Hippocampal subfield strophy in AD progression

Set up

```r
packages <- c("here","readxl","dplyr","ggplot2","psych","lmerTest",
          "lme4","emmeans")

lapply(packages, library, character.only = TRUE)
```

```
## [[1]]
## [1] "here"      "stats"     "graphics"  "grDevices" "utils"     "datasets"
## [7] "methods"   "base"
##
## [[2]]
## [1] "readxl"    "here"      "stats"     "graphics"  "grDevices" "utils"
## [7] "datasets"  "methods"   "base"
##
## [[3]]
##  [1] "dplyr"     "readxl"    "here"      "stats"     "graphics"  "grDevices"
##  [7] "utils"     "datasets"  "methods"   "base"
##
## [[4]]
##  [1] "ggplot2"   "dplyr"     "readxl"    "here"      "stats"     "graphics"
##  [7] "grDevices" "utils"     "datasets"  "methods"   "base"
##
## [[5]]
##  [1] "psych"     "ggplot2"   "dplyr"     "readxl"    "here"      "stats"
##  [7] "graphics"  "grDevices" "utils"     "datasets"  "methods"   "base"
##
## [[6]]
##  [1] "lmerTest"  "lme4"      "Matrix"    "psych"     "ggplot2"   "dplyr"
##  [7] "readxl"    "here"      "stats"     "graphics"  "grDevices" "utils"
## [13] "datasets"  "methods"   "base"
##
## [[7]]
##  [1] "lmerTest"  "lme4"      "Matrix"    "psych"     "ggplot2"   "dplyr"
##  [7] "readxl"    "here"      "stats"     "graphics"  "grDevices" "utils"
## [13] "datasets"  "methods"   "base"
##
## [[8]]
##  [1] "emmeans"   "lmerTest"  "lme4"      "Matrix"    "psych"     "ggplot2"
##  [7] "dplyr"     "readxl"    "here"      "stats"     "graphics"  "grDevices"
## [13] "utils"     "datasets"  "methods"   "base"
```

Load data

```r
# load all sample info
all_sample_info <- readxl::read_xlsx("datafile/Total.sample.final.xlsx", sheet = 1)

# load hippocampus subfields data
segHA_lh_AD <- read.table("datafile/hipposubfields_lh_AD.txt", header = T)
segHA_rh_AD <- read.table("datafile/hipposubfields_rh_AD.txt", header = T)

segHA_lh_CN <- read.table("datafile/hipposubfields.lh.CN.txt", header = T)
segHA_rh_CN <- read.table("datafile/hipposubfields.rh.CN.txt", header = T)

segHA_lh_MCI <- read.table("datafile/hipposubfields.lh.MCI.txt", header = T)
segHA_rh_MCI <- read.table("datafile/hipposubfields.rh.MCI.txt", header = T)

#View(segHA_lh_AD)

# function
segHA_sub <- function(dat, gr,LorR) {

    segHA_tmp <- dat
    segHA_tmp_new <- as.data.frame(matrix(0, dim(segHA_tmp)[1], 1))
    segHA_tmp_new$subjid <- segHA_tmp$Measure.volume

    segHA_tmp_new$CA1 <- segHA_tmp$CA1.head + segHA_tmp$CA1.body
    segHA_tmp_new$CA3 <- segHA_tmp$CA3.head + segHA_tmp$CA3.body
    segHA_tmp_new$CA4 <- segHA_tmp$CA4.head + segHA_tmp$CA4.body
    segHA_tmp_new$subiculum <- segHA_tmp$subiculum.head + segHA_tmp$subiculum.body
    segHA_tmp_new$presubiculum <- segHA_tmp$presubiculum.head + segHA_tmp$presubiculum.body
    segHA_tmp_new$GC.ML.DG <- segHA_tmp$GC.ML.DG.head + segHA_tmp$GC.ML.DG.head
    segHA_tmp_new$molecular_layer_HP <- segHA_tmp$molecular_layer_HP.head +
    segHA_tmp$molecular_layer_HP.body
    segHA_tmp_new$parasubiculum <- segHA_tmp$parasubiculum
    segHA_tmp_new$Hippocampal_tail <- segHA_tmp$Hippocampal_tail
    segHA_tmp_new$HATA <- segHA_tmp$HATA
    segHA_tmp_new$fimbria <- segHA_tmp$fimbria
    segHA_tmp_new$hippocampal.fissure <- segHA_tmp$hippocampal.fissure

    segHA_tmp_new$Whole_hippocampus <- segHA_tmp$Whole_hippocampus

    #View(segHA_tmp_new)
    colnames(segHA_tmp_new)
    segHA_tmp_new <- segHA_tmp_new[,-1]

    # merge
    all_data_subfileds <- merge(all_sample_info, segHA_tmp_new,
                        by.x = "Subject.ID", by.y = "subjid")

    # add estimated TIV
    eTIV <- read.table(paste0("datafile/eTIV_",gr,".txt"), header = T)

    all_data_subfileds <- merge(all_data_subfileds, eTIV, by.x = "Subject.ID",
                        by.y = "Subj")
```

```r
    all_data_subfileds$eTIV_ml <- round(all_data_subfileds$eTIV/1000,2)

    write.csv(all_data_subfileds, paste0("datafile/all_data_subfileds.",
                              gr,".",LorR,".csv"), row.names = F)
}

# run function
segHA_sub(segHA_lh_AD,"AD","lh")
segHA_sub(segHA_rh_AD,"AD","rh")

segHA_sub(segHA_lh_CN,"CN","lh")
segHA_sub(segHA_rh_CN,"CN","rh")

segHA_sub(segHA_lh_MCI,"MCI","lh")
segHA_sub(segHA_rh_MCI,"MCI","rh")

# organize data
gr <- "AD"
LorR <- "lh"
all_data_subfileds_AD.lh <- read.csv(paste0("datafile/all_data_subfileds.",
                          gr,".",LorR,".csv"))
gr <- "AD"
LorR <- "rh"
all_data_subfileds_AD.rh <- read.csv(paste0("datafile/all_data_subfileds.",
                          gr,".",LorR,".csv"))
gr <- "CN"
LorR <- "rh"
all_data_subfileds_CN.rh <- read.csv(paste0("datafile/all_data_subfileds.",
                          gr,".",LorR,".csv"))
gr <- "CN"
LorR <- "lh"
all_data_subfileds_CN.lh <- read.csv(paste0("datafile/all_data_subfileds.",
                          gr,".",LorR,".csv"))

gr <- "MCI"
LorR <- "rh"
all_data_subfileds_MCI.rh <- read.csv(paste0("datafile/all_data_subfileds.",
                          gr,".",LorR,".csv"))
gr <- "MCI"
LorR <- "lh"
all_data_subfileds_MCI.lh <- read.csv(paste0("datafile/all_data_subfileds.",
                          gr,".",LorR,".csv"))

#View(all_data_subfileds_AD.lh)

all_data_subfileds <- rbind.data.frame(cbind(all_data_subfileds_AD.lh, LorR = "Left"),
                          cbind(all_data_subfileds_CN.lh, LorR = "Left"),
                          cbind(all_data_subfileds_MCI.lh, LorR = "Left"),
                          cbind(all_data_subfileds_AD.rh, LorR = "Right"),
                          cbind(all_data_subfileds_CN.rh, LorR = "Right"),
                          cbind(all_data_subfileds_MCI.rh, LorR = "Right"))

dim(all_data_subfileds)
```

```
## [1] 758  30
```

```r
all_data_subfileds$Research.Group <- factor(all_data_subfileds$Research.Group,
                          levels=c("CN","MCI","AD"))

all_data_subfileds$LorR <- factor(all_data_subfileds$LorR, levels = c("Left", "Right"))

dim(all_data_subfileds)
```

```
## [1] 758  30
```

```r
writexl::write_xlsx(all_data_subfileds, "datafile/all_data_subfileds.xlsx", col_names = T)
```

CN converted & non-converted subjects

```r
# all sample
all_data_subfileds_rm <- readxl::read_xlsx("datafile/all_data_subfileds.xlsx", sheet = 1)
dim(all_data_subfileds_rm)
```

```
## [1] 758  30
```

```r
table(all_data_subfileds_rm$Research.Group)
```

```
##
##  AD  CN MCI
## 290 288 180
```

```r
# CN converted
converted_subj <- read.table("datafile/subj_CN_Converted.txt", header = T)

dim(converted_subj)
```

```
## [1] 28  2
```

```r
all_data_subfileds_rm$group <- all_data_subfileds_rm$Research.Group
all_data_subfileds_rm$group <- as.character(all_data_subfileds_rm$group)

all_data_subfileds_rm$group[all_data_subfileds_rm$Research.Group == "CN" &
                all_data_subfileds_rm$Subject.ID %in%
                converted_subj$subjid] <- "Converted_CN"

all_data_subfileds_rm$group[all_data_subfileds_rm$Research.Group == "CN" &
                !all_data_subfileds_rm$Subject.ID %in%
                converted_subj$subjid] <- "NonConverted_CN"

table(all_data_subfileds_rm$group)
```

```
##
##                AD    Converted_CN           MCI NonConverted_CN
##               290              56           180             232
```

```
# period of conversion
summary(converted_subj$period_of_conversion)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    6.00   12.00   24.00   36.18   51.00  105.00
```

```
sd(converted_subj$period_of_conversion)
```

```
## [1] 31.91662
```

```
# save file
write.csv(all_data_subfileds_rm,"datafile/all_data_subfileds_convertedCN.csv",row.names = F)
```

```
table(all_data_subfileds_rm$group)
```

```
##
##                AD    Converted_CN           MCI NonConverted_CN
##               290              56           180             232
```

Demographic and clinical information

```
# summary of demographic and clinical info
all_data_subfileds_L <- all_data_subfileds_rm[all_data_subfileds_rm$LorR == "Left",]
```

```
# CN, MCI, and AD groups
mm <- table(all_data_subfileds_L$Research.Group, all_data_subfileds_L$Gender)
mm
```

```
##
##         1  2
##   AD   61 84
##   CN   78 66
##   MCI  42 48
```

```
mean(all_data_subfileds_L$Age[all_data_subfileds_L$Research.Group == "CN"])
```

```
## [1] 74.77847
```

```
sd(all_data_subfileds_L$Age[all_data_subfileds_L$Research.Group == "CN"])
```

```
## [1] 7.787763
```

```r
mean(all_data_subfileds_L$Age[all_data_subfileds_L$Research.Group == "MCI"])
```

```
## [1] 72.55889
```

```r
sd(all_data_subfileds_L$Age[all_data_subfileds_L$Research.Group == "MCI"])
```

```
## [1] 9.163299
```

```r
mean(all_data_subfileds_L$Age[all_data_subfileds_L$Research.Group == "AD"])
```

```
## [1] 74.94345
```

```r
sd(all_data_subfileds_L$Age[all_data_subfileds_L$Research.Group == "AD"])
```

```
## [1] 7.504441
```

```r
# one way ANOVA
chisq.test(mm)
```

```
## 
##  Pearson's Chi-squared test
## 
## data:  mm
## X-squared = 4.2944, df = 2, p-value = 0.1168
```

```r
ano_age <- anova(lm(Age ~ Research.Group, all_data_subfileds_L))


# CN-NC, CN-C, MCI, and AD
demo_info <- as.data.frame(matrix(0,7,9))

colnames(demo_info) <- c("CN_NC_mean", "CN_NC_range", "CN_C_mean", "CN_C_range",
            "MCI_mean", "MCI_range","AD_mean", "AD_range","p value")

rownames(demo_info) <- c("Gender","Age","Education","CDR","MMSE","MoCA","MIS")

# gender: chi-squared test
gender_gr <- table(all_data_subfileds_L$group, all_data_subfileds_L$Gender)

chisq.test(gender_gr)$p.value
```

```
## [1] 0.2309044
```

```r
demo_info["Gender", 1]  <- paste0(gender_gr["NonConverted_CN",2] ,"/",gender_gr["NonConverted_CN",1])
demo_info["Gender", 3]  <- paste0(gender_gr["Converted_CN",2] ,"/",gender_gr["Converted_CN",1])
demo_info["Gender", 5]  <- paste0(gender_gr["MCI",2] ,"/",gender_gr["MCI",1])
demo_info["Gender", 7]  <- paste0(gender_gr["AD",2] ,"/",gender_gr["AD",1])
demo_info["Gender", 9]  <- round(chisq.test(gender_gr)$p.value,3)
```

```r
# Age, Education, CDR, MMSE, MoCA: mean, sd, range
all_data_subfileds_L$group <- factor(all_data_subfileds_L$group,
                    levels = c("NonConverted_CN","Converted_CN","MCI","AD"))

mean_all <- describeBy(all_data_subfileds_L[, c("group","Age","Education",
                        "Global.CDR","MMSE.Total.Score",
                        "MoCA.Total","MoCA.MIS")],
                        group = "group",
                        digits = 3)

demo_info[c("Age","Education","CDR","MMSE","MoCA","MIS"), 1:8] <-
    cbind(paste0(round(mean_all[[1]]$mean[2:7],2),"±",
            round(mean_all[[1]]$sd[2:7],2)),
        paste0(round(mean_all[[1]]$min[2:7],2),"-",
            round(mean_all[[1]]$max[2:7],2)),
        paste0(round(mean_all[[2]]$mean[2:7],2),"±",
            round(mean_all[[2]]$sd[2:7],2)),
        paste0(round(mean_all[[2]]$min[2:7],2),"-",
            round(mean_all[[2]]$max[2:7],2)),
        paste0(round(mean_all[[3]]$mean[2:7],2),"±",
            round(mean_all[[3]]$sd[2:7],2)),
        paste0(round(mean_all[[3]]$min[2:7],2),"-",
            round(mean_all[[3]]$max[2:7],2)),
         paste0(round(mean_all[[4]]$mean[2:7],2),"±",
            round(mean_all[[4]]$sd[2:7],2)),
        paste0(round(mean_all[[4]]$min[2:7],2),"-",
            round(mean_all[[4]]$max[2:7],2)))

# one way ANOVA
ano_age <- anova(lm(Age ~ group, all_data_subfileds_L))
ano_education <- anova(lm(Education ~ group, all_data_subfileds_L))
ano_MMSE <- anova(lm(MMSE.Total.Score ~ group, all_data_subfileds_L))
ano_MoCA <- anova(lm(MoCA.Total ~ group, all_data_subfileds_L))
ano_MIS <- anova(lm(MoCA.MIS ~ group, all_data_subfileds_L))

demo_info["Age", 9] <- round(ano_age$`Pr(>F)`[1],3)
demo_info["Education", 9] <- round(ano_education$`Pr(>F)`[1],3)
demo_info["MMSE", 9] <- round(ano_MMSE$`Pr(>F)`[1],4)
demo_info["MoCA", 9] <- round(ano_MoCA$`Pr(>F)`[1],4)
demo_info["MIS", 9] <- round(ano_MIS$`Pr(>F)`[1],4)

# Age
t.test(all_data_subfileds_L$Age[all_data_subfileds_L$Research.Group == "AD"],
        all_data_subfileds_L$Age[all_data_subfileds_L$Research.Group == "CN"])
```

```
##
##  Welch Two Sample t-test
##
## data:  all_data_subfileds_L$Age[all_data_subfileds_L$Research.Group == "AD"] and all_data_subfileds_L
## t = 0.18336, df = 286.45, p-value = 0.8546
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -1.606002  1.935954
```

```
## sample estimates:
## mean of x mean of y
##  74.94345  74.77847
```

```
t.test(all_data_subfileds_L$Age[all_data_subfileds_L$Research.Group == "CN"],
       all_data_subfileds_L$Age[all_data_subfileds_L$Research.Group == "MCI"])
```

```
##
##  Welch Two Sample t-test
##
## data:  all_data_subfileds_L$Age[all_data_subfileds_L$Research.Group == "CN"] and all_data_subfileds_L
## t = 1.9074, df = 166.39, p-value = 0.05819
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.07788104  4.51704771
## sample estimates:
## mean of x mean of y
##  74.77847  72.55889
```

```
t.test(all_data_subfileds_L$Age[all_data_subfileds_L$Research.Group == "AD"],
       all_data_subfileds_L$Age[all_data_subfileds_L$Research.Group == "MCI"])
```

```
##
##  Welch Two Sample t-test
##
## data:  all_data_subfileds_L$Age[all_data_subfileds_L$Research.Group == "AD"] and all_data_subfileds_L
## t = 2.0744, df = 161.25, p-value = 0.03963
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.1145476 4.6545712
## sample estimates:
## mean of x mean of y
##  74.94345  72.55889
```

```
# CDR
kruskal.test(Global.CDR ~ Research.Group, data = all_data_subfileds_L)
```

```
##
##  Kruskal-Wallis rank sum test
##
## data:  Global.CDR by Research.Group
## Kruskal-Wallis chi-squared = 325.76, df = 2, p-value < 2.2e-16
```

```
demo_info["CDR", 9] <- "< 0.001"
```

```
knitr::kable(demo_info[, -9])
```

	CN_NC_mean	CN_NC_range	CN_C_mean	CN_C_range	MCI_mean	MCI_range	AD_mean	AD_range
Gender	53/63	0	13/15	0	48/42	0	84/61	0
Age	73.98±7.82	58.4-91.4	78.1±6.8	66.7-93.2	72.56±9.16	55.2-97.4	74.94±7.5	56-91

	CN_NC_mean	CN_NC_range	CN_C_mean	CN_C_range	MCI_mean	MCI_range	AD_mean	AD_range
Education	17.02±2.31	12-20	16±2.09	12-20	16.08±2.67	8-20	15.52±2.66	8-20
CDR	0±0	0-0	0±0	0-0	0.52±0.09	0.5-1	0.81±0.35	0.5-2
MMSE	29.03±1.08	26-30	29±1.19	25-30	27.27±2.43	16-30	22.72±3.04	5-29
MoCA	24.36±1.77	18-28	24.18±1.89	20-27	22.63±3.47	10-29	18.14±4.75	9-27
MIS	7.55±2.57	0-13	7.46±3.25	0-12	7.68±3.58	0-14	6.88±3.72	0-15

```r
writexl::write_xlsx(demo_info,"results/demo_info.xlsx")

# subjects with education
length(which(!is.na(all_data_subfileds_L$Education[all_data_subfileds_L$Research.Group == "CN"])))
```

```
## [1] 138
```

```r
length(which(!is.na(all_data_subfileds_L$Education[all_data_subfileds_L$Research.Group == "MCI"])))
```

```
## [1] 89
```

```r
length(which(!is.na(all_data_subfileds_L$Education[all_data_subfileds_L$Research.Group == "AD"])))
```

```
## [1] 138
```

```r
# subjects with MMSE
length(which(!is.na(all_data_subfileds_L$MMSE.Total.Score[all_data_subfileds_L$Research.Group == "CN"]))
```

```
## [1] 144
```

```r
length(which(!is.na(all_data_subfileds_L$MMSE.Total.Score[all_data_subfileds_L$Research.Group == "MCI"])
```

```
## [1] 90
```

```r
length(which(!is.na(all_data_subfileds_L$MMSE.Total.Score[all_data_subfileds_L$Research.Group == "AD"]))
```

```
## [1] 145
```

```r
# subjects with MoCA
length(which(!is.na(all_data_subfileds_L$MoCA.Total[all_data_subfileds_L$Research.Group == "CN"])))
```

```
## [1] 134
```

```r
length(which(!is.na(all_data_subfileds_L$MoCA.Total[all_data_subfileds_L$Research.Group == "MCI"])))
```

```
## [1] 79
```

```
length(which(!is.na(all_data_subfileds_L$MoCA.Total[all_data_subfileds_L$Research.Group == "AD"]))))
```

```
## [1] 84
```

Hippocampus and hippocampal subfields in CN-NC, CN-C, MCI, and AD groups

```
subfiled_sum <- as.data.frame(matrix(0,length(c(15:27,29)),8))

for (LorR in c( "Left",  "Right")) {
    tmp <- all_data_subfileds_rm[all_data_subfileds_rm$LorR == LorR,]
    mean_all <- describeBy(tmp[, c(15:27,29,31)], group = "group",digits = 3)

    subfiled_sum[1:length(c(15:27,29)), 1:8] <-
        cbind(paste0(round(mean_all[[1]]$mean[1:14],2),"±",
                    round(mean_all[[1]]$sd[1:14],2)),
            paste0(round(mean_all[[1]]$min[1:14],2),"-",
                    round(mean_all[[1]]$max[1:14],2)),
            paste0(round(mean_all[[2]]$mean[1:14],2),"±",
                    round(mean_all[[2]]$sd[1:14],2)),
            paste0(round(mean_all[[2]]$min[1:14],2),"-",
                    round(mean_all[[2]]$max[1:14],2)),
            paste0(round(mean_all[[3]]$mean[1:14],2),"±",
                    round(mean_all[[3]]$sd[1:14],2)),
            paste0(round(mean_all[[3]]$min[1:14],2),"-",
                    round(mean_all[[3]]$max[1:14],2)),
            paste0(round(mean_all[[4]]$mean[1:14],2),"±",
                    round(mean_all[[4]]$sd[1:14],2)),
            paste0(round(mean_all[[4]]$min[1:14],2),"-",
                    round(mean_all[[4]]$max[1:14],2)))

#View(subfiled_sum)
row.names(subfiled_sum) <- colnames(tmp[, c(15:27,29)])

writexl::write_xlsx(subfiled_sum,paste0("results/subfiled_sum_",LorR,".xlsx"))

}
```

Violin plots for each group

```
# plot
all_data_subfileds_rm <- read.csv("datafile/all_data_subfileds_convertedCN.csv")

all_data_subfileds_rm$group <- factor(all_data_subfileds_rm$group,
                    levels= c("NonConverted_CN","Converted_CN",
                        "MCI","AD"))
```

```r
all_data_subfileds_rm$Research.Group <- factor(all_data_subfileds_rm$Research.Group,
                          levels = c("CN","MCI","AD"))

for (sub in colnames(all_data_subfileds_rm)[15:27]) {

    ggplot(all_data_subfileds_rm, aes_string(x = "LorR", y =sub, color = "group")) +
        geom_violin( position = position_dodge(width=0.9))+
        scale_color_brewer(palette = "Set1")+
        #geom_point(position = position_dodge(width=0.9)) +
        stat_summary(fun=mean, size = 3, position = position_dodge(width=0.9),
            geom = "point",shape=4)+   # shape=1
    guides(colour = "none") +
        theme(axis.text = element_text(size = 12, face = "bold"),
            axis.title = element_blank())+
        theme(panel.border = element_blank(),
            panel.background = element_blank(),
            panel.grid = element_blank(),
            axis.line = element_line(colour = "black"))

        ggsave(paste0("results/","boxplot_point.",sub,".convertedCN.png"),
            width = 4, height = 3, units = c("in"), dpi = 200)

}
```

Group differences in subfields

```r
all_data_subfileds_rm <- read.csv("datafile/all_data_subfileds_convertedCN.csv")
all_data_subfileds_rm$group <- factor(all_data_subfileds_rm$group, levels =
                      c("NonConverted_CN","Converted_CN",
                        "MCI","AD"))

table(all_data_subfileds_rm$group)
```

```
##
## NonConverted_CN     Converted_CN          MCI              AD
##             232               56          180             290
```

```r
gr_diff <- as.data.frame(matrix(0, length(15:27), 19))
all_data_subfileds_rm$LorR <- as.character(all_data_subfileds_rm$LorR)
all_data_subfileds_rm$Research.Group <-
    as.factor(all_data_subfileds_rm$Research.Group)

i <- 0
for (subfields in colnames(all_data_subfileds_rm)[15:27]) {
    for (LorR in c("Left", "Right")) {

    i <- i+1
    fl <- formula(paste0(subfields, " ~ group + Age + Gender +
                eTIV_ml + Education"))
```

```r
    diff <- anova(lm(fl, all_data_subfileds_rm[all_data_subfileds_rm$LorR == LorR,]))

    gr_diff[i,1:2] <- c(subfields,LorR)
    gr_diff[i, 3:7] <- round(diff$`Pr(>F)`[1:5],4)

    # post-hoc analysis
    tt <- pairwise.t.test(all_data_subfileds_rm[all_data_subfileds_rm$LorR == LorR,
                            subfields],
                  all_data_subfileds_rm[all_data_subfileds_rm$LorR == LorR,
                            "group"],
            p.adjust="fdr", pool.sd = T)

    gr_diff[i,8] <- "NonConverted_CN vs. Converted_CN"
    gr_diff[i,9] <- round(tt$p.value[1,1],4)

    gr_diff[i,10] <- "NonConverted_CN vs. MCI"
    gr_diff[i,11] <- round(tt$p.value[2,1],4)

    gr_diff[i,12] <- "NonConverted_CN vs. AD"
    gr_diff[i,13] <- round(tt$p.value[3,1],4)

    gr_diff[i,14] <- "Converted_CN vs. MCI"
    gr_diff[i,15] <- round(tt$p.value[2,2],4)

    gr_diff[i,16] <- "Converted_CN vs. AD"
    gr_diff[i,17] <- round(tt$p.value[3,2],4)

    gr_diff[i,18] <- "MCI vs. AD"
    gr_diff[i,19] <- round(tt$p.value[3,3],4)

    }
}

colnames(gr_diff) <- c("subfields","Hemisphere", "Group","Age","Gender","eTIV",
            "Education","contrast1","p-value1","contrast2","p-value2",
            "contrast3","p-value3","contrast4","p-value4","contrast5",
            "p-value5", "contrast6","p-value6")
#View(gr_diff)

writexl::write_xlsx(gr_diff, "results/group_diff_all.xlsx", col_names = T)
```

Correlations between subfields and MoCA MIS

```r
# regression analysis
cor_sum <- as.data.frame(matrix(0, 2*4*length(15:27), 0))

dim(all_data_subfileds_rm)
```

```
## [1] 758  31
```

```r
i <- 0
for (test in c( "MoCA.MIS")) {
for (gr in c("NonConverted_CN","Converted_CN", "MCI","AD")) {
    for (subfields in colnames(all_data_subfileds_rm)[15:27]) {
        for (LorR in c("Left", "Right")) {
        i <- i+1
        tmp <- all_data_subfileds_rm[all_data_subfileds_rm$group == gr
                        & all_data_subfileds_rm$LorR == LorR,]
            mycor <- cor.test(tmp[,subfields], tmp[, test])
    ff <- formula(paste0(test, " ~ ",subfields, "+ Age+ Gender + Education + eTIV_ml"))
    ll <- anova(lm(ff, tmp))

    cor_sum$test[i] <- test
    cor_sum$group[i] <- gr
    cor_sum$subfields[i] <- subfields
    cor_sum$LorR[i] <- LorR
    cor_sum$`r-value`[i] <- round(mycor$estimate,3)
    cor_sum$`p-value`[i] <- round(mycor$p.value,3)
    cor_sum$Pr[i] <- ll$`Pr(>F)`[1]
}
    }


}
}

writexl::write_xlsx(cor_sum, paste0("results/cor_sum_MoCA.xlsx"), col_names = T)

#View(cor_sum)

# check significant results
cor_sum_total <- readxl::read_xlsx(paste0("results/cor_sum_MoCA.xlsx"), sheet = 1)

#View(cor_sum_total)

#knitr::kable(cor_sum_total)

test <- "MoCA.MIS" # "MoCA.MIS"

which(p.adjust(cor_sum_total$Pr[cor_sum_total$test == test &
                cor_sum_total$group == "AD" & cor_sum_total$LorR == "Left"]) < 0.05)
```

```
## integer(0)
```

```r
which(p.adjust(cor_sum_total$Pr[cor_sum_total$test == test &
                cor_sum_total$group == "AD" & cor_sum_total$LorR == "Right"]) < 0.05)
```

```
## [1] 5
```

```r
which(p.adjust(cor_sum_total$Pr[cor_sum_total$test == test &
                cor_sum_total$group == "MCI" &
                cor_sum_total$LorR == "Left"]) < 0.05)
```

```
## integer(0)
```

13

```
which(p.adjust(cor_sum_total$Pr[cor_sum_total$test == test &
                      cor_sum_total$group == "MCI" &
                      cor_sum_total$LorR == "Right"]) < 0.05)
```

```
## integer(0)
```

```
which(p.adjust(cor_sum_total$Pr[cor_sum_total$test == test &
                      cor_sum_total$group == "NonConverted_CN" &
                      cor_sum_total$LorR == "Left"]) < 0.05)
```

```
## [1] 5
```

```
which(p.adjust(cor_sum_total$Pr[cor_sum_total$test == test &
                      cor_sum_total$group == "NonConverted_CN" &
                      cor_sum_total$LorR == "Right"]) < 0.05)
```

```
## [1] 5
```

Scatter plots for significant correlation results in CN-NC and AD

```
# scatter plots
tmp <- all_data_subfileds_rm[all_data_subfileds_rm$Research.Group == "AD" &
                      all_data_subfileds_rm$LorR == "Right",]

plm <- lm(formula = MoCA.MIS ~ presubiculum + Gender + Age +eTIV_ml + Education,
      data=tmp)

anova(plm)
```

```
## Analysis of Variance Table
##
## Response: MoCA.MIS
##               Df Sum Sq Mean Sq F value   Pr(>F)
## presubiculum   1 117.77 117.774  9.1063 0.003492 **
## Gender         1   0.22   0.224  0.0173 0.895608
## Age            1   1.30   1.301  0.1006 0.752026
## eTIV_ml        1   0.00   0.000  0.0000 0.997519
## Education      1   0.12   0.120  0.0093 0.923406
## Residuals     74 957.07  12.933
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
tmp$group <- droplevels(tmp$group)

cor.test(tmp$MoCA.MIS[tmp$group == "AD"],
      tmp$presubiculum[tmp$group == "AD"])
```
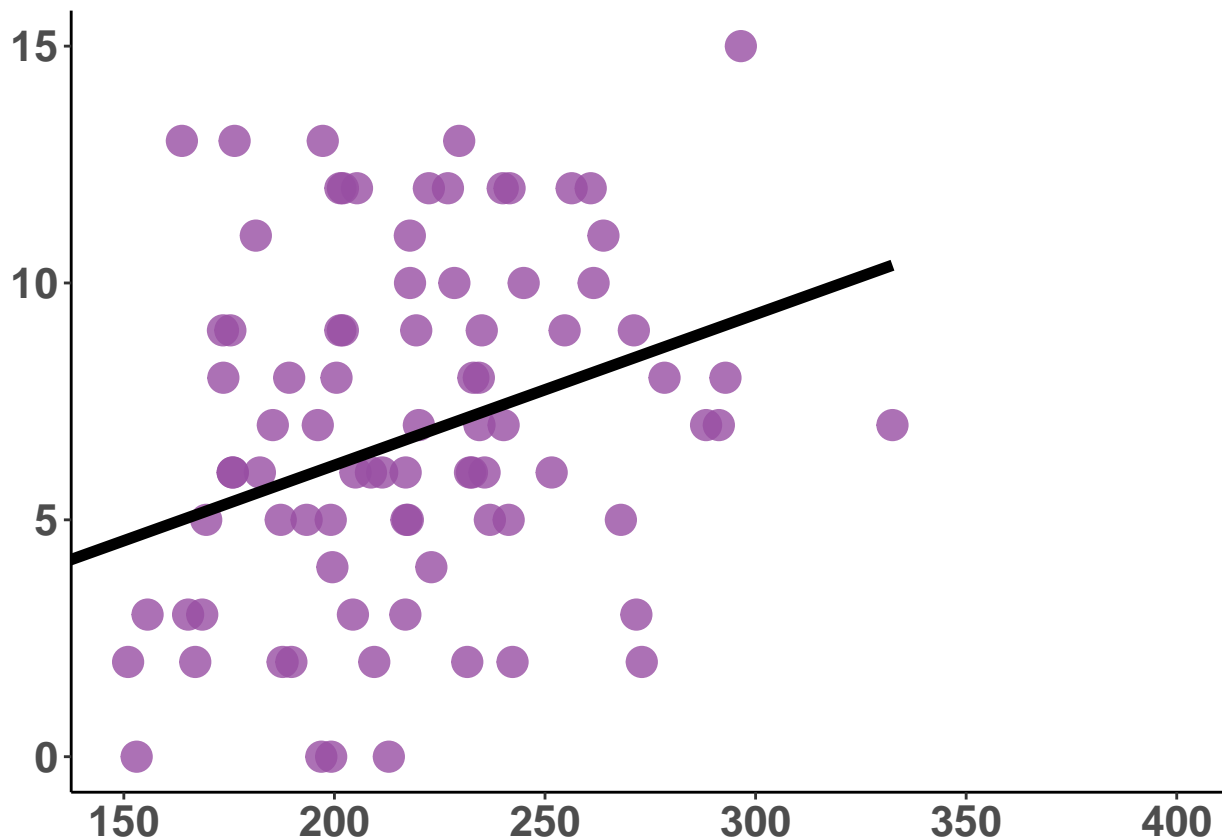
```
##
##  Pearson's product-moment correlation
##
## data:  tmp$MoCA.MIS[tmp$group == "AD"] and tmp$presubiculum[tmp$group == "AD"]
## t = 2.6595, df = 83, p-value = 0.009389
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.07136404 0.46554635
## sample estimates:
##       cor
## 0.2802261
```

```r
# plot with lm for outcome ~ pred + factor

p <- ggplot(data=tmp, aes(x=presubiculum, y=MoCA.MIS, color = group)) +
    geom_point(size = 5, col = '#984EA3', alpha = 0.8) + #col = 'blue',
    geom_smooth(method = "lm", mapping=aes(y=predict(plm,tmp)),
            col = 'black',linewidth = 2, se = FALSE) +
    guides(color = "none") +
    labs(x = "", y = "") +
    theme(legend.title = element_text(colour="black", size=14, face="bold"),
          legend.text = element_text(colour="black", size=14, face="bold")) +
    theme(plot.title = element_text(hjust = 0.5))+
    theme(axis.title = element_blank(),
          axis.text = element_text(size = 16, face = "bold")) +
    theme(panel.background = element_blank(),
          panel.border = element_blank(),
          panel.grid = element_blank(),
          axis.line = element_line(colour = "black")) +
    coord_cartesian(ylim=c(0,15), xlim = c(150, 400)) +
    scale_y_continuous(breaks = seq(0,15,5))

    print(p)
```

```
## `geom_smooth()` using formula = 'y ~ x'
```

```
ggsave(paste0("results/","MoCA.MIS_right.presubiculum.AD.png"), width = 6,
       height = 4, units = c("in"), dpi = 200)
```

`geom_smooth()` using formula = 'y ~ x'

```
# CN-NC, left presubiculum

tmp <- all_data_subfileds_rm[all_data_subfileds_rm$group == "NonConverted_CN" &
                     all_data_subfileds_rm$LorR == "Left",]

cor.test(tmp$MoCA.MIS[tmp$group == "NonConverted_CN"],
    tmp$presubiculum[tmp$group == "NonConverted_CN"])
```

```
##
##  Pearson's product-moment correlation
##
## data:  tmp$MoCA.MIS[tmp$group == "NonConverted_CN"] and tmp$presubiculum[tmp$group == "NonConverted_(
## t = 3.824, df = 104, p-value = 0.0002241
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.1718542 0.5078429
## sample estimates:
##       cor
## 0.3510998
```

```
plm <- lm(formula = MoCA.MIS ~ presubiculum + Gender + Age +eTIV_ml + Education,
        data=tmp)

anova(plm)
```

```
## Analysis of Variance Table
##
## Response: MoCA.MIS
##              Df Sum Sq Mean Sq F value    Pr(>F)
## presubiculum  1  87.01  87.008 14.6707 0.0002259 ***
## Gender        1   0.00   0.005  0.0008 0.9780065
## Age           1  13.69  13.695  2.3091 0.1318382
## eTIV_ml       1   0.41   0.411  0.0693 0.7929319
## Education     1   0.89   0.886  0.1494 0.6999413
## Residuals    98 581.22   5.931
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
tmp$group <- droplevels(tmp$group)

# plot with lm for outcome ~ pred + factor
p <- ggplot(data=tmp, aes(x=presubiculum, y=MoCA.MIS, color = group)) +
    geom_point(size = 5, col = '#E41A1C', alpha = 0.8) + #col = 'blue',
    geom_smooth(method = "lm", mapping=aes(y=predict(plm,tmp)),
            col = 'black',linewidth = 2, se = FALSE) +
    guides(color = "none") +
    labs(x = "", y = "") +
    theme(legend.title = element_text(colour="black", size=14, face="bold"),
          legend.text = element_text(colour="black", size=14, face="bold")) +
    theme(plot.title = element_text(hjust = 0.5))+
    theme(axis.title = element_blank(),
          axis.text = element_text(size = 16, face = "bold")) +
    theme(panel.background = element_blank(),
          panel.border = element_blank(),
          panel.grid = element_blank(),
          axis.line = element_line(colour = "black")) +
    coord_cartesian(ylim=c(0,15), xlim = c(150, 450)) +
    scale_y_continuous(breaks = seq(0,15,5))

return(anova(plm))
```

```
## Analysis of Variance Table
##
## Response: MoCA.MIS
##              Df Sum Sq Mean Sq F value    Pr(>F)
## presubiculum  1  87.01  87.008 14.6707 0.0002259 ***
## Gender        1   0.00   0.005  0.0008 0.9780065
## Age           1  13.69  13.695  2.3091 0.1318382
## eTIV_ml       1   0.41   0.411  0.0693 0.7929319
## Education     1   0.89   0.886  0.1494 0.6999413
## Residuals    98 581.22   5.931
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```
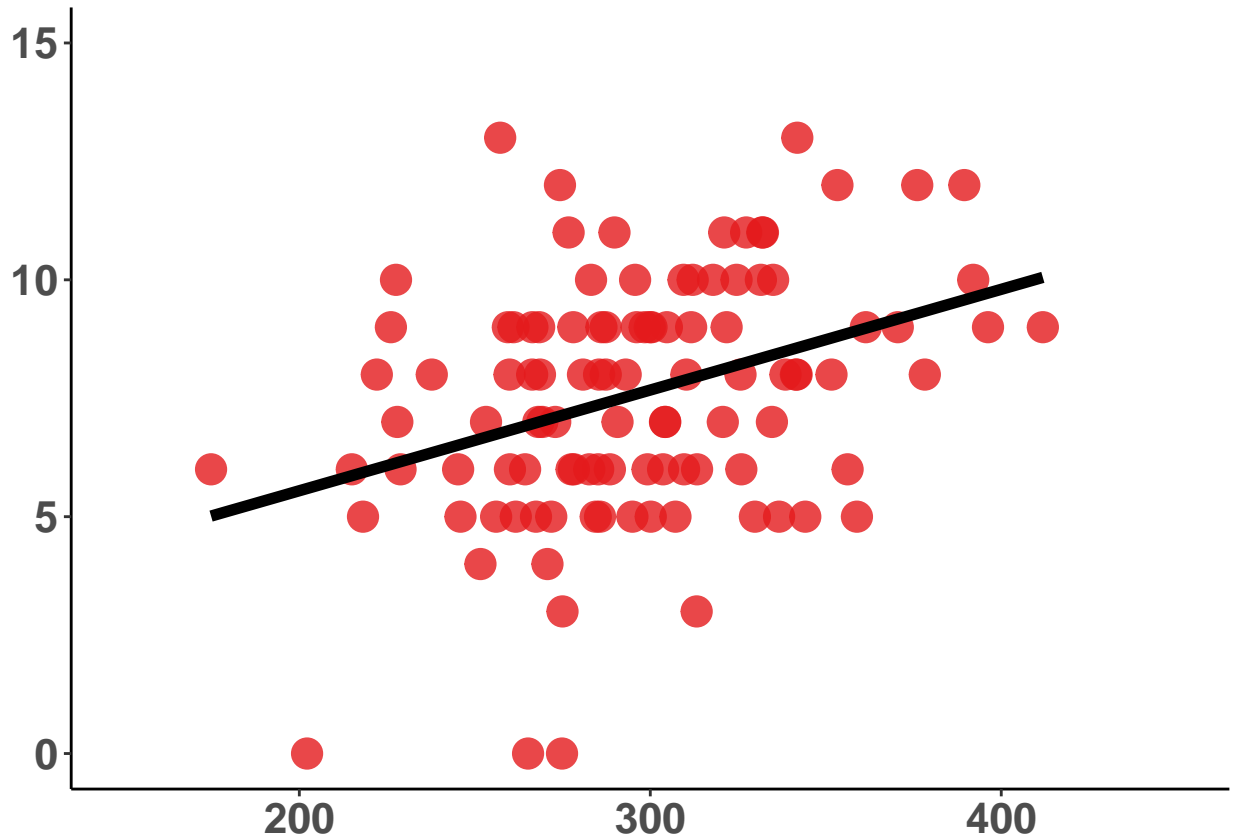
```
print(p)
```

`geom_smooth()` using formula = 'y ~ x'


```
ggsave(paste0("results/","MoCA.MIS.left.presubiculum.CN.png"), width = 6,
        height = 4, units = c("in"), dpi = 200)
```

`geom_smooth()` using formula = 'y ~ x'

```
# CN-NC, right presubiculum

tmp <- all_data_subfileds_rm[all_data_subfileds_rm$group == "NonConverted_CN" &
                all_data_subfileds_rm$LorR == "Right",]

cor.test(tmp$MoCA.MIS[tmp$group == "NonConverted_CN"],
    tmp$presubiculum[tmp$group == "NonConverted_CN"])
```

```
##
##  Pearson's product-moment correlation
##
## data:  tmp$MoCA.MIS[tmp$group == "NonConverted_CN"] and tmp$presubiculum[tmp$group == "NonConverted_
## t = 3.2388, df = 104, p-value = 0.001612
## alternative hypothesis: true correlation is not equal to 0
```

```
## 95 percent confidence interval:
##  0.1187933 0.4665089
## sample estimates:
##       cor
## 0.3026896
```

```r
plm <- lm(formula = MoCA.MIS ~ presubiculum + Gender + Age +eTIV_ml + Education,
      data=tmp)

anova(plm)
```

```
## Analysis of Variance Table
##
## Response: MoCA.MIS
##               Df Sum Sq Mean Sq F value  Pr(>F)
## presubiculum  1  65.28  65.280 10.7136 0.00147 **
## Gender        1   0.00   0.004  0.0006 0.98012
## Age           1  19.15  19.152  3.1432 0.07935 .
## eTIV_ml       1   1.21   1.206  0.1979 0.65741
## Education     1   0.44   0.445  0.0730 0.78765
## Residuals    98 597.14   6.093
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
tmp$group <- droplevels(tmp$group)

# plot with lm for outcome ~ pred + factor
p <- ggplot(data=tmp, aes(x=presubiculum, y=MoCA.MIS, color = group)) +
    geom_point(size = 5, col = '#E41A1C', alpha = 0.8) + #col = 'blue',
    geom_smooth(method = "lm", mapping=aes(y=predict(plm,tmp)),
            col = 'black',linewidth = 2, se = FALSE) +
    guides(color = "none") +
    labs(x = "", y = "") +
    theme(legend.title = element_text(colour="black", size=14, face="bold"),
          legend.text = element_text(colour="black", size=14, face="bold")) +
    theme(plot.title = element_text(hjust = 0.5))+
    theme(axis.title = element_blank(),
          axis.text = element_text(size = 16, face = "bold")) +
    theme(panel.background = element_blank(),
          panel.border = element_blank(),
          panel.grid = element_blank(),
          axis.line = element_line(colour = "black")) +
    coord_cartesian(ylim=c(0,15), xlim = c(150, 450)) +
    scale_y_continuous(breaks = seq(0,15,5))

return(anova(plm))
```
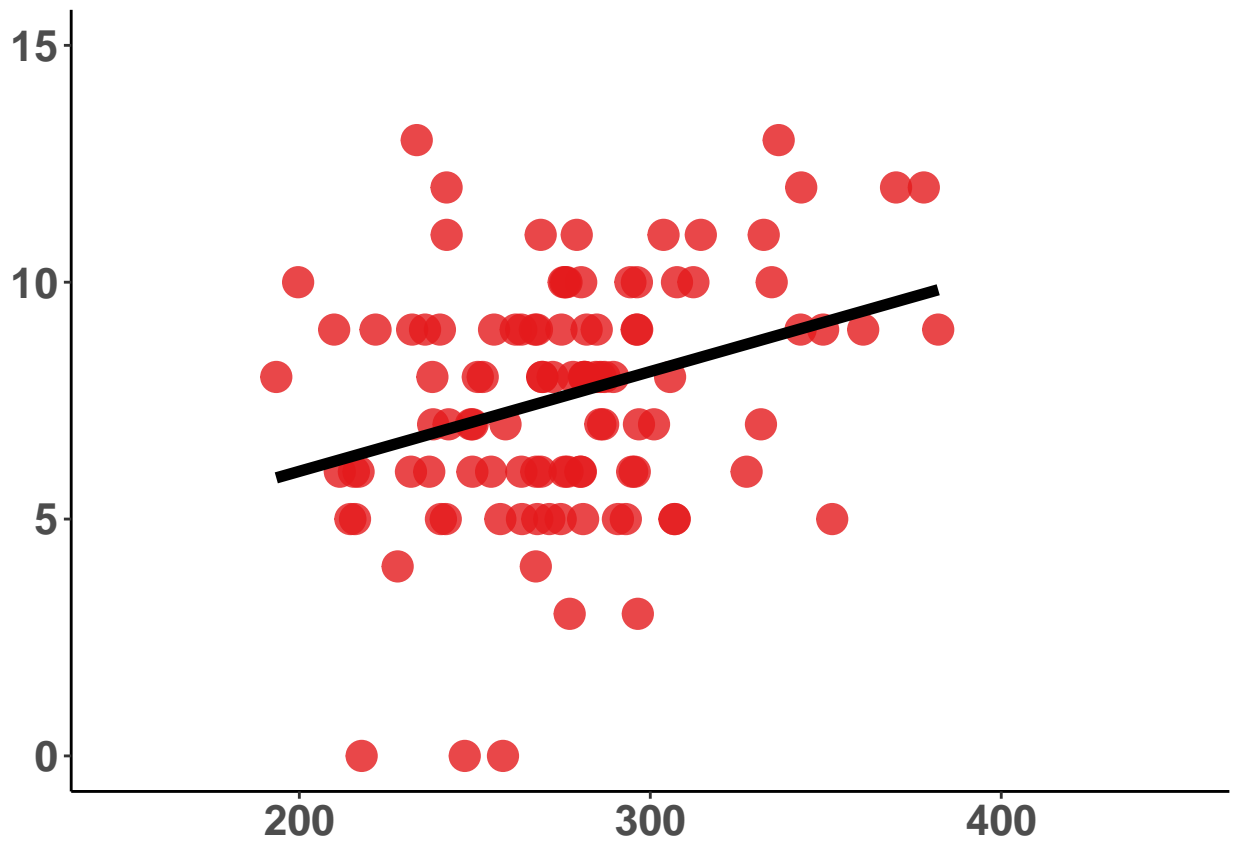
```
## Analysis of Variance Table
##
## Response: MoCA.MIS
##               Df Sum Sq Mean Sq F value  Pr(>F)
## presubiculum  1  65.28  65.280 10.7136 0.00147 **
## Gender        1   0.00   0.004  0.0006 0.98012
```

```
## Age         1  19.15  19.152  3.1432 0.07935 .
## eTIV_ml      1   1.21   1.206  0.1979 0.65741
## Education    1   0.44   0.445  0.0730 0.78765
## Residuals   98 597.14   6.093
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
print(p)
```

```
## `geom_smooth()` using formula = 'y ~ x'
```



```
ggsave(paste0("results/","MoCA.MIS.right.presubiculum.CN.png"), width = 6,
        height = 4, units = c("in"), dpi = 200)
```

```
## `geom_smooth()` using formula = 'y ~ x'
```