

# R Notebook

Loading required libraries and packages.

Setting up roi rsCVR dataframe.

Setting up demographic and other information dataframes.

```
group_list = read_excel(paste0(wd, 'HIV_toba_resting_160.xlsx'), na="na")
# group_list$`HIVDuration(mons)` <- as.numeric(as.character(group_list$`HIVDuration(mons)`))
subset_group_list <- group_list[which(group_list$folder%in%as.character(subject_names$V1)), ]
row.names(subset_group_list) <- subset_group_list$folder
```

## Warning: Setting row names on a tibble is deprecated.

Set up final combined df including all Variables of Interest.

```
final_df <- merge(results, subset_group_list, by="row.names")
final_df_new <- final_df %>%
  mutate(Smoker = if_else(grepl("_S", final_df$group, fixed = TRUE), 'smoker', 'non-smoker')) %>%
  mutate(Disease = if_else(grepl("HIV", final_df$group, fixed = TRUE), 'HIV', 'SN'))
```

ANOVA results.

```
anova_results <- purrr::map(final_df_new[,2:12], ~aov(.x ~ Disease * Smoker, data = final_df_new))
sapply(anova_results, function(x) summary(x))
```

```
## $Frontal
##               Df Sum Sq Mean Sq F value Pr(>F)
## Disease         1 0.01286 0.012859   0.8961 0.34529
## Smoker           1 0.08817 0.088170   6.1445 0.01425 *
## Disease:Smoker   1 0.01253 0.012528   0.8731 0.35154
## Residuals       156 2.23850 0.014349
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $Parietal
##               Df Sum Sq Mean Sq F value Pr(>F)
## Disease         1 0.0074 0.007431   0.2554 0.6140
## Smoker           1 0.0415 0.041458   1.4251 0.2344
## Disease:Smoker   1 0.0063 0.006277   0.2158 0.6429
## Residuals       156 4.5382 0.029091
##
## $Temporal
##               Df Sum Sq Mean Sq F value Pr(>F)
## Disease         1 0.0762 0.076245   3.7412 0.0549 .
## Smoker           1 0.0117 0.011682   0.5732 0.4501
## Disease:Smoker   1 0.0460 0.045995   2.2569 0.1350
## Residuals       156 3.1793 0.020380
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```

##
## $Limbic
##           Df Sum Sq Mean Sq F value    Pr(>F)
## Disease      1  0.13176  0.131762   7.2594 0.007826 **
## Smoker        1  0.07780  0.077795   4.2862 0.040072 *
## Disease:Smoker 1  0.02051  0.020510   1.1300 0.289415
## Residuals    156  2.83146  0.018150
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $Occipital
##           Df Sum Sq Mean Sq F value    Pr(>F)
## Disease      1  0.2405  0.240517   3.2087 0.07519 .
## Smoker        1  0.0067  0.006700   0.0894 0.76536
## Disease:Smoker 1  0.0319  0.031911   0.4257 0.51506
## Residuals    156 11.6935  0.074959
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $Insula
##           Df Sum Sq Mean Sq F value    Pr(>F)
## Disease      1  0.0904  0.090361   3.1041 0.08005 .
## Smoker        1  0.0650  0.065026   2.2338 0.13705
## Disease:Smoker 1  0.0076  0.007557   0.2596 0.61111
## Residuals    156  4.5412  0.029110
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $BasalGanglia
##           Df Sum Sq Mean Sq F value    Pr(>F)
## Disease      1  0.0688  0.068821   2.0706 0.15217
## Smoker        1  0.0006  0.000629   0.0189 0.89077
## Disease:Smoker 1  0.1502  0.150212   4.5193 0.03509 *
## Residuals    156  5.1851  0.033238
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $Thalamus
##           Df Sum Sq Mean Sq F value    Pr(>F)
## Disease      1  0.8238  0.82376 12.5155 0.0005319 ***
## Smoker        1  0.0005  0.00052   0.0079 0.9295056
## Disease:Smoker 1  0.5866  0.58658   8.9120 0.0032903 **
## Residuals    156 10.2678  0.06582
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $BasalForebrain
##           Df Sum Sq Mean Sq F value    Pr(>F)
## Disease      1  0.4909  0.49090   9.2931 0.002703 **
## Smoker        1  0.0008  0.00080   0.0151 0.902199
## Disease:Smoker 1  0.0670  0.06697   1.2678 0.261904
## Residuals    156  8.2406  0.05282
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

```
##
## $Midbrain
##           Df Sum Sq Mean Sq F value Pr(>F)
## Disease      1  0.2860  0.286032   3.6112 0.05924 .
## Smoker        1  0.0357  0.035651   0.4501 0.50328
## Disease:Smoker 1  0.0424  0.042356   0.5347 0.46572
## Residuals    156 12.3564  0.079208
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $Cerebellum
##           Df Sum Sq Mean Sq F value Pr(>F)
## Disease      1  0.0042  0.004229   0.0697 0.7922
## Smoker        1  0.0185  0.018501   0.3048 0.5817
## Disease:Smoker 1  0.0221  0.022150   0.3649 0.5467
## Residuals    156  9.4693  0.060701
```

#### Tukey Results

```
# Limbic
```

```
TukeyHSD(anova_results[[4]])
```

```
## Tukey multiple comparisons of means
## 95% family-wise confidence level
##
## Fit: aov(formula = .x ~ Disease * Smoker, data = final_df_new)
##
## $Disease
##           diff           lwr           upr           p adj
## SN-HIV -0.05741167 -0.09950167 -0.01532168 0.0078255
##
## $Smoker
##           diff           lwr           upr           p adj
## smoker-non-smoker 0.04477207 0.002035741 0.0875084 0.0401597
##
## $`Disease:Smoker`
##           diff           lwr           upr           p adj
## SN:non-smoker-HIV:non-smoker -0.03714082 -0.10931313 0.035031481 0.5412252
## HIV:smoker-HIV:non-smoker      0.06701486 -0.01109907 0.145128785 0.1202353
## SN:smoker-HIV:non-smoker      -0.01617378 -0.09712475 0.064777180 0.9544616
## HIV:smoker-SN:non-smoker      0.10415568  0.02604175 0.182269607 0.0038174
## SN:smoker-SN:non-smoker      0.02096704 -0.05998392 0.101918003 0.9072951
## SN:smoker-HIV:smoker        -0.08318864 -0.16947886 0.003101583 0.0632528
```

```
# Insula
```

```
TukeyHSD(anova_results[[6]])
```

```
## Tukey multiple comparisons of means
## 95% family-wise confidence level
##
## Fit: aov(formula = .x ~ Disease * Smoker, data = final_df_new)
##
## $Disease
##           diff           lwr           upr           p adj
## SN-HIV -0.04754416 -0.1008482 0.005759906 0.0800546
##
```

```
## $Smoker
##               diff               lwr               upr               p adj
## smoker-non-smoker 0.04093293 -0.01318968 0.09505554 0.137219
##
## $`Disease:Smoker`
##               diff               lwr               upr               p adj
## SN:non-smoker-HIV:non-smoker -0.034834824 -0.126236059 0.05656641 0.7554944
## HIV:smoker-HIV:non-smoker      0.054446836 -0.044479056 0.15337273 0.4830503
## SN:smoker-HIV:non-smoker      -0.008339574 -0.110858377 0.09417923 0.9966550
## HIV:smoker-SN:non-smoker       0.089281660 -0.009644232 0.18820755 0.0926835
## SN:smoker-SN:non-smoker        0.026495249 -0.076023553 0.12901405 0.9078350
## SN:smoker-HIV:smoker          -0.062786410 -0.172067017 0.04649420 0.4447387
```

*# Thalamus*

```
TukeyHSD(anova_results[[8]])
```

```
## Tukey multiple comparisons of means
## 95% family-wise confidence level
##
## Fit: aov(formula = .x ~ Disease * Smoker, data = final_df_new)
##
## $Disease
##               diff               lwr               upr               p adj
## SN-HIV -0.143551 -0.2237027 -0.06339935 0.0005319
##
## $Smoker
##               diff               lwr               upr               p adj
## smoker-non-smoker -0.003649129 -0.08503159 0.07773333 0.9295369
##
## $`Disease:Smoker`
##               diff               lwr               upr               p adj
## SN:non-smoker-HIV:non-smoker -0.04229731 -0.179734466 0.09513985 0.8546797
## HIV:smoker-HIV:non-smoker      0.11508485 -0.033666898 0.26383660 0.1890174
## SN:smoker-HIV:non-smoker      -0.17346843 -0.327622726 -0.01931414 0.0206121
## HIV:smoker-SN:non-smoker       0.15738216 0.008630408 0.30613391 0.0335934
## SN:smoker-SN:non-smoker        -0.13117112 -0.285325420 0.02298317 0.1250899
## SN:smoker-HIV:smoker          -0.28855328 -0.452875089 -0.12423147 0.0000603
```

Regression Analysis

```
lm_results_v1 <- purrr::map(final_df_new[,2:12], ~lm(.x ~ final_df_new$sex, data = final_df_new))
lm_results_v2 <- purrr::map(final_df_new[,2:12], ~lm(.x ~ final_df_new$age, data = final_df_new))
lm_results_v3 <- purrr::map(final_df_new[,2:12], ~lm(.x ~ final_df_new$age + final_df_new$sex, data = f
lm_results_v4 <- purrr::map(final_df_new[,2:12], ~lm(.x ~ final_df_new$age + final_df_new$sex + final_d
lm_results_v5 <- purrr::map(final_df_new[,2:12], ~lm(.x ~ final_df_new$age + final_df_new$sex + final_d
lm_results_v6 <- purrr::map(final_df_new[,2:12], ~lm(.x ~ final_df_new$age + final_df_new$sex + final_d
lm_results_v7 <- purrr::map(final_df_new[,2:12], ~lm(.x ~ final_df_new$age + final_df_new$sex + final_d
lm_results_v8 <- purrr::map(final_df_new[,2:12], ~lm(.x ~ final_df_new$age + final_df_new$sex + final_d
```

```

lm_results_v9 <- purrr::map(final_df_new[,2:12], ~lm(.x ~ final_df_new$age + final_df_new$sex + final_d
lm_results_v10 <- purrr::map(final_df_new[,2:12], ~lm(.x ~ final_df_new$age + final_df_new$sex + final_c
lm_results_v11 <- purrr::map(final_df_new[,2:12], ~lm(.x ~ final_df_new$age + final_df_new$sex * final_c
lm_results_v12 <- purrr::map(final_df_new[,2:12], ~lm(.x ~ final_df_new$age + final_df_new$sex * final_c
lm_results_v13 <- purrr::map(final_df_new[,2:12], ~lm(.x ~ final_df_new$sex + final_df_new$age * final_c
lm_results_v14 <- purrr::map(final_df_new[,2:12], ~lm(.x ~ final_df_new$sex + final_df_new$age * final_c
lm_results_v15 <- purrr::map(final_df_new[,2:12], ~lm(.x ~ final_df_new$age * final_df_new$sex * final_c
lm_results_v16 <- purrr::map(final_df_new[,2:12], ~lm(.x ~ final_df_new$age * final_df_new$sex * final_c

```

Comparing models

```

AIC(lm_results_v1$Thalamus, lm_results_v2$Thalamus, lm_results_v3$Thalamus, lm_results_v4$Thalamus, lm_

```

```

## Warning in AIC.default(lm_results_v1$Thalamus, lm_results_v2$Thalamus,
## lm_results_v3$Thalamus, : models are not all fitted to the same number of
## observations

```

| ##                          | df | AIC       |
|-----------------------------|----|-----------|
| ## lm_results_v1\$Thalamus  | 3  | 41.250425 |
| ## lm_results_v2\$Thalamus  | 3  | 40.383446 |
| ## lm_results_v3\$Thalamus  | 4  | 42.311505 |
| ## lm_results_v4\$Thalamus  | 5  | 32.442189 |
| ## lm_results_v5\$Thalamus  | 5  | 31.584173 |
| ## lm_results_v6\$Thalamus  | 5  | 33.473316 |
| ## lm_results_v7\$Thalamus  | 6  | 12.873026 |
| ## lm_results_v8\$Thalamus  | 6  | 15.378239 |
| ## lm_results_v9\$Thalamus  | 7  | 7.930861  |
| ## lm_results_v10\$Thalamus | 7  | 15.088937 |
| ## lm_results_v11\$Thalamus | 8  | 1.626101  |
| ## lm_results_v12\$Thalamus | 8  | 9.764963  |
| ## lm_results_v13\$Thalamus | 10 | 10.174675 |
| ## lm_results_v14\$Thalamus | 10 | 19.384451 |
| ## lm_results_v15\$Thalamus | 11 | 1.566430  |
| ## lm_results_v16\$Thalamus | 11 | 13.811104 |

```

BIC(lm_results_v1$Thalamus, lm_results_v2$Thalamus, lm_results_v3$Thalamus, lm_results_v4$Thalamus, lm_

```

```

## Warning in BIC.default(lm_results_v1$Thalamus, lm_results_v2$Thalamus,
## lm_results_v3$Thalamus, : models are not all fitted to the same number of
## observations

```

| ##                         | df | BIC      |
|----------------------------|----|----------|
| ## lm_results_v1\$Thalamus | 3  | 50.47595 |
| ## lm_results_v2\$Thalamus | 3  | 49.60897 |
| ## lm_results_v3\$Thalamus | 4  | 54.61220 |
| ## lm_results_v4\$Thalamus | 5  | 44.47579 |
| ## lm_results_v5\$Thalamus | 5  | 42.37859 |
| ## lm_results_v6\$Thalamus | 5  | 44.34525 |
| ## lm_results_v7\$Thalamus | 6  | 22.03119 |

```
## lm_results_v8$Thalamus 6 24.53640
## lm_results_v9$Thalamus 7 18.61538
## lm_results_v10$Thalamus 7 25.77346
## lm_results_v11$Thalamus 8 13.83698
## lm_results_v12$Thalamus 8 21.97585
## lm_results_v13$Thalamus 10 25.43828
## lm_results_v14$Thalamus 10 34.64806
## lm_results_v15$Thalamus 11 18.35640
## lm_results_v16$Thalamus 11 30.60107
```

#### # Model 9

```
summary(lm_results_v9$Thalamus)
```

```
##
## Call:
## lm(formula = .x ~ final_df_new$age + final_df_new$sex + final_df_new$`HIVDuration(mons)` *
##     final_df_new$Nic_lifetime_use, data = final_df_new)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.61803 -0.09634 -0.01444  0.10478  0.44784
##
## Coefficients:
##                                     Estimate
## (Intercept)                        7.890e-01
## final_df_new$age                   -2.919e-03
## final_df_new$sexM                   6.996e-02
## final_df_new$`HIVDuration(mons)`    1.175e-03
## final_df_new$Nic_lifetime_use       1.460e-06
## final_df_new$`HIVDuration(mons)`:final_df_new$Nic_lifetime_use -6.064e-09
##                                     Std. Error
## (Intercept)                        2.622e-01
## final_df_new$age                   5.210e-03
## final_df_new$sexM                   1.855e-01
## final_df_new$`HIVDuration(mons)`    7.447e-04
## final_df_new$Nic_lifetime_use       4.886e-07
## final_df_new$`HIVDuration(mons)`:final_df_new$Nic_lifetime_use 2.408e-09
##                                     t value Pr(>|t|)
## (Intercept)                        3.009  0.00550
## final_df_new$age                   -0.560  0.57975
## final_df_new$sexM                    0.377  0.70893
## final_df_new$`HIVDuration(mons)`    1.577  0.12593
## final_df_new$Nic_lifetime_use       2.989  0.00577
## final_df_new$`HIVDuration(mons)`:final_df_new$Nic_lifetime_use -2.518  0.01778
##
## (Intercept)                        **
## final_df_new$age
## final_df_new$sexM
## final_df_new$`HIVDuration(mons)`
## final_df_new$Nic_lifetime_use      **
## final_df_new$`HIVDuration(mons)`:final_df_new$Nic_lifetime_use *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2439 on 28 degrees of freedom
```

```

## (126 observations deleted due to missingness)
## Multiple R-squared: 0.2615, Adjusted R-squared: 0.1296
## F-statistic: 1.983 on 5 and 28 DF, p-value: 0.1122

# Model 11
summary(lm_results_v9$Thalamus)

##
## Call:
## lm(formula = .x ~ final_df_new$age + final_df_new$sex + final_df_new$`HIVDuration(mons)` *
##     final_df_new$Nic_lifetime_use, data = final_df_new)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.61803 -0.09634 -0.01444  0.10478  0.44784
##
## Coefficients:
##                                     Estimate
## (Intercept)                        7.890e-01
## final_df_new$age                   -2.919e-03
## final_df_new$sexM                   6.996e-02
## final_df_new$`HIVDuration(mons)`    1.175e-03
## final_df_new$Nic_lifetime_use       1.460e-06
## final_df_new$`HIVDuration(mons)`:final_df_new$Nic_lifetime_use -6.064e-09
##                                     Std. Error
## (Intercept)                        2.622e-01
## final_df_new$age                   5.210e-03
## final_df_new$sexM                   1.855e-01
## final_df_new$`HIVDuration(mons)`    7.447e-04
## final_df_new$Nic_lifetime_use       4.886e-07
## final_df_new$`HIVDuration(mons)`:final_df_new$Nic_lifetime_use 2.408e-09
##                                     t value Pr(>|t|)
## (Intercept)                        3.009 0.00550
## final_df_new$age                   -0.560 0.57975
## final_df_new$sexM                   0.377 0.70893
## final_df_new$`HIVDuration(mons)`    1.577 0.12593
## final_df_new$Nic_lifetime_use       2.989 0.00577
## final_df_new$`HIVDuration(mons)`:final_df_new$Nic_lifetime_use -2.518 0.01778
##
## (Intercept)                        **
## final_df_new$age
## final_df_new$sexM
## final_df_new$`HIVDuration(mons)`
## final_df_new$Nic_lifetime_use      **
## final_df_new$`HIVDuration(mons)`:final_df_new$Nic_lifetime_use *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2439 on 28 degrees of freedom
## (126 observations deleted due to missingness)
## Multiple R-squared: 0.2615, Adjusted R-squared: 0.1296
## F-statistic: 1.983 on 5 and 28 DF, p-value: 0.1122

```