

```

    if (found) {
      keep_me = c(keep_me, subj_idx[cur_choice])
    }
  }
  filtered_data = a[keep_me, ]

  a = filtered_data
  print(sprintf('After keeping 2 best scans: %d scans, %d subjects, %d
families, %d ADHD',
               nrow(a),
               length(unique(a$Medical.Record...MRN...Subjects)),
               length(unique(a$FAMID)),
               sum(a$DX=='ADHD'))))

```

[1] "After keeping 2 best scans: 504 scans, 252 subjects, 185 families, 213 ADHD"

Let's give the mean and sd of each qc var and tract_var, and then make histograms of the tract_vars ala ENIGMA. Or, even better, show that the scans that stayed are better than the ones that were removed?

```

for (v in qc_vars) {
  print(sprintf('%s: %.2f (+- %.2f)', v, mean(a[, v]), sd(a[, v])))
}

```

[1] "meanX.trans: 0.18 (+- 0.13)" [1] "meanY.trans: 0.38 (+- 0.23)" [1] "meanZ.trans: 0.82 (+- 0.77)" [1]
 "meanX.rot: 0.01 (+- 0.01)" [1] "meanY.rot: 0.01 (+- 0.00)" [1] "meanZ.rot: 0.00 (+- 0.00)" [1]
 "goodVolumes: 59.67 (+- 1.55)"

```

mres$kept = F
mres[mres$Mask.ID...Scan %in% a$Mask.ID...Scan, 'kept'] = T
for (v in qc_vars) {
  print(v)
  print(t.test(mres[mres$kept==T,v], mres[mres$kept==F,v]))
}

```

[1] "meanX.trans"

Welch Two Sample t-test

```

data: mres[mres$kept == T, v] and mres[mres$kept == F, v]
t = -3.901, df = 628.93, p-value = 0.0001061
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.08156248 -0.02694185
sample estimates:
mean of x mean of y
0.1752001 0.2294523

```

```
[1] "meanY.trans"
```

```
Welch Two Sample t-test
```

```
data: mres[mres$kept == T, v] and mres[mres$kept == F, v]
t = -4.597, df = 532.29, p-value = 5.36e-06
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.23135628 -0.09282376
sample estimates:
mean of x mean of y
0.3809131 0.5430031
```

```
[1] "meanZ.trans"
```

```
Welch Two Sample t-test
```

```
data: mres[mres$kept == T, v] and mres[mres$kept == F, v]
t = -4.4124, df = 601.88, p-value = 1.211e-05
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.5710976 -0.2192998
sample estimates:
mean of x mean of y
0.8158318 1.2110305
```

```
[1] "meanX.rot"
```

```
Welch Two Sample t-test
```

```
data: mres[mres$kept == T, v] and mres[mres$kept == F, v]
t = -5.1882, df = 587.54, p-value = 2.927e-07
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.010936408 -0.004930128
sample estimates:
mean of x mean of y
0.01289131 0.02082458
```

```
[1] "meanY.rot"
```

```
Welch Two Sample t-test
```

```
data: mres[mres$kept == T, v] and mres[mres$kept == F, v]
t = -3.4931, df = 717.9, p-value = 0.0005068
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.0022381101 -0.0006274993
sample estimates:
mean of x mean of y
0.005377990 0.006810794
```

```
[1] "meanZ.rot"
```

Welch Two Sample t-test

```
data: mres[mres$kept == T, v] and mres[mres$kept == F, v]
t = -4.6084, df = 582.23, p-value = 4.989e-06
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.003341304 -0.001344337
sample estimates:
 mean of x mean of y
0.004142585 0.006485405

[1] "goodVolumes"
```

Welch Two Sample t-test

```
data: mres[mres$kept == T, v] and mres[mres$kept == F, v]
t = 6.9631, df = 561.16, p-value = 9.338e-12
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 1.029921 1.839285
sample estimates:
mean of x mean of y
 59.6746  58.2400
```

```
a =
read.csv('~data/heritability_change_rev/dti_JHUtracts_ADRDonly_0D0.95_two
TimePoints_noOtherDX.csv')
tract_names = colnames(a)[grepl(colnames(a), pattern="^ad")]
par(mfrow=c(5, 4))
for (v in tract_names) {
  t_str = sprintf('%s: %.2f (+- %.2f)', v, mean(a[, v]), sd(a[, v]))
  hist(a[, v], breaks=25, main=t_str)
}
tract_names = colnames(a)[grepl(colnames(a), pattern="^rd")]
par(mfrow=c(5, 4))
for (v in tract_names) {
  t_str = sprintf('%s: %.2f (+- %.2f)', v, mean(a[, v]), sd(a[, v]))
  hist(a[, v], breaks=25, main=t_str)
}
```



