ResultsT

Freya Acar

4/11/2018

Overview

1. Introduction
2. Startup
3. Descriptives
4. Analysis
5. Posthoc
6. Correlations between T1 and T2
7. Discussion

Introduction

We have a dataset where FreeSurfer was used to determine cortical and subcortical brain anatomy in cis- and transgender population. For every participant two scans were administered (T1 and T2) and an average of both scans was computed. In this report we analyze the data from one scan and the average of both scans. Later on we show the advantage and increase in power obtained by admninistering 2 scans.
One participant (P22) was removed from the analysis because no anatomical data was available for this participant.

Startup

We first read in the data. In data.all all available measurements are stored (descriptive and anatomical) while in data.hyp the anatomical results for every region of interest stored. Both files contain data of T1, T2 and the average.

```
# Libraries
library(knitr)

# Read in data
data.all <- read.csv("../1.Data/Behzad_all.csv", sep=";", dec=",")
data.hyp <- read.csv("../1.Data/Behzad_hyp.csv", sep=";", dec=",")

# Check data
dim(data.hyp)
```

```
## [1] 140  68
```

```
dim(data.all)
```

```
## [1] 140 820
```

The regions we are interested in are the cerebellum, caudate, putamen, nucleus accumenbens, thalamus, fusiform, pre-central gyrus, post-central gyrus, frontal poles and inferior parietal gyrus. Here we list the variables we selected from FreeSurfer that comply with these regions.

```
# Regions of interest
names(data.hyp[,47:68])
```

```
##  [1] "Tavg_L_fusiform_volume"          "Tavg_L_inferiorparietal_volume"
##  [3] "Tavg_L_postcentral_volume"       "Tavg_L_precentral_volume"
##  [5] "Tavg_L_frontalpole_volume"       "Tavg_R_fusiform_volume"
##  [7] "Tavg_R_inferiorparietal_volume"  "Tavg_R_postcentral_volume"
##  [9] "Tavg_R_precentral_volume"        "Tavg_R_frontalpole_volume"
## [11] "Tavg_LeftCerebellumWhiteMatter"  "Tavg_LeftCerebellumCortex"
## [13] "Tavg_RightCerebellumWhiteMatter" "Tavg_RightCerebellumCortex"
## [15] "Tavg_LeftThalamusProper"         "Tavg_LeftCaudate"
## [17] "Tavg_LeftPutamen"                "Tavg_LeftAccumbensarea"
## [19] "Tavg_RightThalamusProper"        "Tavg_RightCaudate"
## [21] "Tavg_RightPutamen"               "Tavg_RightAccumbensarea"
```

Descriptives

In this section population parameters are presented.

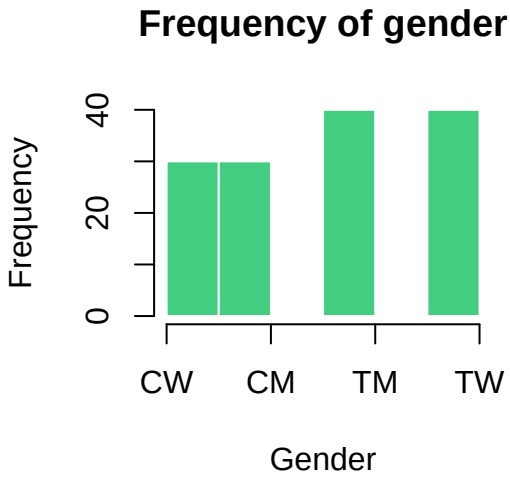
Gender

There were 60 cisgender and 80 transgender participants. One participant (P22) was removed from the analysis because no anatomical data was available for this participant.

```
ng1 <- sum(data.all[,2]==1)
ng2 <- sum(data.all[,2]==2)
ng3 <- sum(data.all[,2]==3)
ng4 <- sum(data.all[,2]==4)

paste("Cisgender women = ", ng1, ", cisgender men = ", ng2,
      ", transgender men = ", ng3, ", transgender women = ",ng4, sep="")
```

```
## [1] "Cisgender women = 30, cisgender men = 30, transgender men = 40, transgender women = 40"
```

```
hist(data.all[,2], breaks = 5, xlab = "Gender", xaxt = "n", col = "seagreen3",
     border = "white", main = "Frequency of gender")
axis(1,at=c(1:4),labels=c("CW", "CM", "TM", "TW"))
```

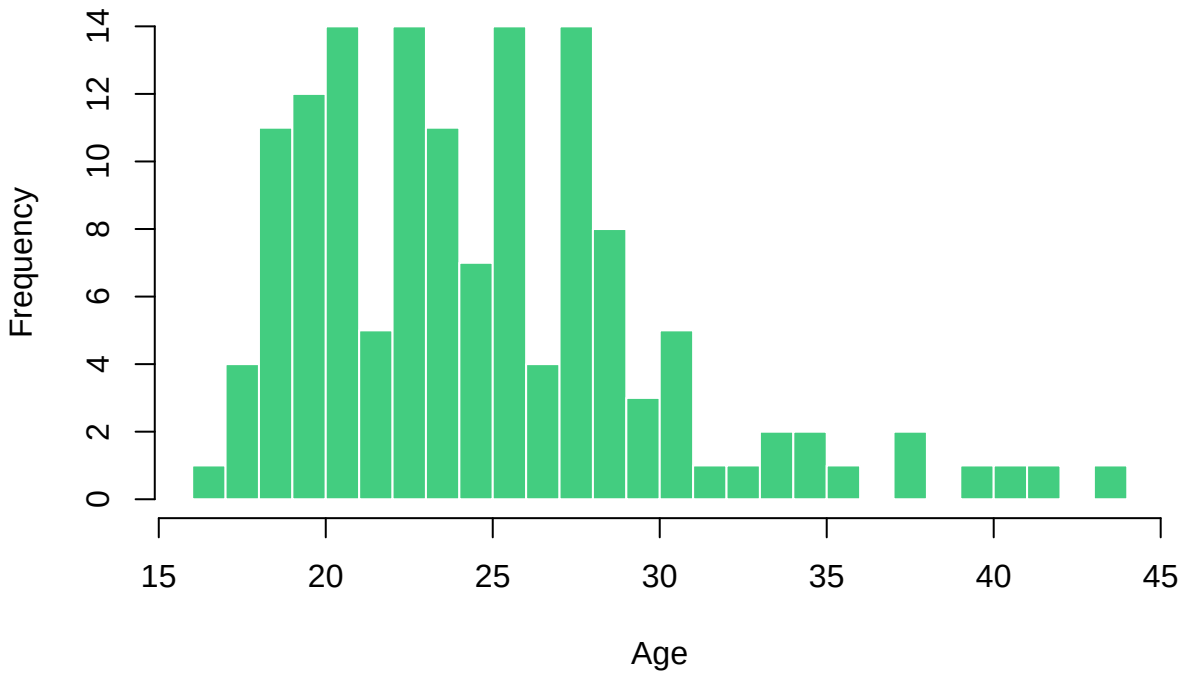


Age

```
summary(data.all[,4])
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   16.00   21.00   24.00   25.18   28.00   44.00
```

```
hist(data.all[,4], xlab = "Age", col = "seagreen3", border = "white",
     main = "Histogram of age distribution", breaks = length(unique(data.all[,4])))
```



Age of the participants ranged from 16 to 44. If we look at the distribution of age in the cis- and transgender group we see that the range is similar in both groups.

```
# Cisgender women
mean(data.all[data.all[,2]==1,4])
```

```
## [1] 25.8
```

```
sd(data.all[data.all[,2]==1,4])
```

```
## [1] 3.836306
```

```
# Cisgender men
mean(data.all[data.all[,2]==2,4])
```

```
## [1] 26.03333
```

```
sd(data.all[data.all[,2]==2,4])
```

```
## [1] 5.25543
```

```
# Transgender men
mean(data.all[data.all[,2]==3,4])
```

```
## [1] 24.375
```

```
sd(data.all[data.all[,2]==3,4])
```

```
## [1] 5.352629
```

```
# Transgender women
mean(data.all[data.all[,2]==4,4])
```

```
## [1] 24.875
```

```r
sd(data.all[data.all[,2]==4,4])
```

```
## [1] 6.202512
```

Analysis of the data

Analysis of one measurement

We compute an ANOVA on all hypothesis regions with the data from T1.

```
bg.one <- 3
nd.one <- 24
ln <- nd.one-bg.one
fac <- c(rep("CW",30), rep("CM",30), rep("TM",40), rep("TW",40))  # factor for participant group

# Object to save p-values of ANOVA
pan.one <- array(data=NA, dim = ln)

# Compute ANOVA for every predictor and save p-value
for(i in bg.one:nd.one){
  tempan <- aov(data.hyp[,i] ~ as.factor(data.hyp[,2]))
  pan.one[i-bg.one+1] <- unlist(summary(tempan))[9]
}

# FDR correction on p-values to correct for multiple testing
pancorr.one <- p.adjust(pan.one, method = "fdr")
sum(pancorr.one < 0.05)
```

```
## [1] 20
```

```
kable(cbind(names(data.hyp[,bg.one:nd.one]), round(pancorr.one, 3)))
```

T1_L_fusiform_volume	0
T1_L_inferiorparietal_volume	0.001
T1_L_postcentral_volume	0.002
T1_L_precentral_volume	0.032
T1_L_frontalpole_volume	0.006
T1_R_fusiform_volume	0
T1_R_inferiorparietal_volume	0.001
T1_R_postcentral_volume	0.014
T1_R_precentral_volume	0.002
T1_R_frontalpole_volume	0.001
T1_LeftCerebellumWhiteMatter	0.001
T1_LeftCerebellumCortex	0
T1_RightCerebellumWhiteMatter	0.001
T1_RightCerebellumCortex	0
T1_LeftThalamusProper	0
T1_LeftCaudate	0.002
T1_LeftPutamen	0.006
T1_LeftAccumbensarea	0.453
T1_RightThalamusProper	0
T1_RightCaudate	0.002
T1_RightPutamen	0
T1_RightAccumbensarea	0.119

Analysis of the average

We computed the same ANOVA on the average.

```
bg.avg <- 47
nd.avg <- 68
ln <- nd.avg-bg.avg+1
fac <- c(rep("CW",30), rep("CM",30), rep("TM",40), rep("TW",40))  # factor for participant group

# Object to save p-values of ANOVA
pan.avg <- array(data=NA, dim = ln)

# Compute ANOVA for every predictor and save p-value
for(i in bg.avg:nd.avg){
  tempan <- aov(data.hyp[,i] ~ as.factor(data.hyp[,2]))
  pan.avg[i-bg.avg+1] <- unlist(summary(tempan))[9]
}

# FDR correction on p-values to correct for multiple testing
pancorr.avg <- p.adjust(pan.avg, method = "fdr")
sum(pancorr.avg < 0.05)
```

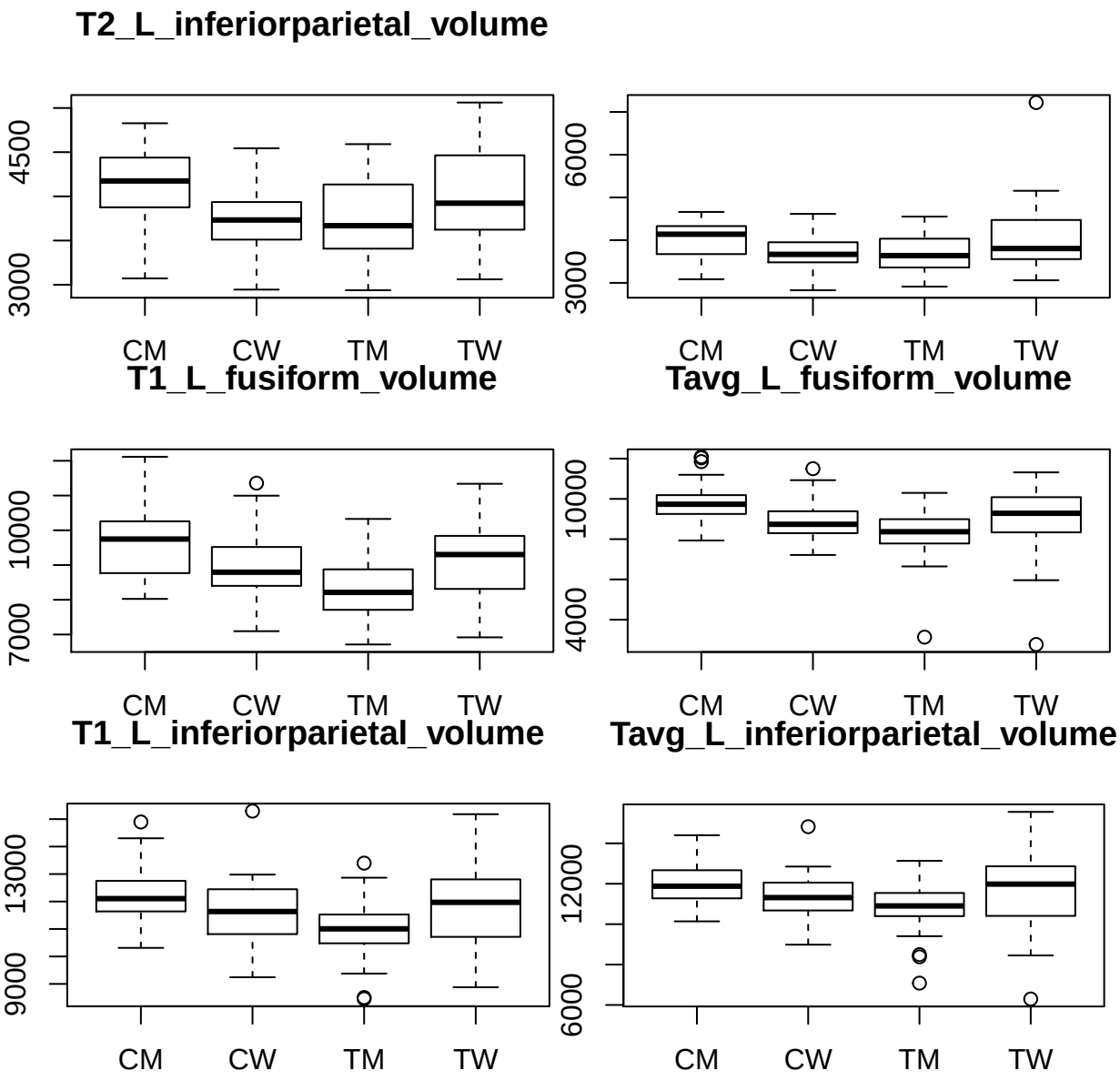
```
## [1] 19
```

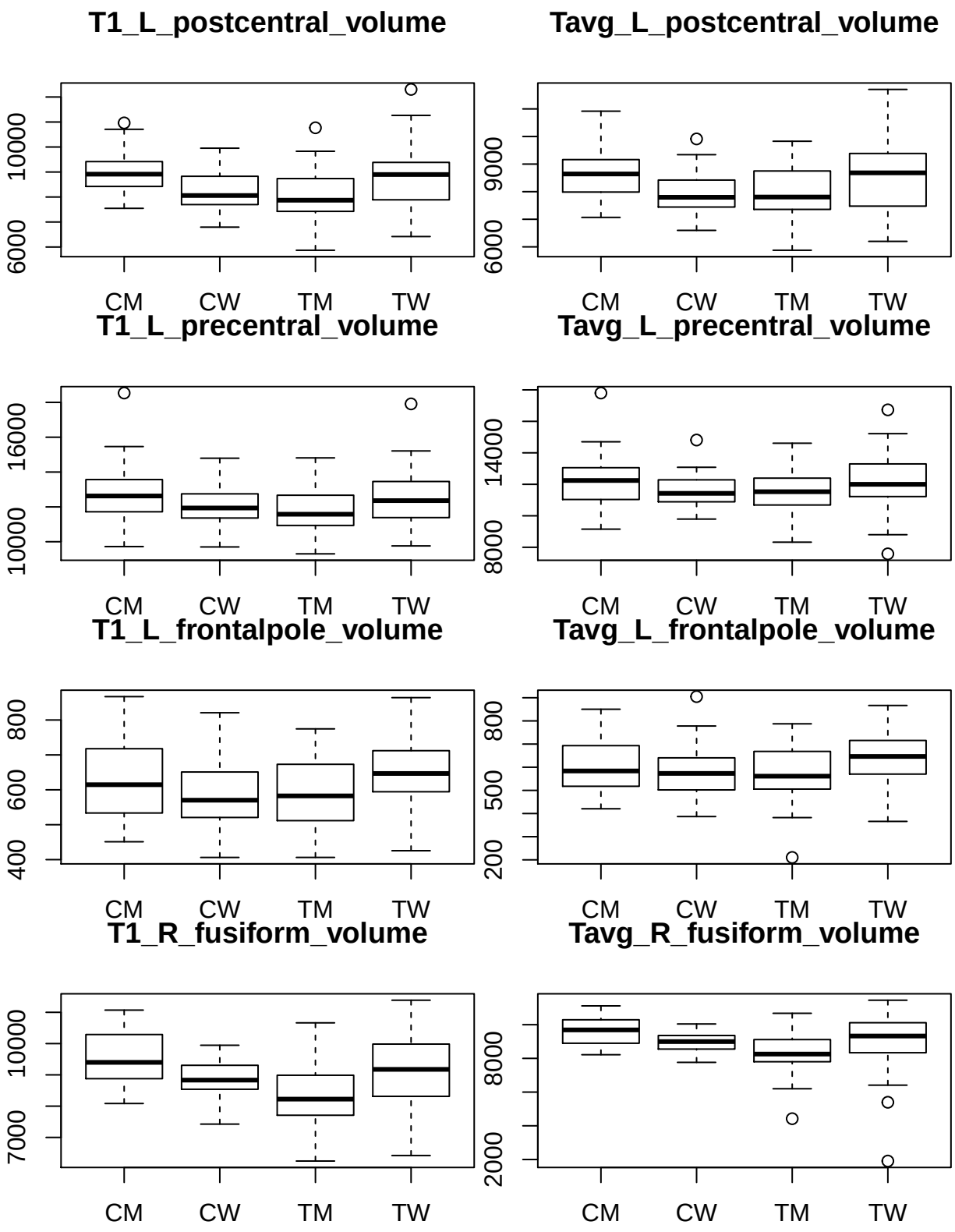
```
kable(cbind(names(data.hyp[,bg.avg:nd.avg]), round(pancorr.avg, 3)))
```

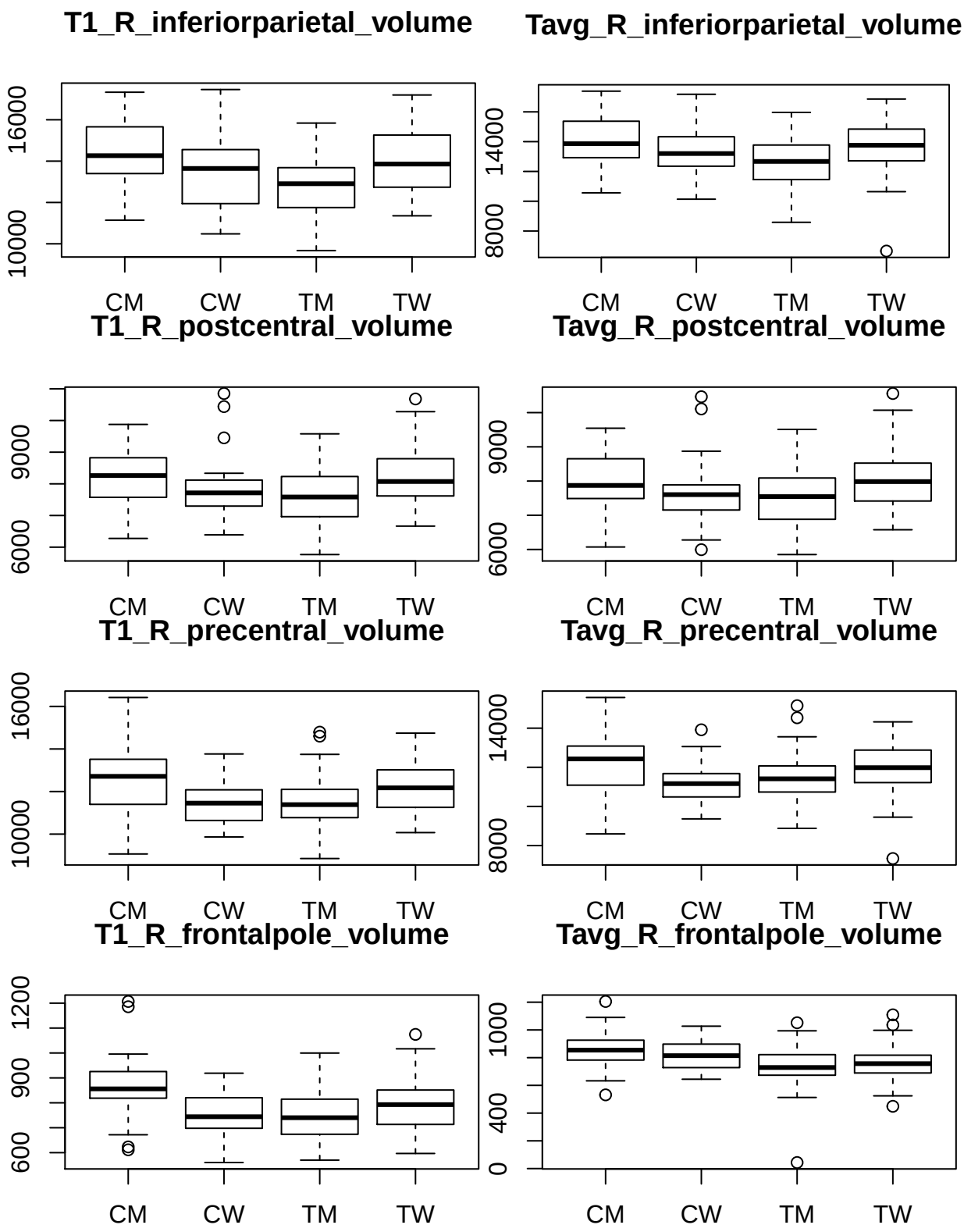
Tavg_L_fusiform_volume	0
Tavg_L_inferiorparietal_volume	0.005
Tavg_L_postcentral_volume	0.005
Tavg_L_precentral_volume	0.128
Tavg_L_frontalpole_volume	0.101
Tavg_R_fusiform_volume	0.001
Tavg_R_inferiorparietal_volume	0.004
Tavg_R_postcentral_volume	0.049
Tavg_R_precentral_volume	0.01
Tavg_R_frontalpole_volume	0.003
Tavg_LeftCerebellumWhiteMatter	0.004
Tavg_LeftCerebellumCortex	0
Tavg_RightCerebellumWhiteMatter	0.003
Tavg_RightCerebellumCortex	0
Tavg_LeftThalamusProper	0
Tavg_LeftCaudate	0.005
Tavg_LeftPutamen	0.001
Tavg_LeftAccumbensarea	0.387
Tavg_RightThalamusProper	0
Tavg_RightCaudate	0.008
Tavg_RightPutamen	0.002
Tavg_RightAccumbensarea	0.014

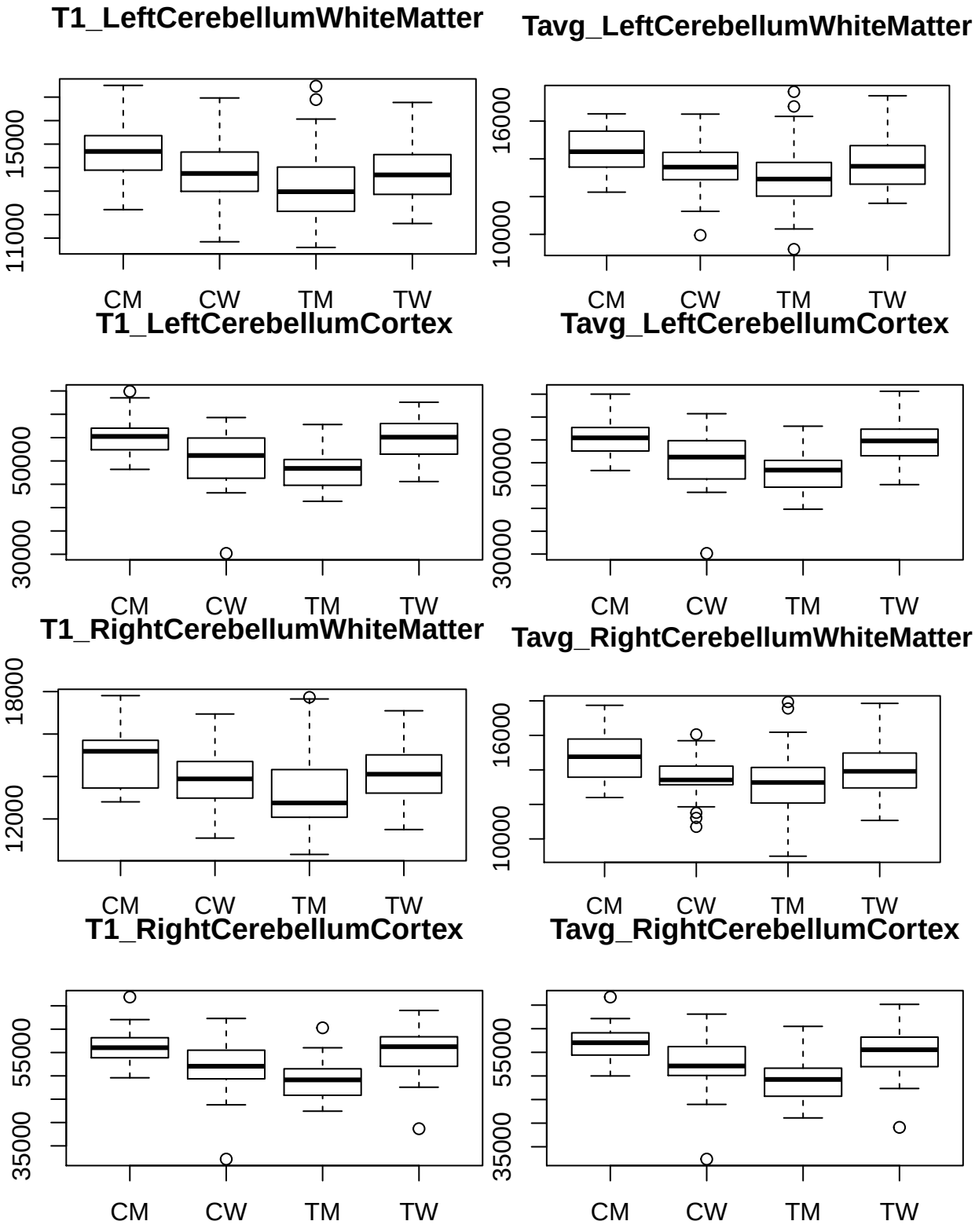
Why are there less regions for which the difference is statistically significant when the average is used compared to when one measure is used? To investigate this we look at the difference between the boxplots for one statistically significant region.

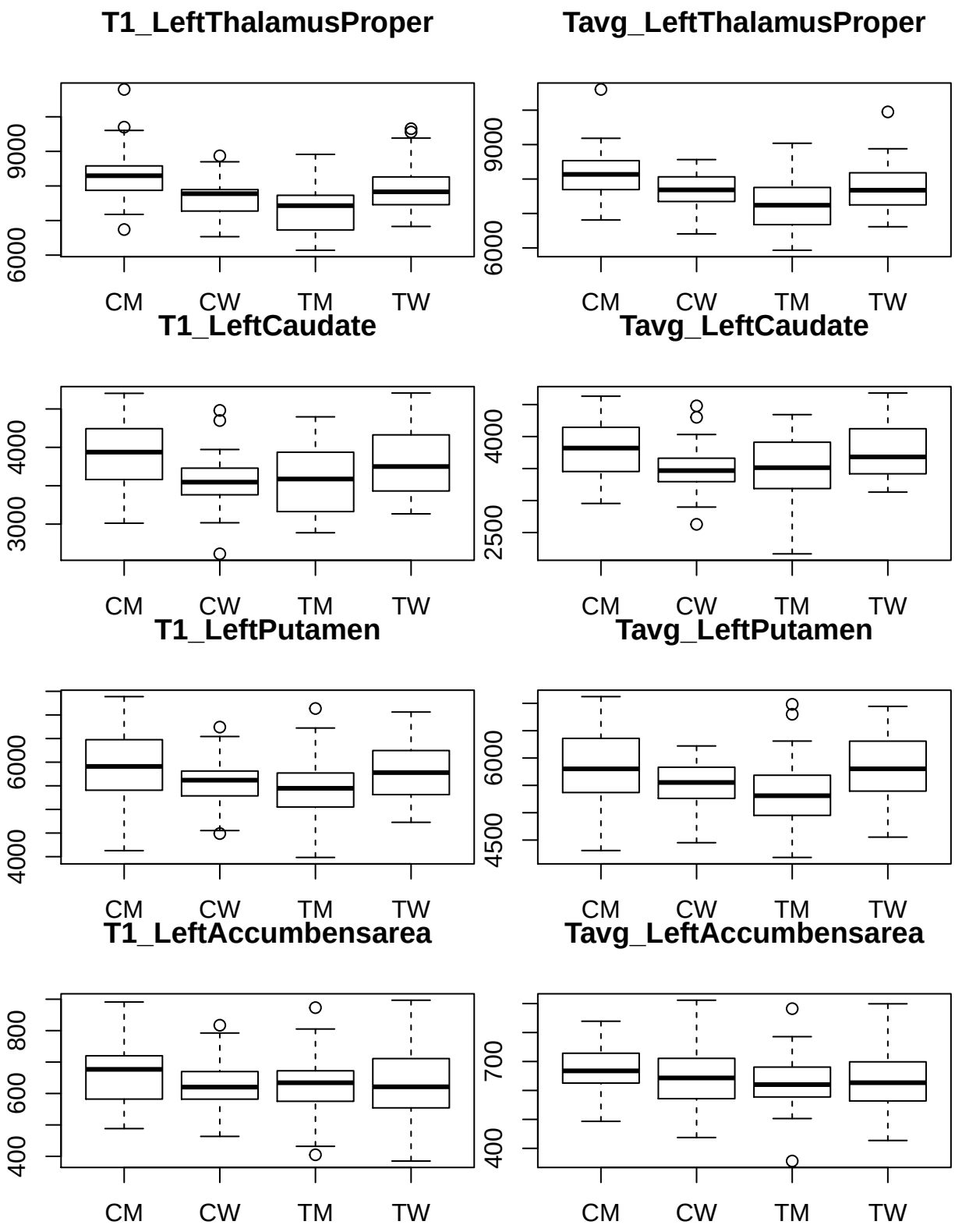
```r
boxplot(data.hyp[,22] ~ fac, main = names(data.hyp)[26])
boxplot(data.hyp[,66] ~ fac, main = names(data.hyp)[78])

for(i in bg.one:nd.one){
  # construct boxplot for every region
  boxplot(data.hyp[,i] ~ fac, main = names(data.hyp)[i])
  boxplot(data.hyp[,i + 44] ~ fac, main = names(data.hyp)[i + 44])
}
```

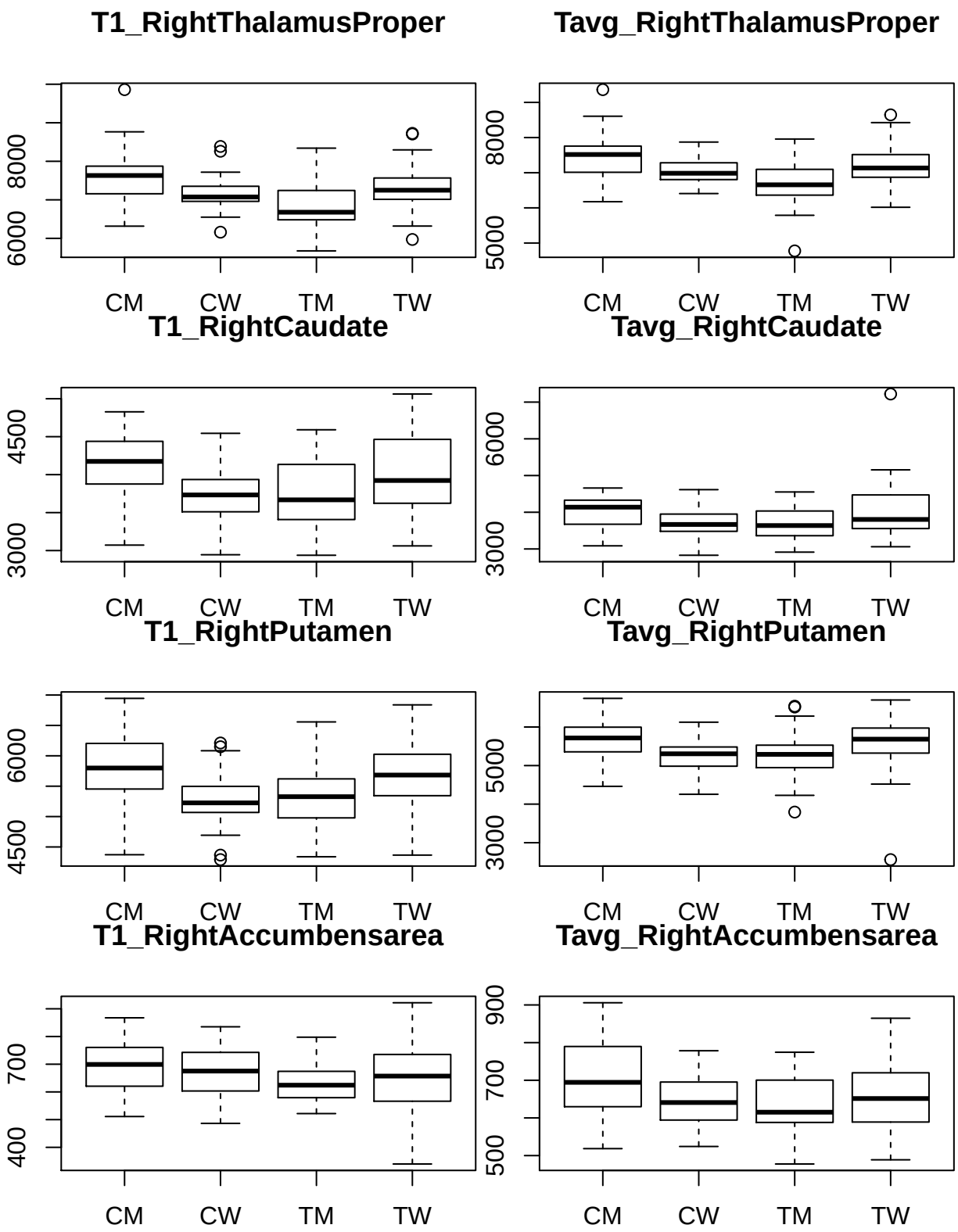



T1_L_postcentral_volume
Tavg_L_postcentral_volume
T1_L_precentral_volume
Tavg_L_precentral_volume
T1_L_frontalpole_volume
Tavg_L_frontalpole_volume
T1_R_fusiform_volume
Tavg_R_fusiform_volume
CM
CW
TM
TW


T1_R_inferiorparietal_volume
Tavg_R_inferiorparietal_volume
T1_R_postcentral_volume
Tavg_R_postcentral_volume
T1_R_precentral_volume
Tavg_R_precentral_volume
T1_R_frontalpole_volume
Tavg_R_frontalpole_volume
CM
CW
TM
TW


T1_LeftCerebellumWhiteMatter
11000
15000
CM
CW
TM
TW
Tavg_LeftCerebellumWhiteMatter
10000
16000
CM
CW
TM
TW
T1_LeftCerebellumCortex
30000
50000
CM
CW
TM
TW
Tavg_LeftCerebellumCortex
30000
50000
CM
CW
TM
TW
T1_RightCerebellumWhiteMatter
12000
18000
CM
CW
TM
TW
Tavg_RightCerebellumWhiteMatter
10000
16000
CM
CW
TM
TW
T1_RightCerebellumCortex
35000
55000
CM
CW
TM
TW
Tavg_RightCerebellumCortex
35000
55000
CM
CW
TM
TW


T1_LeftThalamusProper
6000
9000
CM
CW
TM
TW
Tavg_LeftThalamusProper
6000
9000
CM
CW
TM
TW
T1_LeftCaudate
3000
4000
CM
CW
TM
TW
Tavg_LeftCaudate
2500
4000
CM
CW
TM
TW
T1_LeftPutamen
4000
6000
CM
CW
TM
TW
Tavg_LeftPutamen
4500
6000
CM
CW
TM
TW
T1_LeftAccumbensarea
400
600
800
CM
CW
TM
TW
Tavg_LeftAccumbensarea
400
700
CM
CW
TM
TW


T1_RightThalamusProper
6000
8000
CM
CW
TM
TW
Tavg_RightThalamusProper
5000
8000
CM
CW
TM
TW
T1_RightCaudate
3000
4500
CM
CW
TM
TW
Tavg_RightCaudate
3000
6000
CM
CW
TM
TW
T1_RightPutamen
4500
6000
CM
CW
TM
TW
Tavg_RightPutamen
3000
5000
CM
CW
TM
TW
T1_RightAccumbensarea
400
700
CM
CW
TM
TW
Tavg_RightAccumbensarea
500
700
900
CM
CW
TM
TW

Post-hoc tests

We conduct post-hoc t-tests on the statistically significant regions to determine which group differences cause the effect. The p-values are uncorrected at this point and I computed them for every region.

```
# create an object with all possible combinations
  allcomb <- combn(c(1:4),2)
  allcomb.txt <- array(data=NA, dim = dim(allcomb)[2])
  labels.g <- c("CW", "CM", "TM", "TW")
  for(i in 1:dim(allcomb)[2]){
    allcomb.txt[i] <- paste(labels.g[allcomb[1,i]], " vs ", labels.g[allcomb[2,i]], sep="")
  }


# create object to save results
  pt.one <- array(data = NA, dim = c(length(pan.one),dim(allcomb)[2]))
  pt.avg <- array(data = NA, dim = c(length(pan.avg),dim(allcomb)[2]))

# Left Cerebellum White Matter
  # 1 measure
  bg.one <- 3
  nd.one <- 24
  pt.corr.one <- array(data=NA, dim=dim(pt.one))
  for(r in bg.one:nd.one){
    if (pan.one[r-bg.one+1] > 0.05) {
        pt.one[r-bg.one+1,] <- rep(NA,dim(allcomb)[2])
    }else{
      for(i in 1:dim(allcomb)[2]){
        pt.one[r-bg.one+1,i] <- unlist(t.test(data.hyp[data.hyp[,2]==allcomb[1,i],r], data.hyp[data.hyp
        pt.corr.one[r - bg.one + 1,] <- p.adjust(pt.one[r - bg.one + 1,], method = "bonferroni")
      }
    }
  }

  kable(cbind(c(" ",names(data.hyp[,bg.one:nd.one])),rbind(c("CW vs CM","CW vs TM","CW vs TW","CM vs TM
```

	CW vs CM	CW vs TM	CW vs TW	CM vs TM	CM vs TW	TM vs TW
T1_L_fusiform_volume	0.014	0.047	1	0	0.202	0.001
T1_L_inferiorparietal_volume	0.137	0.158	1	0	0.949	0.025
T1_L_postcentral_volume	0.01	1	0.233	0.002	1	0.06
T1_L_precentral_volume	0.284	1	0.807	0.087	1	0.243
T1_L_frontalpole_volume	0.929	1	0.015	0.998	1	0.009
T1_R_fusiform_volume	0.013	0.02	1	0	0.78	0.006
T1_R_inferiorparietal_volume	0.162	0.568	1	0.001	1	0.006
T1_R_postcentral_volume	0.87	1	0.534	0.067	1	0.018
T1_R_precentral_volume	0.014	1	0.026	0.034	1	0.087
T1_R_frontalpole_volume	0.004	1	0.624	0.005	0.177	0.724
T1_LeftCerebellumWhiteMatter	0.036	0.775	1	0	0.01	0.677
T1_LeftCerebellumCortex	0.005	0.587	0.009	0	1	0
T1_RightCerebellumWhiteMatter	0.04	1	1	0.001	0.133	0.221
T1_RightCerebellumCortex	0.003	0.227	0.058	0	1	0
T1_LeftThalamusProper	0.004	0.196	0.61	0	0.174	0.002
T1_LeftCaudate	0.021	1	0.072	0.02	1	0.066
T1_LeftPutamen	0.14	1	0.417	0.032	1	0.069

T1_LeftAccumbensarea	NA	NA	NA	NA	NA	NA
T1_RightThalamusProper	0.127	0.047	1	0	0.775	0.004
T1_RightCaudate	0.003	1	0.124	0.007	1	0.203
T1_RightPutamen	0	1	0.002	0.001	1	0.03
T1_RightAccumbensarea	NA	NA	NA	NA	NA	NA

```
# Average
bg.avg <- 47
nd.avg <- 68
pt.corr.avg <- array(data=NA, dim=dim(pt.avg))
for(r in bg.avg:nd.avg){
  if (pan.avg[r-bg.avg+1] > 0.05) {
      pt.avg[r-bg.avg+1,] <- rep(NA,dim(allcomb)[2])
  }else{
    for(i in 1:dim(allcomb)[2]){
      pt.avg[r-bg.avg+1,i] <- unlist(t.test(data.hyp[data.hyp[,2]==allcomb[1,i],r], data.hyp[data.hyp
      pt.corr.avg[r - bg.avg + 1,] <- p.adjust(pt.avg[r - bg.avg + 1,], method = "bonferroni")
    }
  }
}

kable(cbind(c(" ",names(data.hyp[,bg.avg:nd.avg])),rbind(c("CW vs CM","CW vs TM","CW vs TW","CM vs TM
```

	CW vs CM	CW vs TM	CW vs TW	CM vs TM	CM vs TW	TM vs TW
Tavg_L_fusiform_volume	0.015	0.137	1	0	0.15	0.102
Tavg_L_inferiorparietal_volume	0.147	0.39	1	0	1	0.074
Tavg_L_postcentral_volume	0.008	1	0.146	0.011	1	0.187
Tavg_L_precentral_volume	NA	NA	NA	NA	NA	NA
Tavg_L_frontalpole_volume	NA	NA	NA	NA	NA	NA
Tavg_R_fusiform_volume	0.005	0.022	1	0	0.389	0.225
Tavg_R_inferiorparietal_volume	0.353	0.597	1	0.004	1	0.02
Tavg_R_postcentral_volume	1	1	0.213	0.729	1	0.071
Tavg_R_precentral_volume	0.023	1	0.012	0.411	1	0.452
Tavg_R_frontalpole_volume	1	0.062	0.236	0.008	0.028	1
Tavg_LeftCerebellumWhiteMatter	0.088	0.916	1	0.001	0.35	0.176
Tavg_LeftCerebellumCortex	0.004	0.282	0.03	0	1	0
Tavg_RightCerebellumWhiteMatter	0.009	1	1	0.002	0.332	0.283
Tavg_RightCerebellumCortex	0.002	0.134	0.098	0	0.97	0
Tavg_LeftThalamusProper	0.037	0.091	1	0	0.144	0.024
Tavg_LeftCaudate	0.048	1	0.021	0.103	1	0.052
Tavg_LeftPutamen	0.043	1	0.045	0.009	1	0.007
Tavg_LeftAccumbensarea	NA	NA	NA	NA	NA	NA
Tavg_RightThalamusProper	0.024	0.073	0.494	0	0.909	0.003
Tavg_RightCaudate	0.009	1	0.087	0.024	1	0.168
Tavg_RightPutamen	0.004	1	0.027	0.015	1	0.081
Tavg_RightAccumbensarea	0.18	1	1	0.02	0.395	1

Correlations

between T1 and T2

```
bg <- 36
nd <- 108
ln <- nd-bg+1

# Object to save correlations
corrall <- array(data=NA, dim = ln)

# Compute ANOVA for every predictor and save p-value
for(i in bg:nd){
  corrall[i-bg+1] <- cor(x = data.all[,i], y = data.all[,i + ln + 1])
}

summary(corrall)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.6334  0.7950  0.8243  0.8216  0.8466  0.9985
```

```
bg.one <- 3
nd.one <- 24
ln <- nd.one-bg.one+1

# Object to save correlations
corrhyp <- array(data=NA, dim = ln)

# Compute ANOVA for every predictor and save p-value
for(i in bg.one:nd.one){
  corrhyp[i-bg.one+1] <- cor(x = data.hyp[,i], y = data.hyp[,i + ln])
}

summary(corrhyp)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.7046  0.8897  0.9313  0.8990  0.9492  0.9963
```

```
# plot with correlations? Which region lies where?
```

Discussion