Hormonal contraceptives and new-onset asthma in women

Kun-Sheng Peng kun-sheng.peng@hotmail.com, Pan Boans, Maya B. Mathur

1. Load Packages

```
library(tidyverse)
library(haven)
library(dplyr)
library(ggplot2)
library(ggcorrplot)
library(tableone)
library(Matrix)
library(survdata)
library(survival)
library(survminer)
library(Hmisc)
library(ggsurvplot)
library(ggfortify)
library(ggpubr)
library(ggsurvfit)
library(survfitplus)
library(survfitcoxph)
library(survfitAJ)
library(survfitKM)
```

2. Import Dataset

```
# Import dataset
dataset <- read.csv("/Users/kunshengpeng/Desktop/Medical Research/Hormonal Contraceptives and Asthma/Hormonal Contraceptives and Asthma - Data
Set/Hormonal Contraceptives and Asthma - Data Set.csv", header = TRUE, sep = ",", stringsAsFactors = TRUE)
# Import dataset_Medical Exam
dataset_ME <- read.csv("/Users/kunshengpeng/Desktop/Medical Research/Hormonal Contraceptives and Asthma/Hormonal Contraceptives and Asthma - Data
Set/Hormonal Contraceptives and Asthma - Data Set_Medical Exam.csv", header = TRUE, sep = ",", stringsAsFactors = TRUE)

# Make a new backup dataset
dat <- read.csv("/Users/kunshengpeng/Desktop/Medical Research/Hormonal Contraceptives and Asthma/Hormonal Contraceptives and Asthma - Data
Set/Hormonal Contraceptives and Asthma - Data Set.csv", header = TRUE, sep = ",", stringsAsFactors = TRUE)

# View data
View(dataset)
# Import dataset_Medical Exam
dat_ME <- read.csv("/Users/kunshengpeng/Desktop/Medical Research/Hormonal Contraceptives and Asthma/Hormonal Contraceptives and Asthma - Data
Set/Hormonal Contraceptives and Asthma - Data Set_Medical Exam.csv", header = TRUE, sep = ",", stringsAsFactors = TRUE)
# View dataset_Medical Exam
View(dat_ME)

# Merge two datasets by "SEQN"
dat_merge <- merge(dat, dat_ME, by = "SEQN")
# Check the number of observations and variables
dim(dat_merge)
# [1] 20293  1344
# Export merged dataset
write.csv(dat_merge, file = "/Users/kunshengpeng/Desktop/Medical Research/Hormonal Contraceptives and Asthma/Hormonal Contraceptives and Asthma - Data
Set/dat_merge.csv")

# Make a new variable "Hormonal_Contraceptives"
dat_merge$Hormonal_Contraceptives <- NA
# If any of the variables "Taking_Birth_Control_Pills", "Ever_taken_Birth_Control_Pills", "Ever_told_taking_Birth_Control_Pills" is "Yes", then
"Hormonal_Contraceptives" = "Yes"
dat_merge$Hormonal_Contraceptives[dat_merge$Taking_Birth_Control_Pills == "Yes" | dat_merge$Ever_taken_Birth_Control_Pills == "Yes" |
dat_merge$Ever_told_taking_Birth_Control_Pills == "Yes"] <- "Yes"
# If all the variables are "No", then "Hormonal_Contraceptives" = "No"
dat_merge$Hormonal_Contraceptives[dat_merge$Taking_Birth_Control_Pills == "No" & dat_merge$Ever_taken_Birth_Control_Pills == "No" &
dat_merge$Ever_told_taking_Birth_Control_Pills == "No"] <- "No"
# If all the variables are "Don't know", then "Hormonal_Contraceptives" = "Don't know"
dat_merge$Hormonal_Contraceptives[dat_merge$Taking_Birth_Control_Pills == "Don't know" & dat_merge$Ever_taken_Birth_Control_Pills == "Don't
know" & dat_merge$Ever_told_taking_Birth_Control_Pills == "Don't know"] <- "Don't know"
# Check the number of observations and variables
dim(dat_merge)
# [1] 20293  1345
# Check the number of "Yes", "No", and "Don't know" in the new variable "Hormonal_Contraceptives"
table(dat_merge$Hormonal_Contraceptives)
# Don't know     No    Yes
# 1.000e+00 9.000e+03 1.029e+04
# Make a new variable "Hormonal_Contraceptives_numeric"
dat_merge$Hormonal_Contraceptives_numeric <- NA
# If "Hormonal_Contraceptives" = "Yes", then "Hormonal_Contraceptives_numeric" = 1
dat_merge$Hormonal_Contraceptives_numeric[dat_merge$Hormonal_Contraceptives == "Yes"] <- 1
# If "Hormonal_Contraceptives" = "No", then "Hormonal_Contraceptives_numeric" = 0
dat_merge$Hormonal_Contraceptives_numeric[dat_merge$Hormonal_Contraceptives == "No"] <- 0
# If "Hormonal_Contraceptives" = "Don't know", then "Hormonal_Contraceptives_numeric" = NA
dat_merge$Hormonal_Contraceptives_numeric[dat_merge$Hormonal_Contraceptives == "Don't know"] <- NA
# Check the number of observations and variables
dim(dat_merge)
# [1] 20293  1346
# Check the number of "1", "0", and "NA" in the new variable "Hormonal_Contraceptives_numeric"
table(dat_merge$Hormonal_Contraceptives_numeric, exclude = FALSE)
#    0    1 <NA>
# 9000 10293    1
# Make a new variable "Hormonal_Contraceptives_numeric_factor"
dat_merge$Hormonal_Contraceptives_numeric_factor <- NA
# If "Hormonal_Contraceptives_numeric" = 1, then "Hormonal_Contraceptives_numeric_factor" = "Yes"
dat_merge$Hormonal_Contraceptives_numeric_factor[dat_merge$Hormonal_Contraceptives_numeric == 1] <- "Yes"
# If "Hormonal_Contraceptives_numeric" = 0, then "Hormonal_Contraceptives_numeric_factor" = "No"
dat_merge$Hormonal_Contraceptives_numeric_factor[dat_merge$Hormonal_Contraceptives_numeric == 0] <- "No"
# Check the number of observations and variables
dim(dat_merge)
# [1] 20293  1347
# Check the number of "Yes", "No", and "NA" in the new variable "Hormonal_Contraceptives_numeric_factor"
table(dat_merge$Hormonal_Contraceptives_numeric_factor, exclude = FALSE)
#   No  Yes <NA>
# 9000 10293    1
# Make a new variable "Hormonal_Contraceptives_numeric_factor_Yes_Only"
dat_merge$Hormonal_Contraceptives_numeric_factor_Yes_Only <- NA
# If "Hormonal_Contraceptives_numeric_factor" = "Yes", then "Hormonal_Contraceptives_numeric_factor_Yes_Only" = "Yes"
dat_merge$Hormonal_Contraceptives_numeric_factor_Yes_Only[dat_merge$Hormonal_Contraceptives_numeric_factor == "Yes"] <- "Yes"
# Check the number of observations and variables
dim(dat_merge)
# [1] 20293  1348
# Check the number of "Yes", "NA" in the new variable "Hormonal_Contraceptives_numeric_factor_Yes_Only"
table(dat_merge$Hormonal_Contraceptives_numeric_factor_Yes_Only, exclude = FALSE)
#   Yes  <NA>
# 10293 10000
# Make a new variable "Hormonal_Contraceptives_numeric_factor_Yes_Only_Ever_Taken"
dat_merge$Hormonal_Contraceptives_numeric_factor_Yes_Only_Ever_Taken <- NA
# If "Hormonal_Contraceptives_numeric_factor" = "Yes", then "Hormonal_Contraceptives_numeric_factor_Yes_Only_Ever_Taken" = "Yes"
dat_merge$Hormonal_Contraceptives_numeric_factor_Yes_Only_Ever_Taken[dat_merge$Hormonal_Contraceptives_numeric_factor == "Yes"] <- "Yes"
# Check the number of observations and variables
dim(dat_merge)
# [1] 20293  1349
# Check the number of "Yes", "NA" in the new variable "Hormonal_Contraceptives_numeric_factor_Yes_Only_Ever_Taken"
table(dat_merge$Hormonal_Contraceptives_numeric_factor_Yes_Only_Ever_Taken, exclude = FALSE)
#   Yes  <NA>
# 10293 10000
# Make a new variable "Hormonal_Contraceptives_numeric_factor_Yes_Only_Ever_Taken_Ever_Told"
dat_merge$Hormonal_Contraceptives_numeric_factor_Yes_Only_Ever_Taken_Ever_Told <- NA
# If "Hormonal_Contraceptives_numeric_factor" = "Yes", then "Hormonal_Contraceptives_numeric_factor_Yes_Only_Ever_Taken_Ever_Told" = "Yes"
dat_merge$Hormonal_Contraceptives_numeric_factor_Yes_Only_Ever_Taken_Ever_Told[dat_merge$Hormonal_Contraceptives_numeric_factor == "Yes"] <- "Yes"
# Check the number of observations and variables
dim(dat_merge)
# [1] 20293  1350
# Check the number of "Yes", "NA" in the new variable "Hormonal_Contraceptives_numeric_factor_Yes_Only_Ever_Taken_Ever_Told"
table(dat_merge$Hormonal_Contraceptives_numeric_factor_Yes_Only_Ever_Taken_Ever_Told, exclude = FALSE)
#   Yes  <NA>
# 10293 10000
# Make a new variable "Hormonal_Contraceptives_Ever_Taken"
dat_merge$Hormonal_Contraceptives_Ever_Taken <- NA
# If "Taking_Birth_Control_Pills" = "Yes", then "Hormonal_Contraceptives_Ever_Taken" = "Yes"
dat_merge$Hormonal_Contraceptives_Ever_Taken[dat_merge$Taking_Birth_Control_Pills == "Yes"] <- "Yes"
# If "Ever_taken_Birth_Control_Pills" = "Yes", then "Hormonal_Contraceptives_Ever_Taken" = "Yes"
dat_merge$Hormonal_Contraceptives_Ever_Taken[dat_merge$Ever_taken_Birth_Control_Pills == "Yes"] <- "Yes"
# If "Ever_told_taking_Birth_Control_Pills" = "Yes", then "Hormonal_Contraceptives_Ever_Taken" = "Yes"
dat_merge$Hormonal_Contraceptives_Ever_Taken[dat_merge$Ever_told_taking_Birth_Control_Pills == "Yes"] <- "Yes"
# If all the variables are "No", then "Hormonal_Contraceptives_Ever_Taken" = "No"
dat_merge$Hormonal_Contraceptives_Ever_Taken[dat_merge$Taking_Birth_Control_Pills == "No" & dat_merge$Ever_taken_Birth_Control_Pills == "No" &
dat_merge$Ever_told_taking_Birth_Control_Pills == "No"] <- "No"
# If all the variables are "Don't know", then "Hormonal_Contraceptives_Ever_Taken" = "Don't know"
dat_merge$Hormonal_Contraceptives_Ever_Taken[dat_merge$Taking_Birth_Control_Pills == "Don't know" & dat_merge$Ever_taken_Birth_Control_Pills
== "Don't know" & dat_merge$Ever_told_taking_Birth_Control_Pills == "Don't know"] <- "Don't know"
# Check the number of observations and variables
dim(dat_merge)
# [1] 20293  1351
# Check the number of "Yes", "No", and "Don't know" in the new variable "Hormonal_Contraceptives_Ever_Taken"
table(dat_merge$Hormonal_Contraceptives_Ever_Taken, exclude = FALSE)
# Don't know     No    Yes
# 1.000e+00 9.000e+03 1.029e+04
# Make a new variable "Hormonal_Contraceptives_Ever_Taken_Yes_Only"
dat_merge$Hormonal_Contraceptives_Ever_Taken_Yes_Only <- NA
# If "Hormonal_Contraceptives_Ever_Taken" = "Yes", then "Hormonal_Contraceptives_Ever_Taken_Yes_Only" = "Yes"
dat_merge$Hormonal_Contraceptives_Ever_Taken_Yes_Only[dat_merge$Hormonal_Contraceptives_Ever_Taken == "Yes"] <- "Yes"
# Check the number of observations and variables
dim(dat_merge)
# [1] 20293  1352
# Check the number of "Yes", "NA" in the new variable "Hormonal_Contraceptives_Ever_Taken_Yes_Only"
table(dat_merge$Hormonal_Contraceptives_Ever_Taken_Yes_Only, exclude = FALSE)
#   Yes  <NA>
# 10293 10000
# Make a new variable "Hormonal_Contraceptives_Ever_Taken_Ever_Told"
dat_merge$Hormonal_Contraceptives_Ever_Taken_Ever_Told <- NA
# If "Hormonal_Contraceptives_Ever_Taken" = "Yes", then "Hormonal_Contraceptives_Ever_Taken_Ever_Told" = "Yes"
dat_merge$Hormonal_Contraceptives_Ever_Taken_Ever_Told[dat_merge$Hormonal_Contraceptives_Ever_Taken == "Yes"] <- "Yes"
# Check the number of observations and variables
dim(dat_merge)
# [1] 20293  1353
# Check the number of "Yes", "NA" in the new variable "Hormonal_Contraceptives_Ever_Taken_Ever_Told"
table(dat_merge$Hormonal_Contraceptives_Ever_Taken_Ever_Told, exclude = FALSE)
#   Yes  <NA>
# 10293 10000
# Make a new variable "Hormonal_Contraceptives_Ever_Taken_Ever_Told_Numeric"
dat_merge$Hormonal_Contraceptives_Ever_Taken_Ever_Told_Numeric <- NA
# If "Hormonal_Contraceptives_Ever_Taken_Ever_Told" = "Yes", then "Hormonal_Contraceptives_Ever_Taken_Ever_Told_Numeric" = 1
dat_merge$Hormonal_Contraceptives_Ever_Taken_Ever_Told_Numeric[dat_merge$Hormonal_Contraceptives_Ever_Taken_Ever_Told == "Yes"] <- 1
# Check the number of observations and variables
dim(dat_merge)
# [1] 20293  1354
# Check the number of "1", "NA" in the new variable "Hormonal_Contraceptives_Ever_Taken_Ever_Told_Numeric"
table(dat_merge$Hormonal_Contraceptives_Ever_Taken_Ever_Told_Numeric, exclude = FALSE)
#    1 <NA>
# 10293 10000
# Make a new variable "Hormonal_Contraceptives_Ever_Taken_Ever_Told_Numeric_Factor"
dat_merge$Hormonal_Contraceptives_Ever_Taken_Ever_Told_Numeric_Factor <- NA
# If "Hormonal_Contraceptives_Ever_Taken_Ever_Told_Numeric" = 1, then "Hormonal_Contraceptives_Ever_Taken_Ever_Told_Numeric_Factor" = "Yes"
dat_merge$Hormonal_Contraceptives_Ever_Taken_Ever_Told_Numeric_Factor[dat_merge$Hormonal_Contraceptives_Ever_Taken_Ever_Told_Numeric == 1] <-
"Yes"
# Check the number of observations and variables
dim(dat_merge)
# [1] 20293  1355
# Check the number of "Yes", "NA" in the new variable "Hormonal_Contraceptives_Ever_Taken_Ever_Told_Numeric_Factor"
table(dat_merge$Hormonal_Contraceptives_Ever_Taken_Ever_Told_Numeric_Factor, exclude = FALSE)
#   Yes  <NA>
# 10293 10000
# Export the new dataset
write.csv(dat_merge, file = "/Users/kunshengpeng/Desktop/Medical Research/Hormonal Contraceptives and Asthma/Hormonal Contraceptives and Asthma - Data
Set/dat_merge_1.csv")
```

3. Import Dataset

```
# Import dataset
dat_merge <- read.csv("/Users/kunshengpeng/Desktop/Medical Research/Hormonal Contraceptives and Asthma/Hormonal Contraceptives and Asthma - Data
Set/dat_merge_1.csv", header = TRUE, sep = ",", stringsAsFactors = TRUE)
# Make a new backup dataset
dat <- read.csv("/Users/kunshengpeng/Desktop/Medical Research/Hormonal Contraceptives and Asthma/Hormonal Contraceptives and Asthma - Data
Set/dat_merge_1.csv", header = TRUE, sep = ",", stringsAsFactors = TRUE)
# View data
View(dat_merge)
# Check the number of observations and variables
dim(dat_merge)
# [1] 20293  1355
# Check the number of "Yes", "NA" in the new variable "Hormonal_Contraceptives_Ever_Taken_Ever_Told_Numeric_Factor"
table(dat_merge$Hormonal_Contraceptives_Ever_Taken_Ever_Told_Numeric_Factor, exclude = FALSE)
#   Yes  <NA>
# 10293 10000
# Make a new variable "Hormonal_Contraceptives_Ever_Taken_Ever_Told_Numeric_Factor_Yes_Only"
dat_merge$Hormonal_Contraceptives_Ever_Taken_Ever_Told_Numeric_Factor_Yes_Only <- NA
# If "Hormonal_Contraceptives_Ever_Taken_Ever_Told_Numeric_Factor" = "Yes", then "Hormonal_Contraceptives_Ever_Taken_Ever_Told_Numeric_Factor_Yes_Only"
= "Yes"
dat_merge$Hormonal_Contraceptives_Ever_Taken_Ever_Told_Numeric_Factor_Yes_Only[dat_merge$Hormonal_Contraceptives_Ever_Taken_Ever_Told_Numeric_Factor
== "Yes"] <- "Yes"
# Check the number of observations and variables
dim(dat_merge)
# [1] 20293  1356
# Check the number of "Yes", "NA" in the new variable "Hormonal_Contraceptives_Ever_Taken_Ever_Told_Numeric_Factor_Yes_Only"
table(dat_merge$Hormonal_Contraceptives_Ever_Taken_Ever_Told_Numeric_Factor_Yes_Only, exclude = FALSE)
#   Yes  <NA>
# 10293 10000
# Make a new variable "Hormonal_Contraceptives_Ever_Taken_Ever_Told_Numeric_Factor_Yes_Only_Ever_Taken"
dat_merge$Hormonal_Contraceptives_Ever_Taken_Ever_Told_Numeric_Factor_Yes_Only_Ever_Taken <- NA
# If "Hormonal_Contraceptives_Ever_Taken" = "Yes", then "Hormonal_Contraceptives_Ever_Taken_Ever_Told_Numeric_Factor_Yes_Only_Ever_Taken" =
"Yes"
dat_merge$Hormonal_Contraceptives_Ever_Taken_Ever_Told_Numeric_Factor_Yes_Only_Ever_Taken[dat_merge$Hormonal_Contraceptives_Ever_Taken == "Yes"]
<- "Yes"
# Check the number of observations and variables
dim(dat_merge)
# [1] 20293  1357
# Check the number of "Yes", "NA" in the new variable "Hormonal_Contraceptives_Ever_Taken_Ever_Told_Numeric_Factor_Yes_Only_Ever_Taken"
table(dat_merge$Hormonal_Contraceptives_Ever_Taken_Ever_Told_Numeric_Factor_Yes_Only_Ever_Taken, exclude = FALSE)
#   Yes  <NA>
# 10293 10000
# Export the new dataset
write.csv(dat_merge, file = "/Users/kunshengpeng/Desktop/Medical Research/Hormonal Contraceptives and Asthma/Hormonal Contraceptives and Asthma - Data
Set/dat_merge_2.csv")
```

4. Complete-Case Analysis

4.1. Frequentist conditional logistic regression

```
# Make sure R knows
dat <- factor(dat$Hormonal_Contraceptives_Ever_Taken_Ever_Told_Numeric_Factor, levels = c("Yes", "No"))
# Make a table
table(dat, exclude = FALSE)
# dat
#   Yes  <NA>
# 10293 10000
# Logistic regression analysis by age in years at screening
# Change variable "Hormonal_Contraceptives_Ever_Taken_Ever_Told_Numeric_Factor" to factor
dat$Hormonal_Taken <- dat$Hormonal_Contraceptives_Ever_Taken_Ever_Told_Numeric_Factor
# Make a table
table(dat$Hormonal_Taken, exclude = FALSE)
# Yes  <NA>
# 10293 10000
# Logistic regression analysis by age in years at screening
# Change variable "Hormonal_Taken" to factor
dat$Hormonal_Taken <- factor(dat$Hormonal_Taken, levels = c("Yes", "No"))
# Make a table
table(dat$Hormonal_Taken, exclude = FALSE)
# Yes  <NA>
# 10293 10000
# Logistic regression analysis by age in years at screening
# Change variable "Age_in_years_at_screening" to factor
dat$Age_in_years_at_screening <- factor(dat$Age_in_years_at_screening, levels = c("Yes", "No"))
# Make a table
table(dat$Age_in_years_at_screening, exclude = FALSE)
# <NA>
# 20293
# Logistic regression analysis by age in years at screening
# Change variable "Asthma" to factor
dat$Asthma <- factor(dat$Asthma, levels = c("Yes", "No"))
# Make a table
table(dat$Asthma, exclude = FALSE)
# Yes  No <NA>
# 2127 8166 9000
# Logistic regression analysis by age in years at screening
# Change variable "Hay_Fever" to factor
dat$Hay_Fever <- factor(dat$Hay_Fever, levels = c("Yes", "No"))
# Make a table
table(dat$Hay_Fever, exclude = FALSE)
# Yes  No <NA>
# 4127 7166 9000
# Logistic regression analysis by age in years at screening
# Change variable "Hay_Fever_Ever_Told" to factor
dat$Hay_Fever_Ever_Told <- factor(dat$Hay_Fever_Ever_Told, levels = c("Yes", "No"))
# Make a table
table(dat$Hay_Fever_Ever_Told, exclude = FALSE)
# Yes  No <NA>
# 4127 7166 9000
# Logistic regression analysis by age in years at screening
# Change variable "Hay_Fever_Taken" to factor
dat$Hay_Fever_Taken <- factor(dat$Hay_Fever_Taken, levels = c("Yes", "No"))
# Make a table
table(dat$Hay_Fever_Taken, exclude = FALSE)
# Yes  No <NA>
# 4127 7166 9000
# Logistic regression analysis by age in years at screening
# Change variable "Ever_told_allergic_rhinitis" to factor
dat$Ever_told_allergic_rhinitis <- factor(dat$Ever_told_allergic_rhinitis, levels = c("Yes", "No"))
# Make a table
table(dat$Ever_told_allergic_rhinitis, exclude = FALSE)
# Yes  No <NA>
# 4127 7166 9000
# Logistic regression analysis by age in years at screening
# Change variable "Ever_taken_allergic_rhinitis" to factor
dat$Ever_taken_allergic_rhinitis <- factor(dat$Ever_taken_allergic_rhinitis, levels = c("Yes", "No"))
# Make a table
table(dat$Ever_taken_allergic_rhinitis, exclude = FALSE)
# Yes  No <NA>
# 4127 7166 9000
# Logistic regression analysis by age in years at screening
# Change variable "Ever_told_allergic_rhinitis_Taken" to factor
dat$Ever_told_allergic_rhinitis_Taken <- factor(dat$Ever_told_allergic_rhinitis_Taken, levels = c("Yes", "No"))
# Make a table
table(dat$Ever_told_allergic_rhinitis_Taken, exclude = FALSE)
# Yes  No <NA>
# 4127 7166 9000
# Logistic regression analysis by age in years at screening
# Change variable "Ever_taken_allergic_rhinitis_Taken" to factor
dat$Ever_taken_allergic_rhinitis_Taken <- factor(dat$Ever_taken_allergic_rhinitis_Taken, levels = c("Yes", "No"))
# Make a table
table(dat$Ever_taken_allergic_rhinitis_Taken, exclude = FALSE)
# Yes  No <NA>
# 4127 7166 9000
# Logistic regression analysis by age in years at screening
# Change variable "Ever_told_allergic_rhinitis_Ever_Taken" to factor
dat$Ever_told_allergic_rhinitis_Ever_Taken <- factor(dat$Ever_told_allergic_rhinitis_Ever_Taken, levels = c("Yes", "No"))
# Make a table
table(dat$Ever_told_allergic_rhinitis_Ever_Taken, exclude = FALSE)
# Yes  No <NA>
# 4127 7166 9000
# Logistic regression analysis by age in years at screening
# Change variable "Ever_taken_allergic_rhinitis_Ever_Taken" to factor
dat$Ever_taken_allergic_rhinitis_Ever_Taken <- factor(dat$Ever_taken_allergic_rhinitis_Ever_Taken, levels = c("Yes", "No"))
# Make a table
table(dat$Ever_taken_allergic_rhinitis_Ever_Taken, exclude = FALSE)
# Yes  No <NA>
# 4127 7166 9000
# Logistic regression analysis by age in years at screening
# Change variable "Ever_told_asthma" to factor
dat$Ever_told_asthma <- factor(dat$Ever_told_asthma, levels = c("Yes", "No"))
# Make a table
table(dat$Ever_told_asthma, exclude = FALSE)
# Yes  No <NA>
# 2127 8166 9000
# Logistic regression analysis by age in years at screening
# Change variable "Ever_taken_asthma" to factor
dat$Ever_taken_asthma <- factor(dat$Ever_taken_asthma, levels = c("Yes", "No"))
# Make a table
table(dat$Ever_taken_asthma, exclude = FALSE)
# Yes  No <NA>
# 2127 8166 9000
# Logistic regression analysis by age in years at screening
# Change variable "Ever_told_asthma_Taken" to factor
dat$Ever_told_asthma_Taken <- factor(dat$Ever_told_asthma_Taken, levels = c("Yes", "No"))
# Make a table
table(dat$Ever_told_asthma_Taken, exclude = FALSE)
# Yes  No <NA>
# 2127 8166 9000
# Logistic regression analysis by age in years at screening
# Change variable "Ever_taken_asthma_Taken" to factor
dat$Ever_taken_asthma_Taken <- factor(dat$Ever_taken_asthma_Taken, levels = c("Yes", "No"))
# Make a table
table(dat$Ever_taken_asthma_Taken, exclude = FALSE)
# Yes  No <NA>
# 2127 8166 9000
# Logistic regression analysis by age in years at screening
# Change variable "Ever_told_asthma_Ever_Taken" to factor
dat$Ever_told_asthma_Ever_Taken <- factor(dat$Ever_told_asthma_Ever_Taken, levels = c("Yes", "No"))
# Make a table
table(dat$Ever_told_asthma_Ever_Taken, exclude = FALSE)
# Yes  No <NA>
# 2127 8166 9000
# Logistic regression analysis by age in years at screening
# Change variable "Ever_taken_asthma_Ever_Taken" to factor
dat$Ever_taken_asthma_Ever_Taken <- factor(dat$Ever_taken_asthma_Ever_Taken, levels = c("Yes", "No"))
# Make a table
table(dat$Ever_taken_asthma_Ever_Taken, exclude = FALSE)
# Yes  No <NA>
# 2127 8166 9000
# Logistic regression analysis by age in years at screening
# Change variable "Ever_told_asthma_Taken" to factor
dat$Ever_told_asthma_Taken <- factor(dat$Ever_told_asthma_Taken, levels = c("Yes", "No"))
# Make a table
table(dat$Ever_told_asthma_Taken, exclude = FALSE)
# Yes  No <NA>
# 2127 8166 9000
# Logistic regression analysis by age in years at screening
# Change variable "Ever_taken_asthma_Taken" to factor
dat$Ever_taken_asthma_Taken <- factor(dat$Ever_taken_asthma_Taken, levels = c("Yes", "No"))
# Make a table
table(dat$Ever_taken_asthma_Taken, exclude = FALSE)
# Yes  No <NA>
# 2127 8166 9000
# Logistic regression analysis by age in years at screening
# Change variable "Ever_told_asthma_Ever_Taken" to factor
dat$Ever_told_asthma_Ever_Taken <- factor(dat$Ever_told_asthma_Ever_Taken, levels = c("Yes", "No"))
# Make a table
table(dat$Ever_told_asthma_Ever_Taken, exclude = FALSE)
# Yes  No <NA>
# 2127 8166 9000
# Logistic regression analysis by age in years at screening
# Change variable "Ever_taken_asthma_Ever_Taken" to factor
dat$Ever_taken_asthma_Ever_Taken <- factor(dat$Ever_taken_asthma_Ever_Taken, levels = c("Yes", "No"))
# Make a table
table(dat$Ever_taken_asthma_Ever_Taken, exclude = FALSE)
# Yes  No <NA>
# 2127 8166 9000
# Logistic regression analysis by age in years at screening
# Change variable "Ever_told_asthma_Taken_Ever_Taken" to factor
dat$Ever_told_asthma_Taken_Ever_Taken <- factor(dat$Ever_told_asthma_Taken_Ever_Taken, levels = c("Yes", "No"))
# Make a table
table(dat$Ever_told_asthma_Taken_Ever_Taken, exclude = FALSE)
# Yes  No <NA>
# 2127 8166 9000
# Logistic regression analysis by age in years at screening
# Change variable "Ever_taken_asthma_Taken_Ever_Taken" to factor
dat$Ever_taken_asthma_Taken_Ever_Taken <- factor(dat$Ever_taken_asthma_Taken_Ever_Taken, levels = c("Yes", "No"))
# Make a table
table(dat$Ever_taken_asthma_Taken_Ever_Taken, exclude = FALSE)
# Yes  No <NA>
# 2127 8166 9000
# Logistic regression analysis by age in years at screening
# Change variable "Ever_told_asthma_Taken_Ever_Taken" to factor
dat$Ever_told_asthma_Taken_Ever_Taken <- factor(dat$Ever_told_asthma_Taken_Ever_Taken, levels = c("Yes", "No"))
# Make a table
table(dat$Ever_told_asthma_Taken_Ever_Taken, exclude = FALSE)
# Yes  No <NA>
# 2127 8166 9000
# Logistic regression analysis by age in years at screening
# Change variable "Ever_taken_asthma_Taken_Ever_Taken" to factor
dat$Ever_taken_asthma_Taken_Ever_Taken <- factor(dat$Ever_taken_asthma_Taken_Ever_Taken, levels = c("Yes", "No"))
# Make a table
table(dat$Ever_taken_asthma_Taken_Ever_Taken, exclude = FALSE)
# Yes  No <NA>
# 2127 8166 9000
# Logistic regression analysis by age in years at screening
# Change variable "Ever_told_asthma_Ever_Taken" to factor
dat$Ever_told_asthma_Ever_Taken <- factor(dat$Ever_told_asthma_Ever_Taken, levels = c("Yes", "No"))
# Make a table
table(dat$Ever_told_asthma_Ever_Taken, exclude = FALSE)
# Yes  No <NA>
# 2127 8166 9000
# Logistic regression analysis by age in years at screening
# Change variable "Ever_taken_asthma_Ever_Taken" to factor
dat$Ever_taken_asthma_Ever_Taken <- factor(dat$Ever_taken_asthma_Ever_Taken, levels = c("Yes", "No"))
# Make a table
table(dat$Ever_taken_asthma
```

```
        panel.grid.minor = element_blank(),
        panel.border = element_blank(),
        plot.title = element_text(size = 18, face = "bold"),
        plot.subtitle = element_text(size = 18),
        axis.title.x = element_text(size = 14),
        axis.title.y = element_text(size = 14),
        axis.text = element_text(size = 14),
        legend.text = element_text(size = 12),
        text = element_text(family = "Arial"),
        axis.line = element_line(color='black'))
```

6. Sensitivity Analysis for Residual Confounding

Hide

```
effect.estimate <-  # on odds ratio scale

# if the outcome is relatively rare (< 15% at the end of follow-up)
evalues.OR(est = effect.estimate, rare = TRUE)

# if the outcome is common (>= 15% at the end of follow-up)
evalues.OR(est = effect.estimate, rare = FALSE)
```

Hide

```
# some helper scripts
save.image(file = "Global_Environment_HC.RData") # save work environment
load(file = "Global_Environment_HC.RData") # load work environment

# record the versions of R packages used
library(renv)
renv::init()
renv::snapshot() # update renv.lock lockfile
```