Analysis of HDL(High-Density Lipoprotein) Cholesterol Level and Symptoms in Dementia and Alzheimer's Disease.

Chaeeun Shin

Overview

High level of HDL(High-Density Lipoprotien) cholesterol is associated with an increased risk of dementia. In EDA procedure for this project, it is observed that there are relatively more participants who have diagnosed Alzheimer's disease than un-diagnosed participants for above HDL level of 80 mg/dL.

The aim of this project is to evaluate relationship between development of symptoms and level of HDL cholesterol.

Below are symptoms used in this analysis.
1. Confusion 2. Disorientation 3. Personality Changes 4. Difficulty in Completing Tasks 5. Forgetfulness

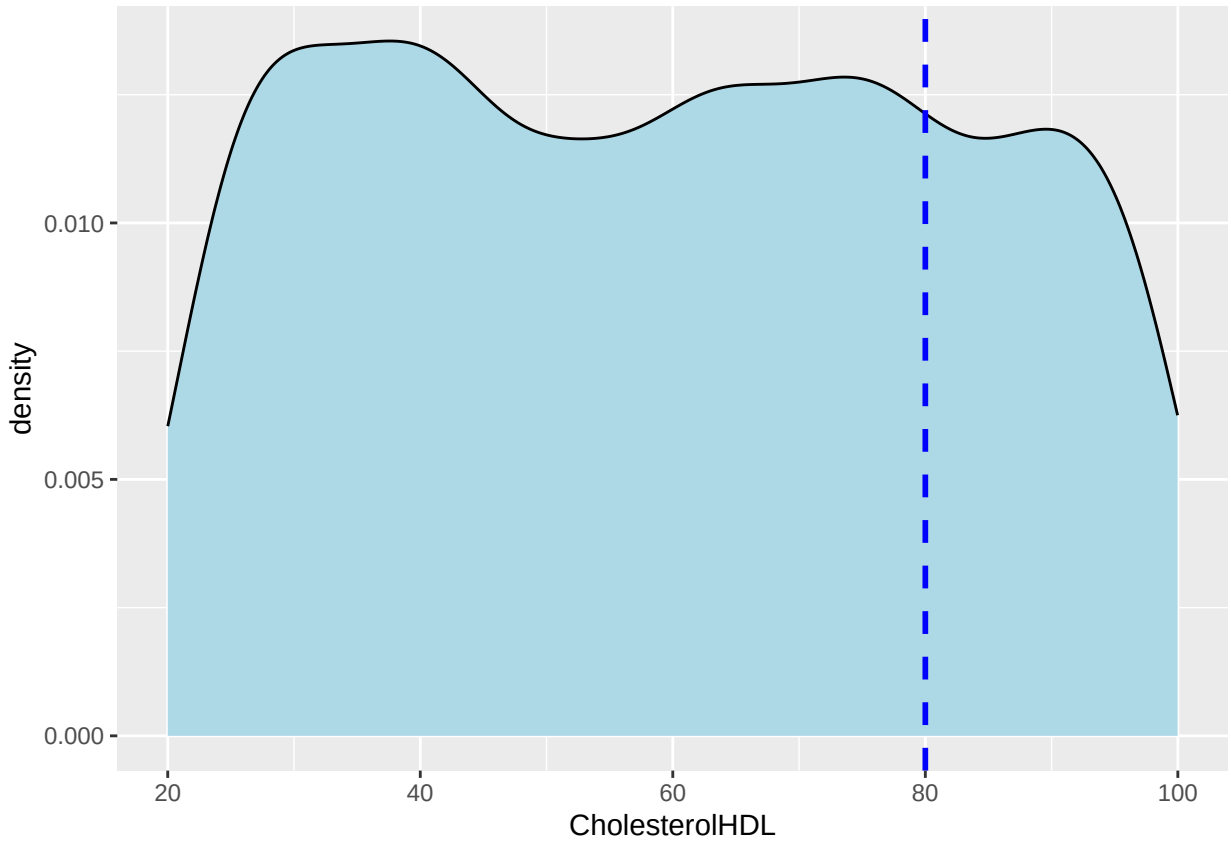
```
library(ggplot2)
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v dplyr     1.1.4     v readr     2.1.5
## v forcats   1.0.0     v stringr   1.5.1
## v lubridate 1.9.4     v tibble    3.2.1
## v purrr     1.0.4     v tidyr     1.3.1
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
data <- read_rds('data.rds')
p <- ggplot(data, aes(x=CholesterolHDL)) +
  geom_density(color="black", fill = 'lightblue') +
  geom_vline(aes(xintercept=80), color = 'blue', linetype = 'dashed', size = 1)
```

```
## Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.
## i Please use `linewidth` instead.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.
```

```
print(p)
```



The figure above shows a density of patient's LDL cholesterol level.

```
data$HDLlevel <- data$CholesterolHDL
data$HDLlevel <- ifelse(data$HDLlevel>80,'Above80','Below80')
t <- table(data$HDLlevel)
print(t)
```

```
##
## Above80 Below80
##     509    1640
```

There are 509 participants with high HDL cholesterol level and 1640 participants with HDL cholesterol level lower than 80 mg/dL. As the dataset is a large sample, normal distribution is applied for this test.

1. Confusion and HDL Cholesterol

Null Hypothesis: The proportion of developing confusion among participants whose HDL cholesterol density is greater than 80 mg/dL is equal to the proportion among participants whose HDL cholesterol density is lower than 80 mg/dL

Hypothesis Testing

```
highHDL <- data[data$HDLlevel == 'Above80',]
optimalHDL <- data[data$HDLlevel == 'Below80',]

a <- nrow(highHDL[highHDL$Confusion==1,])
b <- nrow(highHDL[highHDL$Confusion==0,])
c <- nrow(optimalHDL[optimalHDL$Confusion==1,])
```

```r
d <- nrow(optimalHDL[optimalHDL$Confusion==0,])

confusion <- matrix(c(a,b,c,d),nrow=2,byrow=T)
rownames(confusion) <- c('HDL above 80 mg/dL', 'HDL below 80 mg/dL')
colnames(confusion) <- c('Confusion', 'No confusion')
cat("\nContingency Table:\n")
```

```
##
## Contingency Table:
```

```r
print(confusion)
```

```
##                     Confusion No confusion
## HDL above 80 mg/dL        105          404
## HDL below 80 mg/dL        336         1304
```

```r
n1 <- a+b
n2 <- c+d

p1 <- a/(a+b)
p2 <- c/(c+d)

S <- sqrt((p1*(1-p1)/n1)+(p2*(1-p2)/n2))
z0 <- (p1-p2)/S
cv <- qnorm(0.95)

if (z0>cv) {
  cat('Reject null hypothesis: High HDL cholesterol level is associated with development of confusion\n
} else {
  cat('Accept null hypothesis: High HDL cholesterol level is not associated with development of confusi
}
```

```
## Accept null hypothesis: High HDL cholesterol level is not associated with development of confusion
```

#2. Disorientation
Nul Hypothesis: The participants with high HDL cholesterol density are equally likely to have disorientation than the patients with optimal HDL cholesterol density.
Hypothesis Testing

```r
a <- nrow(highHDL[highHDL$Disorientation==1,])
b <- nrow(highHDL[highHDL$Disorientation==0,])
c <- nrow(optimalHDL[optimalHDL$Disorientation==1,])
d <- nrow(optimalHDL[optimalHDL$Disorientation==0,])

disorientation <- matrix(c(a,b,c,d),nrow=2,byrow=T)
rownames(disorientation) <- c('HDL above 80 mg/dL', 'HDL below 80 mg/dL')
colnames(disorientation) <- c('Disorientation', 'Disorientation')
cat("\nContingency Table:\n")
```

```
##
## Contingency Table:
```

```r
print(disorientation)
```

```
##                     Disorientation Disorientation
## HDL above 80 mg/dL              91            418
## HDL below 80 mg/dL             249           1391
```

```r
n1 <- a+b
n2 <- c+d

p1 <- a/(a+b)
p2 <- c/(c+d)

S <- sqrt((p1*(1-p1)/n1)+(p2*(1-p2)/n2))
z0 <- (p1-p2)/S
cv <- qnorm(0.95)

if (z0>cv) {
  cat('Reject null hypothesis: High HDL cholesterol level is associated with development of disorientati
} else {
  cat('Accept null hypothesis: High HDL cholesterol level is not associated with development of disorier
}
```

```
## Accept null hypothesis: High HDL cholesterol level is not associated with development of disorientati
```

#3. Personality Changes
Nul Hypothesis: The participants with high HDL cholesterol density are equally likely to have personality changes than the patients with optimal HDL cholesterol density.
Hypothesis Testing

```r
a <- nrow(highHDL[highHDL$PersonalityChanges==1,])
b <- nrow(highHDL[highHDL$PersonalityChanges==0,])
c <- nrow(optimalHDL[optimalHDL$PersonalityChanges==1,])
d <- nrow(optimalHDL[optimalHDL$PersonalityChanges==0,])

personalitychanges <- matrix(c(a,b,c,d),nrow=2,byrow=T)
rownames(personalitychanges) <- c('HDL above 80 mg/dL', 'HDL below 80 mg/dL')
colnames(personalitychanges) <- c('Personality changes', 'No personality changes')
cat("\nContingency Table:\n")
```

```
##
## Contingency Table:
```

```r
print(personalitychanges)
```

```
##                     Personality changes No personality changes
## HDL above 80 mg/dL                   61                     448
## HDL below 80 mg/dL                  263                    1377
```

```r
n1 <- a+b
n2 <- c+d

p1 <- a/(a+b)
p2 <- c/(c+d)

S <- sqrt((p1*(1-p1)/n1)+(p2*(1-p2)/n2))
z0 <- (p1-p2)/S
cv <- qnorm(0.95)

if (z0>cv) {
  cat('Reject null hypothesis: High HDL cholesterol level is associated with personality changes\n')
} else {
  cat('Accept null hypothesis: High HDL cholesterol level is not associated with personality changes\n')
```

```
}
```

```
## Accept null hypothesis: High HDL cholesterol level is not associated with personality changes
```

#4. Difficulty in completing tasks
Null Hypothesis: The participants with high HDL cholesterol density are equally likely to have difficulty in completing tasks than the patients with optimal HDL cholesterol density.
Hypothesis Testing

```
a <- nrow(highHDL[highHDL$DifficultyCompletingTasks==1,])
b <- nrow(highHDL[highHDL$DifficultyCompletingTasks==0,])
c <- nrow(optimalHDL[optimalHDL$DifficultyCompletingTasks==1,])
d <- nrow(optimalHDL[optimalHDL$DifficultyCompletingTasks==0,])

diffcompletingtasks <- matrix(c(a,b,c,d),nrow=2,byrow=T)
rownames(diffcompletingtasks) <- c('HDL above 80 mg/dL', 'HDL below 80 mg/dL')
colnames(diffcompletingtasks) <- c('Difficulty in completing tasks', 'No difficulty in completing tasks
cat("\nContingency Table:\n")
```

```
##
## Contingency Table:
```

```
print(diffcompletingtasks)
```

```
##                     Difficulty in completing tasks
## HDL above 80 mg/dL                              65
## HDL below 80 mg/dL                             276
##                     No difficulty in completing tasks
## HDL above 80 mg/dL                               444
## HDL below 80 mg/dL                              1364
```

```
n1 <- a+b
n2 <- c+d

p1 <- a/(a+b)
p2 <- c/(c+d)

S <- sqrt((p1*(1-p1)/n1)+(p2*(1-p2)/n2))
z0 <- (p1-p2)/S
cv <- qnorm(0.95)

if (z0>cv) {
  cat('Reject null hypothesis: High HDL cholesterol level is associated with difficulty in completing ta
} else {
  cat('Accept null hypothesis: High HDL cholesterol level is not associated with difficulty in completi
}
```

```
## Accept null hypothesis: High HDL cholesterol level is not associated with difficulty in completing ta
```

#5. Forgetfulness
Null Hypothesis: The participants with high HDL cholesterol density are equally likely to develop forgetfulness than the participants with optimal HDL cholesterol density.
Hypothesis Testing

```
a <- nrow(highHDL[highHDL$Forgetfulness==1,])
b <- nrow(highHDL[highHDL$Forgetfulness==0,])
c <- nrow(optimalHDL[optimalHDL$Forgetfulness==1,])
d <- nrow(optimalHDL[optimalHDL$Forgetfulness==0,])
```

```r
forgetfulness <- matrix(c(a,b,c,d),nrow=2,byrow=T)
rownames(forgetfulness) <- c('HDL above 80 mg/dL', 'HDL below 80 mg/dL')
colnames(forgetfulness) <- c('Forgetfulness', 'No forgetfulness')
cat("\nContingency Table:\n")
```

```
##
## Contingency Table:
```

```r
print(forgetfulness)
```

```
##                     Forgetfulness No forgetfulness
## HDL above 80 mg/dL            143               366
## HDL below 80 mg/dL            505              1135
```

```r
n1 <- a+b
n2 <- c+d

p1 <- a/(a+b)
p2 <- c/(c+d)

S <- sqrt((p1*(1-p1)/n1)+(p2*(1-p2)/n2))
z0 <- (p1-p2)/S
cv <- qnorm(0.95)

if (z0>cv) {
  cat('Reject null hypothesis: High HDL cholesterol level is associated with development of forgetfulnes
} else {
  cat('Accept null hypothesis: High HDL cholesterol level is not associated with development of forgetfu
}
```

```
## Accept null hypothesis: High HDL cholesterol level is not associated with development of forgetfulnes
```

Conclusion

The test concludes that the high density of HDL cholesterol is associated with development of five major symptoms.