ADNI Project

Soham Dighe

2022-09-26

Analysis of our datasets

Checking if the difference in age is significant

```
control = read.csv(file = "/Users/sohamdighe/Documents/SCIE90017/Data/ADNI/control.csv")
diabetes = read.csv(file = "/Users/sohamdighe/Documents/SCIE90017/Data/ADNI/diabetes.csv")
control$DX = as.factor(control$DX)
control$APOE4 = as.factor(control$APOE4)
control$ABETA = as.numeric(control$ABETA)
```

```
## Warning: NAs introduced by coercion
```

```
control$PTGENDER = as.factor(control$PTGENDER)
control$PTETHCAT = as.factor(control$PTETHCAT)
control$PTRACCAT = as.factor(control$PTRACCAT)

diabetes$DX = as.factor(diabetes$DX)
diabetes$APOE4 = as.factor(diabetes$APOE4)
diabetes$ABETA = as.numeric(diabetes$ABETA)
```
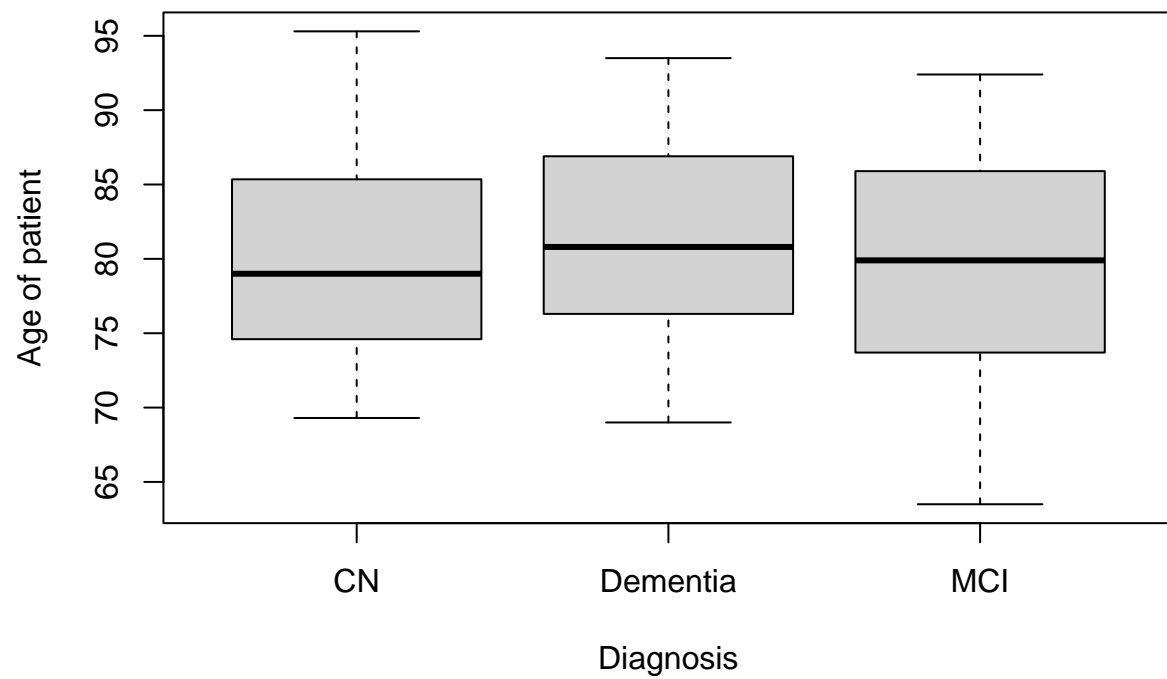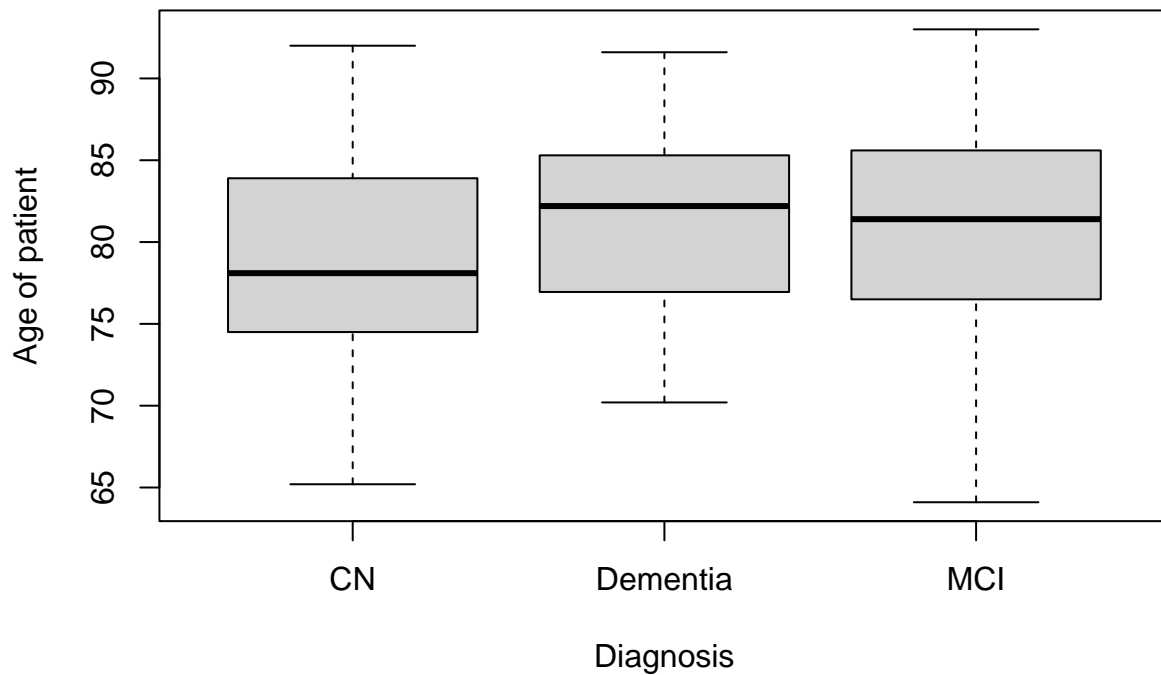
```
## Warning: NAs introduced by coercion
```

```
diabetes$PTGENDER = as.factor(diabetes$PTGENDER)
diabetes$PTETHCAT = as.factor(diabetes$PTETHCAT)
diabetes$PTRACCAT = as.factor(diabetes$PTRACCAT)
diabetes$PTAU = as.numeric(diabetes$PTAU)
```

```
## Warning: NAs introduced by coercion
```

```
plot(y = control$AGE, x = control$DX, xlab = "Diagnosis", ylab = "Age of patient")
```

```
plot(y = diabetes$AGE, x = diabetes$DX, xlab = "Diagnosis", ylab = "Age of patient")
```

Checking normality of age through Shapiro-Wilk test

```
shapiro.test(control$AGE)
```
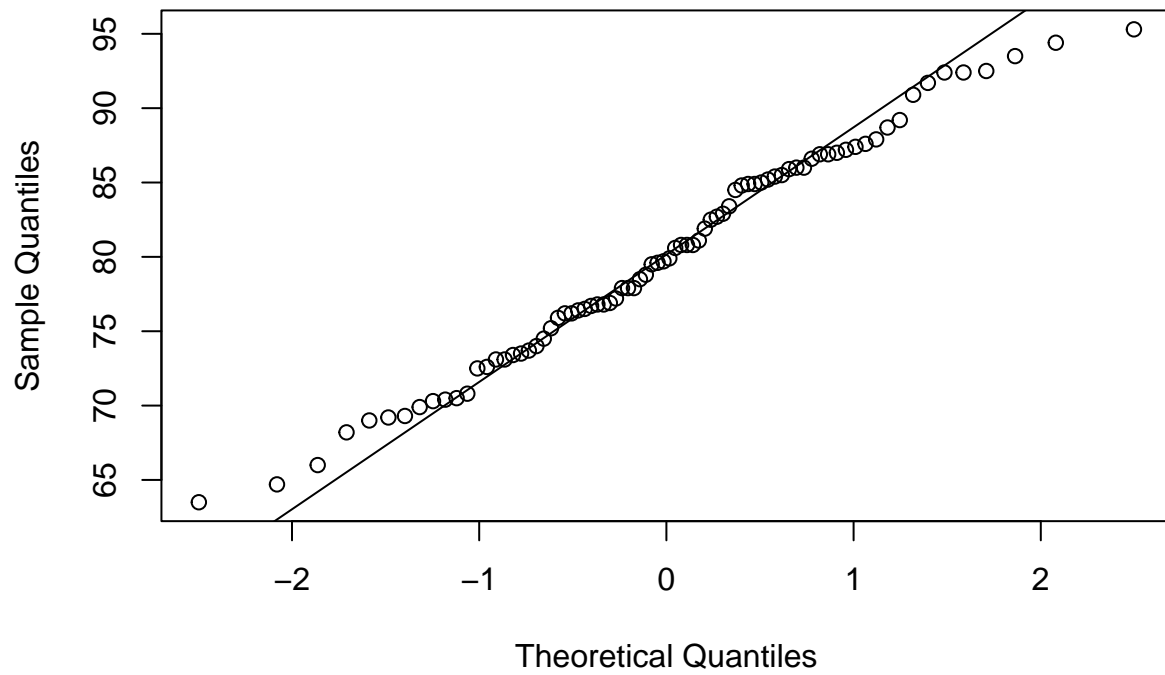
```
##
##  Shapiro-Wilk normality test
##
## data:  control$AGE
## W = 0.98419, p-value = 0.4291
```
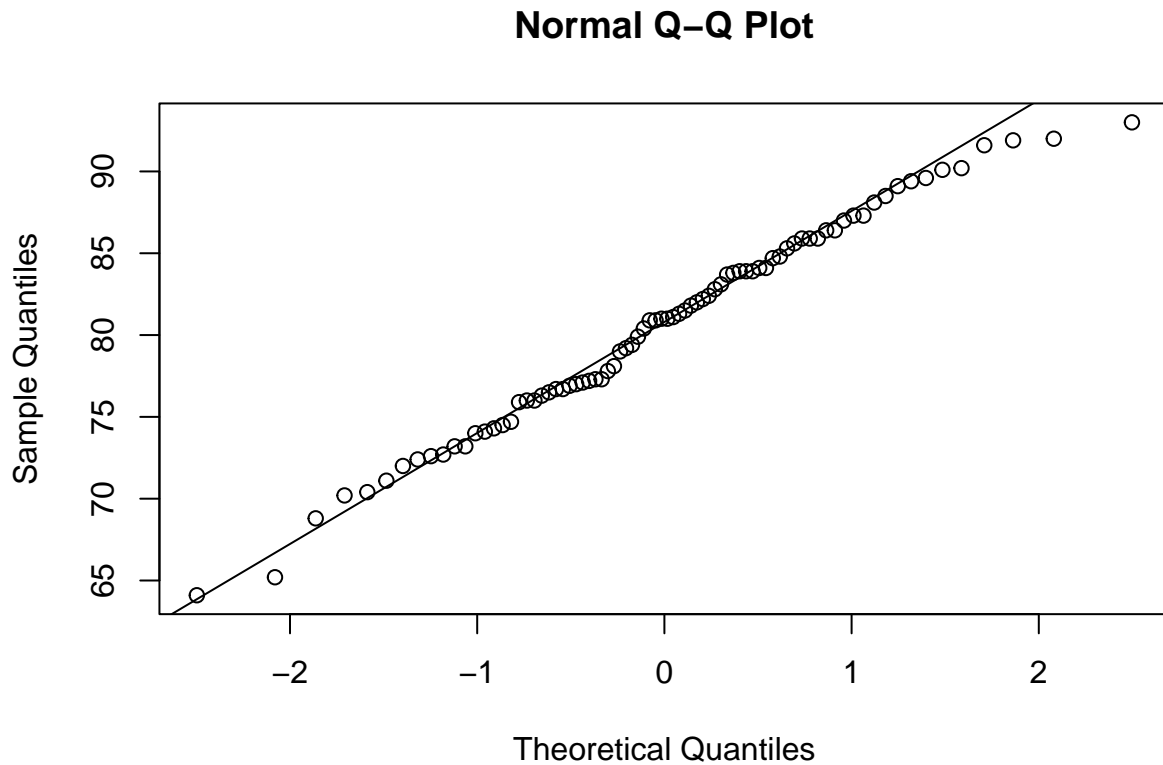
```
shapiro.test(diabetes$AGE)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  diabetes$AGE
## W = 0.98648, p-value = 0.5649
```

```
# trying to plot the data anyways
qqnorm(control$AGE)
qqline(control$AGE)
```

Normal Q–Q Plot


```
qqnorm(diabetes$AGE)
qqline(diabetes$AGE)
```

Normal Q–Q Plot

Now, we have established both are normally distributed so we will do t test on them.

```
t.test(control$AGE, diabetes$AGE)
```

```
##
##  Welch Two Sample t-test
##
## data:  control$AGE and diabetes$AGE
## t = -0.41517, df = 154.27, p-value = 0.6786
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -2.648763  1.728763
## sample estimates:
## mean of x mean of y
##    80.085    80.545
```

From the above test we reject the alternativce hypothesis and hence the true difference between means is zero and hence is not statistically significant.

Table 1

Here, we have a caveat. We still have not replaced missing data in the files so the values are not final but should give a decent estimate. FDG,PIB,FBB,AV45 and CSF markers are the ones with most null values. This is still to be addressed.

```
variables_selected = c("AGE", "MMSE", "CDRSB", "FDG", "PIB", "AV45", "PTGENDER",
    "PTEDUCAT", "PTETHCAT", "PTRACCAT", "APOE4", "ABETA", "PTAU", "TAU", "Ventricles",
    "Hippocampus", "mPACCdigit", "mPACCtrailsB")
control_tab <- CreateTableOne(vars = variables_selected, data = control, strata = "DX")
control_tab
```

```
##                              Stratified by DX
##                               CN                     Dementia
##   n                                 24                     23
##   AGE (mean (SD))              80.00 (7.08)           80.73 (7.24)
##   MMSE (mean (SD))             28.54 (1.32)           19.48 (6.27)
##   CDRSB (mean (SD))             0.21 (0.49)            8.20 (4.54)
##   FDG (mean (SD))               1.28 (0.07)            1.16 (0.10)
##   PIB (mean (SD))               NaN (NA)               NaN (NA)
##   AV45 (mean (SD))              1.21 (0.32)            1.43 (0.34)
##   PTGENDER = Male (%)             12 (50.0)              11 ( 47.8)
##   PTEDUCAT (mean (SD))         16.12 (2.31)           15.43 (2.48)
##   PTETHCAT = Not Hisp/Latino (%)  23 (95.8)              23 (100.0)
##   PTRACCAT (%)
##      Black                         0 ( 0.0)               1 (  4.3)
##      More than one                 1 ( 4.2)               0 (  0.0)
##      White                        23 (95.8)              22 ( 95.7)
##   APOE4 (%)
##      0                            20 (83.3)               5 ( 21.7)
##      1                             4 (16.7)              11 ( 47.8)
##      2                             0 ( 0.0)               7 ( 30.4)
##   ABETA (mean (SD))           918.44 (334.52)        612.79 (160.71)
##   PTAU (mean (SD))             22.88 (12.22)          38.47 (19.68)
##   TAU (mean (SD))             244.80 (107.77)        383.86 (172.22)
##   Ventricles (mean (SD))    35200.81 (19822.89) 52609.23 (19037.64)
##   Hippocampus (mean (SD))    7051.67 (990.10)     5368.87 (1046.44)
##   mPACCdigit (mean (SD))        -0.54 (3.36)          -21.15 (9.81)
##   mPACCtrailsB (mean (SD))      -0.77 (3.27)          -19.26 (10.43)
##                              Stratified by DX
##                               MCI                    p      test
##   n                                 33
##   AGE (mean (SD))              79.70 (8.22)          0.883
##   MMSE (mean (SD))             27.85 (2.00)         <0.001
##   CDRSB (mean (SD))             1.11 (0.91)         <0.001
##   FDG (mean (SD))               1.28 (0.06)         <0.001
##   PIB (mean (SD))               1.59 (0.48)            NA
##   AV45 (mean (SD))              1.15 (0.22)          0.204
##   PTGENDER = Male (%)             21 ( 63.6)          0.424
##   PTEDUCAT (mean (SD))         16.18 (2.91)          0.537
##   PTETHCAT = Not Hisp/Latino (%)  33 (100.0)          0.307
##   PTRACCAT (%)                                        0.511
##      Black                         1 (  3.0)
##      More than one                 0 (  0.0)
```

```
##       White                              32 ( 97.0)
##    APOE4 (%)                                              <0.001
##       0                                  21 ( 63.6)
##       1                                  11 ( 33.3)
##       2                                   1 (  3.0)
##    ABETA (mean (SD))                 892.64 (323.37)    0.003
##    PTAU (mean (SD))                   21.70 (10.30)    <0.001
##    TAU (mean (SD))                   231.75 (95.87)    <0.001
##    Ventricles (mean (SD))         49233.90 (26546.13)   0.022
##    Hippocampus (mean (SD))          6719.16 (1376.91)  <0.001
##    mPACCdigit (mean (SD))               -4.28 (5.43)   <0.001
##    mPACCtrailsB (mean (SD))             -3.96 (4.80)   <0.001
```

```r
diabetes_tab <- CreateTableOne(vars = variables_selected, data = diabetes, strata = "DX")
diabetes_tab
```

```
##                              Stratified by DX
##                               CN                   Dementia
##    n                           27                       23
##    AGE (mean (SD))             79.52 (6.81)         81.40 (5.56)
##    MMSE (mean (SD))            28.56 (2.95)         20.78 (4.88)
##    CDRSB (mean (SD))            1.02 (3.12)          9.91 (4.56)
##    FDG (mean (SD))              1.29 (0.06)          1.21 (0.11)
##    PIB (mean (SD))               NaN (NA)            2.19 (0.00)
##    AV45 (mean (SD))             0.98 (0.04)          1.52 (0.30)
##    PTGENDER = Male (%)         16 (59.3)            16 (69.6)
##    PTEDUCAT (mean (SD))        16.44 (2.36)         15.57 (2.57)
##    PTETHCAT = Not Hisp/Latino (%)   26 (96.3)       22 (95.7)
##    PTRACCAT (%)
##       Am Indian/Alaskan         1 ( 3.7)             0 ( 0.0)
##       Asian                     1 ( 3.7)             1 ( 4.3)
##       Black                     3 (11.1)             0 ( 0.0)
##       More than one             1 ( 3.7)             1 ( 4.3)
##       White                    21 (77.8)            21 (91.3)
##    APOE4 (%)
##       0                        18 (66.7)             9 (39.1)
##       1                         7 (25.9)            13 (56.5)
##       2                         2 ( 7.4)             1 ( 4.3)
##    ABETA (mean (SD))         1073.12 (341.06)     893.26 (464.81)
##    PTAU (mean (SD))            20.04 (6.88)         31.84 (13.27)
##    TAU (mean (SD))            218.37 (81.24)       336.64 (121.48)
##    Ventricles (mean (SD))  46500.65 (26484.13) 58264.68 (27618.48)
##    Hippocampus (mean (SD))  7193.09 (831.50)    5569.21 (1519.57)
##    mPACCdigit (mean (SD))       0.62 (6.20)        -18.74 (7.30)
##    mPACCtrailsB (mean (SD))    -0.10 (5.98)        -16.67 (6.48)
##                              Stratified by DX
##                               MCI                  p       test
##    n                           30
##    AGE (mean (SD))             80.82 (6.81)         0.571
##    MMSE (mean (SD))            27.47 (1.94)        <0.001
##    CDRSB (mean (SD))            1.93 (1.76)        <0.001
##    FDG (mean (SD))              1.26 (0.08)         0.010
##    PIB (mean (SD))             1.86 (0.74)          0.593
##    AV45 (mean (SD))            1.29 (0.27)          0.014
```

```
##    PTGENDER = Male (%)                   25 ( 83.3)      0.131
##    PTEDUCAT (mean (SD))              16.30 (2.34)        0.398
##    PTETHCAT = Not Hisp/Latino (%)      30 (100.0)        0.535
##    PTRACCAT (%)                                          0.676
##       Am Indian/Alaskan                1 (  3.3)
##       Asian                            0 (  0.0)
##       Black                            1 (  3.3)
##       More than one                    1 (  3.3)
##       White                           27 ( 90.0)
##    APOE4 (%)                                             0.059
##       0                               22 ( 73.3)
##       1                                8 ( 26.7)
##       2                                0 (  0.0)
##    ABETA (mean (SD))               884.03 (332.59)       0.316
##    PTAU (mean (SD))                 24.22 (10.31)        0.005
##    TAU (mean (SD))                 257.51 (100.68)       0.004
##    Ventricles (mean (SD))       53215.74 (23361.72)      0.272
##    Hippocampus (mean (SD))       6817.54 (1077.71)      <0.001
##    mPACCdigit (mean (SD))           -5.19 (5.58)        <0.001
##    mPACCtrailsB (mean (SD))         -4.74 (4.87)        <0.001
```

```
cohort = rbind(control, diabetes)
tab <- CreateTableOne(vars = variables_selected, data = cohort, strata = "DX")
tab
```

```
##                                 Stratified by DX
##                                  CN                   Dementia
##    n                                51                   46
##    AGE (mean (SD))               79.74 (6.87)         81.06 (6.39)
##    MMSE (mean (SD))              28.55 (2.31)         20.13 (5.60)
##    CDRSB (mean (SD))             0.64 (2.31)           9.05 (4.58)
##    FDG (mean (SD))               1.29 (0.06)           1.19 (0.11)
##    PIB (mean (SD))                NaN (NA)             2.19 (0.00)
##    AV45 (mean (SD))              1.13 (0.28)           1.48 (0.31)
##    PTGENDER = Male (%)             28 (54.9)             27 (58.7)
##    PTEDUCAT (mean (SD))          16.29 (2.32)         15.50 (2.50)
##    PTETHCAT = Not Hisp/Latino (%)   49 (96.1)            45 (97.8)
##    PTRACCAT (%)
##       Black                        3 ( 5.9)              1 ( 2.2)
##       More than one                2 ( 3.9)              1 ( 2.2)
##       White                       44 (86.3)             43 (93.5)
##       Am Indian/Alaskan            1 ( 2.0)              0 ( 0.0)
##       Asian                        1 ( 2.0)              1 ( 2.2)
##    APOE4 (%)
##       0                           38 (74.5)             14 (30.4)
##       1                           11 (21.6)             24 (52.2)
##       2                            2 ( 3.9)              8 (17.4)
##    ABETA (mean (SD))           992.91 (340.35)      726.73 (343.82)
##    PTAU (mean (SD))             21.50 (9.96)         35.74 (17.41)
##    TAU (mean (SD))             231.59 (95.26)       364.41 (153.12)
##    Ventricles (mean (SD))   41183.08 (24040.29) 55436.96 (23627.90)
##    Hippocampus (mean (SD))   7126.54 (903.08)     5469.04 (1293.35)
##    mPACCdigit (mean (SD))       0.07 (5.05)         -19.94 (8.64)
##    mPACCtrailsB (mean (SD))    -0.41 (4.86)         -17.97 (8.68)
```

8

```
##                              Stratified by DX
##                               MCI                    p       test
##   n                               63
##   AGE (mean (SD))               80.23 (7.54)        0.649
##   MMSE (mean (SD))              27.67 (1.97)       <0.001
##   CDRSB (mean (SD))             1.50 (1.43)        <0.001
##   FDG (mean (SD))               1.27 (0.07)        <0.001
##   PIB (mean (SD))               1.73 (0.53)         0.311
##   AV45 (mean (SD))             1.22 (0.25)         0.009
##   PTGENDER = Male (%)           46 ( 73.0)          0.105
##   PTEDUCAT (mean (SD))         16.24 (2.63)         0.218
##   PTETHCAT = Not Hisp/Latino (%)    63 (100.0)      0.303
##   PTRACCAT (%)                                      0.860
##     Black                         2 (  3.2)
##     More than one                 1 (  1.6)
##     White                        59 ( 93.7)
##     Am Indian/Alaskan             1 (  1.6)
##     Asian                         0 (  0.0)
##   APOE4 (%)                                        <0.001
##     0                            43 ( 68.3)
##     1                            19 ( 30.2)
##     2                             1 (  1.6)
##   ABETA (mean (SD))           888.03 (324.44)      0.010
##   PTAU (mean (SD))             23.01 (10.28)      <0.001
##   TAU (mean (SD))             245.14 (98.26)      <0.001
##   Ventricles (mean (SD))   51130.02 (24960.22)    0.013
##   Hippocampus (mean (SD))   6766.01 (1234.60)    <0.001
##   mPACCdigit (mean (SD))       -4.71 (5.48)      <0.001
##   mPACCtrailsB (mean (SD))     -4.33 (4.81)      <0.001
```