatic +

cardiac\_arrhythmias + kidney + hf + Coronary + Dementia + Depression +

Hepatitis + HTN + Osteoporosis + Stroke + SAD + smk\_g +

drk\_g +pa\_g +low\_incm +bmi\_g2 +mx\_ER +treatment +age\_g,

data = Entire)

k = summary(model\_glm)

df = k$coefficients[-1,1] %>% data.frame()

names(df)[1] = "beta"

df$point = round(df$beta / min(abs(df$beta)))

df$name\_before = lag(rownames(df))

ttt = data.frame(beta = "1(Ref.)", point = "", name\_before = "")

for(i in 1:nrow(df)){

tmp = df[i,]

tmp$beta = round(tmp$beta,4)

if(i == 1) all = tmp

else if(gsub("[0-9]","",df$name\_before[i]) != gsub("[0-9]","",rownames(df)[i])){

all = rbind(all,ttt)

all = rbind(all,tmp)

}else{

all = rbind(all,tmp)

}

}

write.csv(all,"Male\_beta\_model1\_3.csv")

#############################################################

# Female beta값

#############################################################

Entire = data %>% filter(mx\_Seoul == 1 & SEX\_TYPE == 2)

Entire$bmi\_g2 = factor(Entire$bmi\_g, levels = c(2,1,3))

Entire$admin\_day\_g = factor(Entire$admin\_day\_g)

model\_glm = glm(dth ~ Arthritis + Asthma + Cancer + metastatic +

cardiac\_arrhythmias + kidney + pulmonary + hf + Coronary + Dementia + Depression +

Diabetes + Hepatitis + HTN + Osteoporosis + Schizophrenia + Stroke + smk\_g +

drk\_g +pa\_g +low\_incm +bmi\_g2 +mx\_ER +admin\_day\_g +treatment +age\_g,

data = Entire)

k = summary(model\_glm)

df = k$coefficients[-1,1] %>% data.frame()

names(df)[1] = "beta"

df$point = round(df$beta / min(abs(df$beta)))

df$name\_before = lag(rownames(df))

ttt = data.frame(beta = "1(Ref.)", point = "", name\_before = "")

for(i in 1:nrow(df)){

tmp = df[i,]

tmp$beta = round(tmp$beta,4)

if(i == 1) all = tmp

else if(gsub("[0-9]","",df$name\_before[i]) != gsub("[0-9]","",rownames(df)[i])){

all = rbind(all,ttt)

all = rbind(all,tmp)

}else{

all = rbind(all,tmp)

}

}

write.csv(all,"Female\_beta\_model1\_1.csv")

model\_glm = glm(dth ~ Arthritis + Asthma + Cancer + metastatic +

cardiac\_arrhythmias + kidney + pulmonary + hf + Coronary + Dementia + Depression +

Diabetes + Hepatitis + HTN + Osteoporosis + Schizophrenia + Stroke + smk\_g +

drk\_g +pa\_g +low\_incm +bmi\_g2 +mx\_ER +treatment +age\_g,

data = Entire)

k = summary(model\_glm)

df = k$coefficients[-1,1] %>% data.frame()

names(df)[1] = "beta"

df$point = round(df$beta / min(abs(df$beta)))

df$name\_before = lag(rownames(df))

ttt = data.frame(beta = "1(Ref.)", point = "", name\_before = "")

for(i in 1:nrow(df)){

tmp = df[i,]

tmp$beta = round(tmp$beta,4)

if(i == 1) all = tmp

else if(gsub("[0-9]","",df$name\_before[i]) != gsub("[0-9]","",rownames(df)[i])){

all = rbind(all,ttt)

all = rbind(all,tmp)

}else{

all = rbind(all,tmp)

}

}

write.csv(all,"Female\_beta\_model1\_2.csv")

model\_glm = glm(dth ~ Arthritis + Cancer + metastatic +

kidney + hf + Coronary + Dementia +

Diabetes + Hepatitis + HTN + Stroke + smk\_g +

pa\_g +low\_incm +bmi\_g2 +mx\_ER +treatment +age\_g,

data = Entire)

k = summary(model\_glm)

df = k$coefficients[-1,1] %>% data.frame()

names(df)[1] = "beta"

df$point = round(df$beta / min(abs(df$beta)))

df$name\_before = lag(rownames(df))

ttt = data.frame(beta = "1(Ref.)", point = "", name\_before = "")

for(i in 1:nrow(df)){

tmp = df[i,]

tmp$beta = round(tmp$beta,4)

if(i == 1) all = tmp

else if(gsub("[0-9]","",df$name\_before[i]) != gsub("[0-9]","",rownames(df)[i])){

all = rbind(all,ttt)

all = rbind(all,tmp)

}else{

all = rbind(all,tmp)

}

}

write.csv(all,"Female\_beta\_model1\_3.csv")

#############################################################

# Table3 beta값

#############################################################

Entire = data %>% filter(mx\_Seoul == 1)

Entire$bmi\_g2 = factor(Entire$bmi\_g, levels = c(2,1,3))

Entire$admin\_day\_g = factor(Entire$admin\_day\_g)

model\_glm = glm(dth ~ nocd\_g + SEX\_TYPE + smk\_g +

drk\_g +pa\_g +low\_incm +bmi\_g2 +mx\_ER +admin\_day\_g +treatment +age\_g,

data = Entire)

k = summary(model\_glm)

df = k$coefficients[-1,1] %>% data.frame()

names(df)[1] = "beta"

df$point = round(df$beta / min(abs(df$beta)))

df$name\_before = lag(rownames(df))

ttt = data.frame(beta = "1(Ref.)", point = "", name\_before = "")

for(i in 1:nrow(df)){

tmp = df[i,]

tmp$beta = round(tmp$beta,4)

if(i == 1) all = tmp

else if(gsub("[0-9]","",df$name\_before[i]) != gsub("[0-9]","",rownames(df)[i])){

all = rbind(all,ttt)

all = rbind(all,tmp)

}else{

all = rbind(all,tmp)

}

}

write.csv(all,"Table3\_beta\_model2\_1.csv")

model\_glm = glm(dth ~ nocd\_g + SEX\_TYPE + smk\_g +

drk\_g +pa\_g +low\_incm +bmi\_g2 +mx\_ER +treatment +age\_g,

data = Entire)

k = summary(model\_glm)

df = k$coefficients[-1,1] %>% data.frame()

names(df)[1] = "beta"

df$point = round(df$beta / min(abs(df$beta)))

df$name\_before = lag(rownames(df))

ttt = data.frame(beta = "1(Ref.)", point = "", name\_before = "")

for(i in 1:nrow(df)){

tmp = df[i,]

tmp$beta = round(tmp$beta,4)

if(i == 1) all = tmp

else if(gsub("[0-9]","",df$name\_before[i]) != gsub("[0-9]","",rownames(df)[i])){

all = rbind(all,ttt)

all = rbind(all,tmp)

}else{

all = rbind(all,tmp)

}

}

write.csv(all,"Table3\_beta\_model2\_2.csv")

#############################################################

# Male beta값

#############################################################

Entire = data %>% filter(mx\_Seoul == 1 & SEX\_TYPE == 1)

Entire$bmi\_g2 = factor(Entire$bmi\_g, levels = c(2,1,3))

Entire$admin\_day\_g = factor(Entire$admin\_day\_g)

model\_glm = glm(dth ~ nocd\_g + smk\_g +

drk\_g +pa\_g +low\_incm +bmi\_g2 +mx\_ER +admin\_day\_g +treatment +age\_g,

data = Entire)

k = summary(model\_glm)

df = k$coefficients[-1,1] %>% data.frame()

names(df)[1] = "beta"

df$point = round(df$beta / min(abs(df$beta)))

df$name\_before = lag(rownames(df))

ttt = data.frame(beta = "1(Ref.)", point = "", name\_before = "")

for(i in 1:nrow(df)){

tmp = df[i,]

tmp$beta = round(tmp$beta,4)

if(i == 1) all = tmp

else if(gsub("[0-9]","",df$name\_before[i]) != gsub("[0-9]","",rownames(df)[i])){

all = rbind(all,ttt)

all = rbind(all,tmp)

}else{

all = rbind(all,tmp)

}

}

write.csv(all,"Male\_beta\_model2\_1.csv")

model\_glm = glm(dth ~ nocd\_g + smk\_g +

drk\_g +pa\_g +low\_incm +bmi\_g2 +mx\_ER +treatment +age\_g,

data = Entire)

k = summary(model\_glm)

df = k$coefficients[-1,1] %>% data.frame()

names(df)[1] = "beta"

df$point = round(df$beta / min(abs(df$beta)))

df$name\_before = lag(rownames(df))

ttt = data.frame(beta = "1(Ref.)", point = "", name\_before = "")

for(i in 1:nrow(df)){

tmp = df[i,]

tmp$beta = round(tmp$beta,4)

if(i == 1) all = tmp

else if(gsub("[0-9]","",df$name\_before[i]) != gsub("[0-9]","",rownames(df)[i])){

all = rbind(all,ttt)

all = rbind(all,tmp)

}else{

all = rbind(all,tmp)

}

}

write.csv(all,"Male\_beta\_model2\_2.csv")

#############################################################

# Female beta값

#############################################################

Entire = data %>% filter(mx\_Seoul == 1 & SEX\_TYPE == 2)

Entire$bmi\_g2 = factor(Entire$bmi\_g, levels = c(2,1,3))

Entire$admin\_day\_g = factor(Entire$admin\_day\_g)

model\_glm = glm(dth ~ nocd\_g + smk\_g +

drk\_g +pa\_g +low\_incm +bmi\_g2 +mx\_ER +admin\_day\_g +treatment +age\_g,

data = Entire)

k = summary(model\_glm)

df = k$coefficients[-1,1] %>% data.frame()

names(df)[1] = "beta"

df$point = round(df$beta / min(abs(df$beta)))

df$name\_before = lag(rownames(df))

ttt = data.frame(beta = "1(Ref.)", point = "", name\_before = "")

for(i in 1:nrow(df)){

tmp = df[i,]

tmp$beta = round(tmp$beta,4)

if(i == 1) all = tmp

else if(gsub("[0-9]","",df$name\_before[i]) != gsub("[0-9]","",rownames(df)[i])){

all = rbind(all,ttt)

all = rbind(all,tmp)

}else{

all = rbind(all,tmp)

}

}

write.csv(all,"Female\_beta\_model2\_1.csv")

model\_glm = glm(dth ~ nocd\_g + smk\_g +

drk\_g +pa\_g +low\_incm +bmi\_g2 +mx\_ER +treatment +age\_g,

data = Entire)

k = summary(model\_glm)

df = k$coefficients[-1,1] %>% data.frame()

names(df)[1] = "beta"

df$point = round(df$beta / min(abs(df$beta)))

df$name\_before = lag(rownames(df))

ttt = data.frame(beta = "1(Ref.)", point = "", name\_before = "")

for(i in 1:nrow(df)){

tmp = df[i,]

tmp$beta = round(tmp$beta,4)

if(i == 1) all = tmp

else if(gsub("[0-9]","",df$name\_before[i]) != gsub("[0-9]","",rownames(df)[i])){

all = rbind(all,ttt)

all = rbind(all,tmp)

}else{

all = rbind(all,tmp)

}

}

write.csv(all,"Female\_beta\_model2\_2.csv")

##################################################################

#Table4

##################################################################

get\_tt <- function(data, event){

final = list()

set.seed(2022)

kkk = 5

res = c()

event\_rate = data.frame(matrix(rep(0,3\*kkk),3,kkk))

Kfold\_res = createFolds(Seoul$event, k = kkk,list=T, returnTrain = F)

val\_list = paste0("Fold",1:length(Kfold\_res))

data$event = data[,event]

for(i in 1:kkk){

glm\_stat <- glm(event ~ risk\_g, data = data[-unlist(Kfold\_res[val\_list[i]]),])

tmp = data[unlist(Kfold\_res[val\_list[i]]),] %>% select(event, risk\_g)

ypred = ifelse(predict(glm\_stat,tmp)<=0.5,0,1)

tmp$W = 1

e = aggregate(W~risk\_g, sum, data = tmp)

dth\_rate = aggregate(event~risk\_g, sum, data = tmp)

mm = merge(e,dth\_rate)

mm$rate = round(mm$event / mm$W,2)

event\_rate[,i] = mm$rate

cm = table(ypred, tmp$event)

res = c(res,sum(diag(cm)) / sum(cm))

}

m =mean(res) %>% round(2)

CI = (mean(res) + c(-1,1)\*1.96\*sd(res)) %>% round(2)

m\_d = apply(event\_rate,1,mean)

sd\_d = apply(event\_rate,1,sd)

CI\_d\_H = (m\_d + 1.96\*sd\_d) %>% round(2)\* 100

CI\_d\_L = (m\_d - 1.96\*sd\_d) %>% round(2)\* 100

tt = paste0((m\_d%>% round(2))\* 100," (",CI\_d\_L,"-",CI\_d\_H,")") %>% data.frame()

final$tt <- tt

#C statistic

Cstat = paste0(m," (",CI[1],"-",CI[2],")")

final$Cstat <-Cstat

return(final)

}

#############################################################

# Table4 Total 만성질환 종류

#############################################################

Seoul = data %>% filter(mx\_Seoul == 1)

Seoul$event = Seoul$dth

Seoul$bmi\_g2 = factor(Seoul$bmi\_g, levels = c(2,1,3))

Busan = data %>% filter(mx\_Seoul == 0)

Busan$event = Busan$dth

Busan$bmi\_g2 = factor(Busan$bmi\_g, levels = c(2,1,3))

Seoul$event\_1y = ifelse(Seoul$dth\_fu >= 365,0,Seoul$event)

Seoul$event\_3y = ifelse(Seoul$dth\_fu >= 365\*3,0,Seoul$event)

Busan$event\_1y = ifelse(Busan$dth\_fu >= 365,0,Busan$event)

Busan$event\_3y = ifelse(Busan$dth\_fu >= 365\*3,0,Busan$event)

model\_glm = glm(event ~ Arthritis + Asthma + Cancer + metastatic +

cardiac\_arrhythmias + kidney + hf + Coronary + Dementia + Depression +

Diabetes + Hepatitis + HTN + Stroke + SAD + SEX\_TYPE + smk\_g +

drk\_g +pa\_g +low\_incm +bmi\_g2 +mx\_ER +treatment +age\_g,

data = Seoul)

k = summary(model\_glm)

Seoul$risk = predict(model\_glm)

Busan$risk = predict(model\_glm, newdata = Busan)

min\_max <- function(x){

(x - min(x)) / (max(x) - min(x))

}

Seoul$risk <- min\_max(Seoul$risk)

Busan$risk <- min\_max(Busan$risk)

q1 = 0.4

q3 = 0.6

Seoul$risk\_g <- ifelse(Seoul$risk <= q1,1,

ifelse(Seoul$risk <= q3,2,3)) %>% factor()

Busan$risk\_g <- ifelse(Busan$risk <= q1,1,

ifelse(Busan$risk <= q3,2,3)) %>% factor()

tt5y\_Seoul = get\_tt(Seoul, "event")

tt1y\_Seoul = get\_tt(Seoul, "event\_1y")

tt3y\_Seoul = get\_tt(Seoul, "event\_3y")

Seoul$W=1

Seoul$W1 = Seoul$W

N = mytable(W1~risk\_g, data = Seoul)

mycsv(N,"N.csv",row.names = F)

N = read.csv("N.csv")

res\_Seoul <- cbind(N[3:5,1:2],tt5y\_Seoul$tt,tt1y\_Seoul$tt,tt3y\_Seoul$tt)

write.csv(res\_Seoul,"res\_Seoul.csv",row.names = F)

tt5y\_Seoul$Cstat

tt5y\_Busan = get\_tt(Busan, "event")

tt1y\_Busan = get\_tt(Busan, "event\_1y")

tt3y\_Busan = get\_tt(Busan, "event\_3y")

Busan$W=1

Busan$W1 = Busan$W

N = mytable(W1~risk\_g, data = Busan)

mycsv(N,"N.csv",row.names = F)

N = read.csv("N.csv")

res\_Busan <- cbind(N[3:5,1:2],tt5y\_Busan$tt,tt1y\_Busan$tt,tt3y\_Busan$tt)

write.csv(res\_Busan,"res\_Busan.csv",row.names = F)

tt5y\_Busan$Cstat

########

#KM

#######

Seoul$time\_yr = Seoul$dth\_fu/ 365

fit <- survfit(Surv(time\_yr, dth) ~ risk\_g, data = Seoul)

res <- ggsurvplot(fit,

conf.int = F,

palette = "lancet",

legend.labs = c(1,2,3),

legend.title = "",

risk.table = T,

risk.table.col = "strata",

risk.table.fontsize = 4.5,

censor = F,

p.val = T,

break.time.by = 1,

xlab = "Time (Years)", ylab = "Survivial", data = Seoul)

res

graph2ppt(file = "Total\_KM.pptx",width = 10, height =10)

#############################################################

# Table4 Male 만성질환 종류

#############################################################

Seoul = data %>% filter(mx\_Seoul == 1 & SEX\_TYPE == 1)

Seoul$event = Seoul$dth

Seoul$bmi\_g2 = factor(Seoul$bmi\_g, levels = c(2,1,3))

Seoul$admin\_day\_g = factor(Seoul$admin\_day\_g)

Busan = data %>% filter(mx\_Seoul == 0 & SEX\_TYPE == 1)

Busan$event = Busan$dth

Busan$bmi\_g2 = factor(Busan$bmi\_g, levels = c(2,1,3))

Busan$admin\_day\_g = factor(Busan$admin\_day\_g)

Seoul$event\_1y = ifelse(Seoul$dth\_fu >= 365,0,Seoul$event)

Seoul$event\_3y = ifelse(Seoul$dth\_fu >= 365\*3,0,Seoul$event)

Busan$event\_1y = ifelse(Busan$dth\_fu >= 365,0,Busan$event)

Busan$event\_3y = ifelse(Busan$dth\_fu >= 365\*3,0,Busan$event)

model\_glm = glm(event ~ Arthritis + Asthma + Cancer + metastatic +

cardiac\_arrhythmias + kidney + hf + Coronary + Dementia + Depression +

Hepatitis + HTN + Osteoporosis + Stroke + SAD + smk\_g +

drk\_g +pa\_g +low\_incm +bmi\_g2 +mx\_ER +treatment +age\_g,

data = Seoul)

k = summary(model\_glm)

Seoul$risk = predict(model\_glm)

Busan$risk = predict(model\_glm, newdata = Busan)

min\_max <- function(x){

(x - min(x)) / (max(x) - min(x))

}

Seoul$risk <- min\_max(Seoul$risk)

Busan$risk <- min\_max(Busan$risk)

q1 = 0.4

q3 = 0.6

Seoul$risk\_g <- ifelse(Seoul$risk <= q1,1,

ifelse(Seoul$risk <= q3,2,3)) %>% factor()

Busan$risk\_g <- ifelse(Busan$risk <= q1,1,

ifelse(Busan$risk <= q3,2,3)) %>% factor()

tt5y\_Seoul = get\_tt(Seoul, "event")

tt1y\_Seoul = get\_tt(Seoul, "event\_1y")

tt3y\_Seoul = get\_tt(Seoul, "event\_3y")

Seoul$W=1

Seoul$W1 = Seoul$W

N = mytable(W1~risk\_g, data = Seoul)

mycsv(N,"N.csv",row.names = F)

N = read.csv("N.csv")

res\_Seoul <- cbind(N[3:5,1:2],tt5y\_Seoul$tt,tt1y\_Seoul$tt,tt3y\_Seoul$tt)

write.csv(res\_Seoul,"res\_Seoul\_male.csv",row.names = F)

tt5y\_Seoul$Cstat

tt5y\_Busan = get\_tt(Busan, "event")

tt1y\_Busan = get\_tt(Busan, "event\_1y")

tt3y\_Busan = get\_tt(Busan, "event\_3y")

Busan$W=1

Busan$W1 = Busan$W

N = mytable(W1~risk\_g, data = Busan)

mycsv(N,"N.csv",row.names = F)

N = read.csv("N.csv")

res\_Busan <- cbind(N[3:5,1:2],tt5y\_Busan$tt,tt1y\_Busan$tt,tt3y\_Busan$tt)

write.csv(res\_Busan,"res\_Busan\_male.csv",row.names = F)

tt5y\_Busan$Cstat

########

#KM

#######

Seoul$time\_yr = Seoul$dth\_fu/ 365

fit <- survfit(Surv(time\_yr, dth) ~ risk\_g, data = Seoul)

res <- ggsurvplot(fit,

conf.int = F,

palette = "lancet",

legend.labs = c(1,2,3),

legend.title = "",

risk.table = T,

risk.table.col = "strata",

risk.table.fontsize = 4.5,

censor = F,

p.val = T,

break.time.by = 1,

xlab = "Time (Years)", ylab = "Survivial", data = Seoul)

res

graph2ppt(file = "Male\_KM.pptx",width = 10, height =10)

#############################################################

# Table4 female 만성질환 종류

#############################################################

Seoul = data %>% filter(mx\_Seoul == 1 & SEX\_TYPE == 2)

Seoul$event = Seoul$dth

Seoul$bmi\_g2 = factor(Seoul$bmi\_g, levels = c(2,1,3))

Seoul$admin\_day\_g = factor(Seoul$admin\_day\_g)

Busan = data %>% filter(mx\_Seoul == 0 & SEX\_TYPE == 2)

Busan$event = Busan$dth

Busan$bmi\_g2 = factor(Busan$bmi\_g, levels = c(2,1,3))

Busan$admin\_day\_g = factor(Busan$admin\_day\_g)

Seoul$event\_1y = ifelse(Seoul$dth\_fu >= 365,0,Seoul$event)

Seoul$event\_3y = ifelse(Seoul$dth\_fu >= 365\*3,0,Seoul$event)

Busan$event\_1y = ifelse(Busan$dth\_fu >= 365,0,Busan$event)

Busan$event\_3y = ifelse(Busan$dth\_fu >= 365\*3,0,Busan$event)

model\_glm = glm(event ~ Arthritis + Cancer + metastatic +

kidney + hf + Coronary + Dementia +

Diabetes + Hepatitis + HTN + Stroke + smk\_g +

pa\_g +low\_incm +bmi\_g2 +mx\_ER +treatment +age\_g,

data = Seoul)

k = summary(model\_glm)

Seoul$risk = predict(model\_glm)

Busan$risk = predict(model\_glm, newdata = Busan)

min\_max <- function(x){

(x - min(x)) / (max(x) - min(x))

}

Seoul$risk <- min\_max(Seoul$risk)

Busan$risk <- min\_max(Busan$risk)

q1 = 0.33

q3 = 0.53

Seoul$risk\_g <- ifelse(Seoul$risk <= q1,1,

ifelse(Seoul$risk <= q3,2,3)) %>% factor()

Busan$risk\_g <- ifelse(Busan$risk <= q1,1,

ifelse(Busan$risk <= q3,2,3)) %>% factor()

tt5y\_Seoul = get\_tt(Seoul, "event")

tt1y\_Seoul = get\_tt(Seoul, "event\_1y")

tt3y\_Seoul = get\_tt(Seoul, "event\_3y")

Seoul$W=1

Seoul$W1 = Seoul$W

N = mytable(W1~risk\_g, data = Seoul)

mycsv(N,"N.csv",row.names = F)

N = read.csv("N.csv")

res\_Seoul <- cbind(N[3:5,1:2],tt5y\_Seoul$tt,tt1y\_Seoul$tt,tt3y\_Seoul$tt)

write.csv(res\_Seoul,"res\_Seoul\_female.csv",row.names = F)

tt5y\_Seoul$Cstat

tt5y\_Busan = get\_tt(Busan, "event")

tt1y\_Busan = get\_tt(Busan, "event\_1y")

tt3y\_Busan = get\_tt(Busan, "event\_3y")

Busan$W=1

Busan$W1 = Busan$W

N = mytable(W1~risk\_g, data = Busan)

mycsv(N,"N.csv",row.names = F)

N = read.csv("N.csv")

res\_Busan <- cbind(N[3:5,1:2],tt5y\_Busan$tt,tt1y\_Busan$tt,tt3y\_Busan$tt)

write.csv(res\_Busan,"res\_Busan\_female.csv",row.names = F)

tt5y\_Busan$Cstat

########

#KM

#######

Seoul$time\_yr = Seoul$dth\_fu/ 365

fit <- survfit(Surv(time\_yr, dth) ~ risk\_g, data = Seoul)

res <- ggsurvplot(fit,

conf.int = F,

palette = "lancet",

legend.labs = c(1,2,3),

legend.title = "",

risk.table = T,

risk.table.col = "strata",

risk.table.fontsize = 4.5,

censor = F,

p.val = T,

break.time.by = 1,

xlab = "Time (Years)", ylab = "Survivial", data = Seoul)

res

graph2ppt(file = "Female\_KM.pptx",width = 10, height =10)

#############################################################

# Table4 Total 만성질환 갯수

#############################################################

Seoul = data %>% filter(mx\_Seoul == 1)

Seoul$event = Seoul$dth

Seoul$bmi\_g2 = factor(Seoul$bmi\_g, levels = c(2,1,3))

Busan = data %>% filter(mx\_Seoul == 0)

Busan$event = Busan$dth

Busan$bmi\_g2 = factor(Busan$bmi\_g, levels = c(2,1,3))

Seoul$event\_1y = ifelse(Seoul$dth\_fu >= 365,0,Seoul$event)

Seoul$event\_3y = ifelse(Seoul$dth\_fu >= 365\*3,0,Seoul$event)

Busan$event\_1y = ifelse(Busan$dth\_fu >= 365,0,Busan$event)

Busan$event\_3y = ifelse(Busan$dth\_fu >= 365\*3,0,Busan$event)

model\_glm = glm(event ~ nocd\_g + SEX\_TYPE + smk\_g +

drk\_g +pa\_g +low\_incm +bmi\_g2 +mx\_ER +treatment +age\_g,

data = Seoul)

k = summary(model\_glm)

Seoul$risk = predict(model\_glm)

Busan$risk = predict(model\_glm, newdata = Busan)

min\_max <- function(x){

(x - min(x)) / (max(x) - min(x))

}

Seoul$risk <- min\_max(Seoul$risk)

Busan$risk <- min\_max(Busan$risk)

q1 = 0.45

q3 = 0.68

Seoul$risk\_g <- ifelse(Seoul$risk <= q1,1,

ifelse(Seoul$risk <= q3,2,3)) %>% factor()

Busan$risk\_g <- ifelse(Busan$risk <= q1,1,

ifelse(Busan$risk <= q3,2,3)) %>% factor()

tt5y\_Seoul = get\_tt(Seoul, "event")

tt1y\_Seoul = get\_tt(Seoul, "event\_1y")

tt3y\_Seoul = get\_tt(Seoul, "event\_3y")

Seoul$W=1

Seoul$W1 = Seoul$W

N = mytable(W1~risk\_g, data = Seoul)

mycsv(N,"N.csv",row.names = F)

N = read.csv("N.csv")

res\_Seoul <- cbind(N[3:5,1:2],tt5y\_Seoul$tt,tt1y\_Seoul$tt,tt3y\_Seoul$tt)

write.csv(res\_Seoul,"res\_nocd\_Seoul.csv",row.names = F)

tt5y\_Seoul$Cstat

tt5y\_Busan = get\_tt(Busan, "event")

tt1y\_Busan = get\_tt(Busan, "event\_1y")

tt3y\_Busan = get\_tt(Busan, "event\_3y")

Busan$W=1

Busan$W1 = Busan$W

N = mytable(W1~risk\_g, data = Busan)

mycsv(N,"N.csv",row.names = F)

N = read.csv("N.csv")

res\_Busan <- cbind(N[3:5,1:2],tt5y\_Busan$tt,tt1y\_Busan$tt,tt3y\_Busan$tt)

write.csv(res\_Busan,"res\_nocd\_Busan.csv",row.names = F)

tt5y\_Busan$Cstat

########

#KM

#######

Seoul$time\_yr = Seoul$dth\_fu/ 365

fit <- survfit(Surv(time\_yr, dth) ~ risk\_g, data = Seoul)

res <- ggsurvplot(fit,

conf.int = F,

palette = "lancet",

legend.labs = c(1,2,3),

legend.title = "",

risk.table = T,

risk.table.col = "strata",

risk.table.fontsize = 4.5,

censor = F,

p.val = T,

break.time.by = 1,

xlab = "Time (Years)", ylab = "Survivial", data = Seoul)

res

graph2ppt(file = "nocd\_Total\_KM.pptx",width = 10, height =10)

#############################################################

# Table4 Male 만성질환 갯수

#############################################################

Seoul = data %>% filter(mx\_Seoul == 1 & SEX\_TYPE == 1)

Seoul$event = Seoul$dth

Seoul$bmi\_g2 = factor(Seoul$bmi\_g, levels = c(2,1,3))

Seoul$admin\_day\_g = factor(Seoul$admin\_day\_g)

Busan = data %>% filter(mx\_Seoul == 0 & SEX\_TYPE == 1)

Busan$event = Busan$dth

Busan$bmi\_g2 = factor(Busan$bmi\_g, levels = c(2,1,3))

Busan$admin\_day\_g = factor(Busan$admin\_day\_g)

Seoul$event\_1y = ifelse(Seoul$dth\_fu >= 365,0,Seoul$event)

Seoul$event\_3y = ifelse(Seoul$dth\_fu >= 365\*3,0,Seoul$event)

Busan$event\_1y = ifelse(Busan$dth\_fu >= 365,0,Busan$event)

Busan$event\_3y = ifelse(Busan$dth\_fu >= 365\*3,0,Busan$event)

model\_glm = glm(event ~ nocd\_g + smk\_g +

drk\_g +pa\_g +low\_incm +bmi\_g2 +mx\_ER +treatment +age\_g,

data = Seoul)

k = summary(model\_glm)

Seoul$risk = predict(model\_glm)

Busan$risk = predict(model\_glm, newdata = Busan)

min\_max <- function(x){

(x - min(x)) / (max(x) - min(x))

}

Seoul$risk <- min\_max(Seoul$risk)

Busan$risk <- min\_max(Busan$risk)

q1 = 0.45

q3 = 0.68

Seoul$risk\_g <- ifelse(Seoul$risk <= q1,1,

ifelse(Seoul$risk <= q3,2,3)) %>% factor()

Busan$risk\_g <- ifelse(Busan$risk <= q1,1,

ifelse(Busan$risk <= q3,2,3)) %>% factor()

tt5y\_Seoul = get\_tt(Seoul, "event")

tt1y\_Seoul = get\_tt(Seoul, "event\_1y")

tt3y\_Seoul = get\_tt(Seoul, "event\_3y")

Seoul$W=1

Seoul$W1 = Seoul$W

N = mytable(W1~risk\_g, data = Seoul)

mycsv(N,"N.csv",row.names = F)

N = read.csv("N.csv")

res\_Seoul <- cbind(N[3:5,1:2],tt5y\_Seoul$tt,tt1y\_Seoul$tt,tt3y\_Seoul$tt)

write.csv(res\_Seoul,"res\_nocd\_Seoul\_male.csv",row.names = F)

tt5y\_Seoul$Cstat

tt5y\_Busan = get\_tt(Busan, "event")

tt1y\_Busan = get\_tt(Busan, "event\_1y")

tt3y\_Busan = get\_tt(Busan, "event\_3y")

Busan$W=1

Busan$W1 = Busan$W

N = mytable(W1~risk\_g, data = Busan)

mycsv(N,"N.csv",row.names = F)

N = read.csv("N.csv")

res\_Busan <- cbind(N[3:5,1:2],tt5y\_Busan$tt,tt1y\_Busan$tt,tt3y\_Busan$tt)

write.csv(res\_Busan,"res\_nocd\_Busan\_male.csv",row.names = F)

tt5y\_Busan$Cstat

########

#KM

#######

Seoul$time\_yr = Seoul$dth\_fu/ 365

fit <- survfit(Surv(time\_yr, dth) ~ risk\_g, data = Seoul)

res <- ggsurvplot(fit,

conf.int = F,

palette = "lancet",

legend.labs = c(1,2,3),

legend.title = "",

risk.table = T,

risk.table.col = "strata",

risk.table.fontsize = 4.5,

censor = F,

p.val = T,

break.time.by = 1,

xlab = "Time (Years)", ylab = "Survivial", data = Seoul)

res

graph2ppt(file = "nocd\_Male\_KM.pptx",width = 10, height =10)

#############################################################

# Table4 female 만성질환 갯수

#############################################################

Seoul = data %>% filter(mx\_Seoul == 1 & SEX\_TYPE == 2)

Seoul$event = Seoul$dth

Seoul$bmi\_g2 = factor(Seoul$bmi\_g, levels = c(2,1,3))

Seoul$admin\_day\_g = factor(Seoul$admin\_day\_g)

Busan = data %>% filter(mx\_Seoul == 0 & SEX\_TYPE == 2)

Busan$event = Busan$dth

Busan$bmi\_g2 = factor(Busan$bmi\_g, levels = c(2,1,3))

Busan$admin\_day\_g = factor(Busan$admin\_day\_g)

Seoul$event\_1y = ifelse(Seoul$dth\_fu >= 365,0,Seoul$event)

Seoul$event\_3y = ifelse(Seoul$dth\_fu >= 365\*3,0,Seoul$event)

Busan$event\_1y = ifelse(Busan$dth\_fu >= 365,0,Busan$event)

Busan$event\_3y = ifelse(Busan$dth\_fu >= 365\*3,0,Busan$event)

model\_glm = glm(event ~ nocd\_g + smk\_g + drk\_g +

pa\_g +low\_incm +bmi\_g2 +mx\_ER +treatment +age\_g,

data = Seoul)

k = summary(model\_glm)

Seoul$risk = predict(model\_glm)

Busan$risk = predict(model\_glm, newdata = Busan)

min\_max <- function(x){

(x - min(x)) / (max(x) - min(x))

}

Seoul$risk <- min\_max(Seoul$risk)

Busan$risk <- min\_max(Busan$risk)

q1 = 0.3

q3 = 0.55

Seoul$risk\_g <- ifelse(Seoul$risk <= q1,1,

ifelse(Seoul$risk <= q3,2,3)) %>% factor()

Busan$risk\_g <- ifelse(Busan$risk <= q1,1,

ifelse(Busan$risk <= q3,2,3)) %>% factor()

tt5y\_Seoul = get\_tt(Seoul, "event")

tt1y\_Seoul = get\_tt(Seoul, "event\_1y")

tt3y\_Seoul = get\_tt(Seoul, "event\_3y")

Seoul$W=1

Seoul$W1 = Seoul$W

N = mytable(W1~risk\_g, data = Seoul)

mycsv(N,"N.csv",row.names = F)

N = read.csv("N.csv")

res\_Seoul <- cbind(N[3:5,1:2],tt5y\_Seoul$tt,tt1y\_Seoul$tt,tt3y\_Seoul$tt)

write.csv(res\_Seoul,"res\_nocd\_Seoul\_female.csv",row.names = F)

tt5y\_Seoul$Cstat

tt5y\_Busan = get\_tt(Busan, "event")

tt1y\_Busan = get\_tt(Busan, "event\_1y")

tt3y\_Busan = get\_tt(Busan, "event\_3y")

Busan$W=1

Busan$W1 = Busan$W

N = mytable(W1~risk\_g, data = Busan)

mycsv(N,"N.csv",row.names = F)

N = read.csv("N.csv")

res\_Busan <- cbind(N[3:5,1:2],tt5y\_Busan$tt,tt1y\_Busan$tt,tt3y\_Busan$tt)

write.csv(res\_Busan,"res\_nocd\_Busan\_female.csv",row.names = F)

tt5y\_Busan$Cstat

########

#KM

#######

Seoul$time\_yr = Seoul$dth\_fu/ 365

fit <- survfit(Surv(time\_yr, dth) ~ risk\_g, data = Seoul)

res <- ggsurvplot(fit,

conf.int = F,

palette = "lancet",

legend.labs = c(1,2,3),

legend.title = "",

risk.table = T,

risk.table.col = "strata",

risk.table.fontsize = 4.5,

censor = F,

p.val = T,

break.time.by = 1,

xlab = "Time (Years)", ylab = "Survivial", data = Seoul)

res

graph2ppt(file = "nocd\_Female\_KM.pptx",width = 10, height =10)