Brain Paper Title Here

Jonathan Tay

November 15, 2017

Introduction

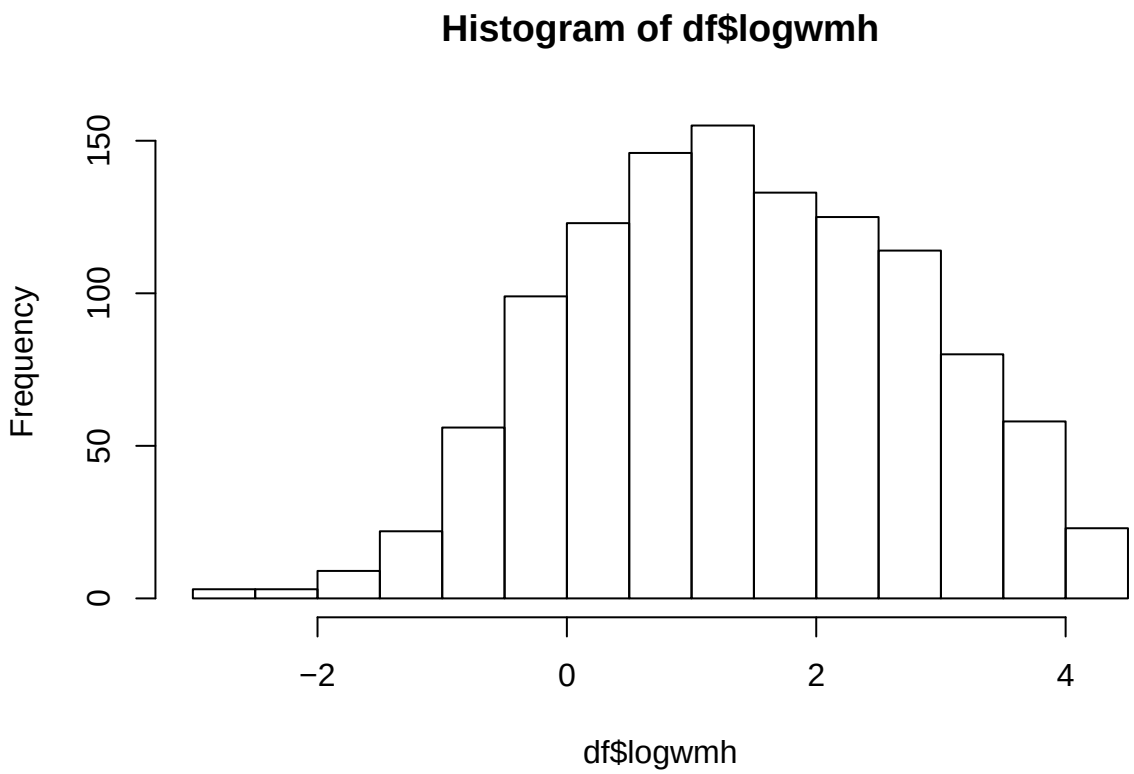
In our study, we sought to examine the longitudinal relationships between white matter hyperintensities, hippocampal volumes, and age, and how they may explain cognitive decline.

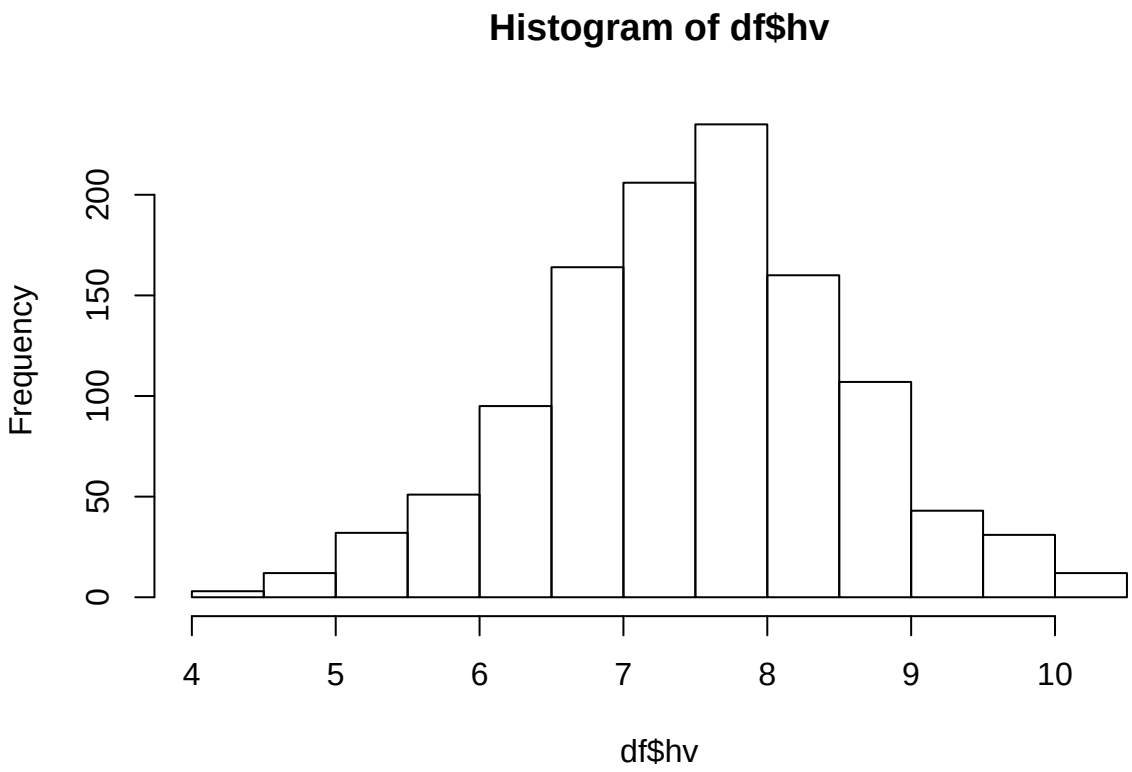
First, load up the data and apply transforms.

```
df <- read.csv("../data/RUNDMC_data_long.csv", sep=";", dec=",")
df$subject <- factor(df$rundmcs)
df$logwmh <- log(df$wmh)
df$age06 <- scale(df$age06)
df$agesq <- df$age06^2
df$timesq <- df$time^2
```

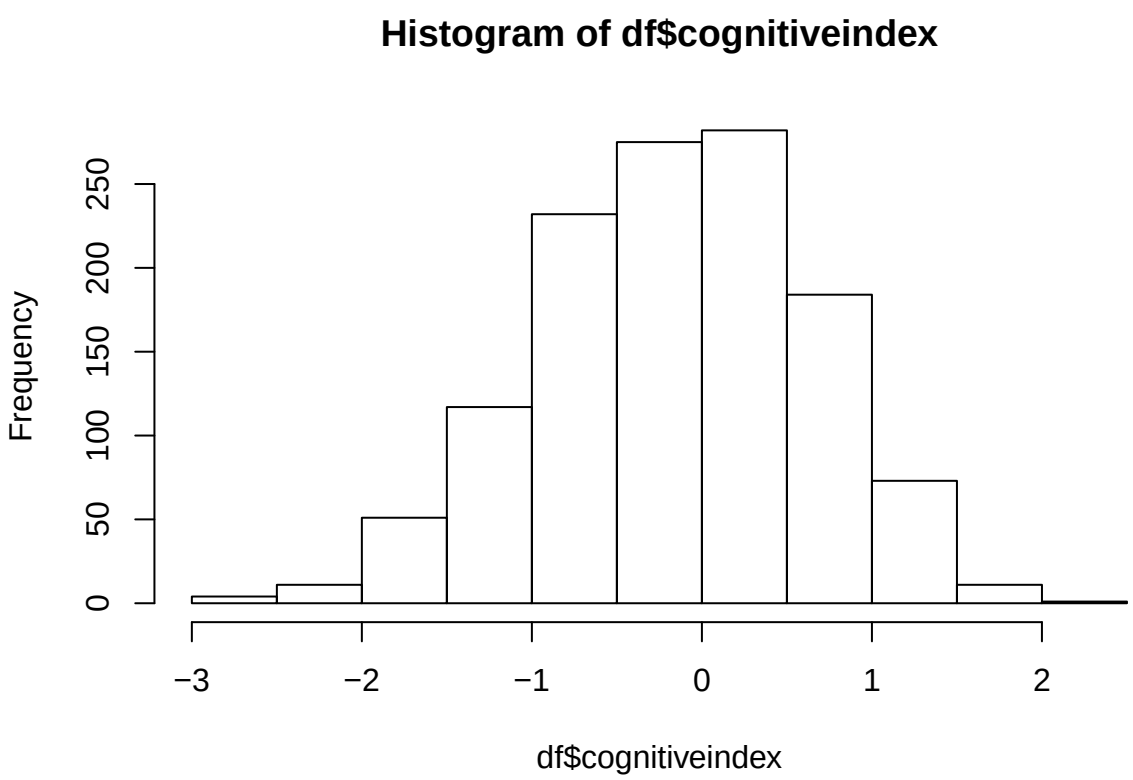
Visualizing the data

Since we want to fit a linear mixed model, let's make sure that all the variables are (roughly) normally distributed:




Histogram of df$hv
Frequency
0
50
100
150
200
4
5
6
7
8
9
10
df$hv

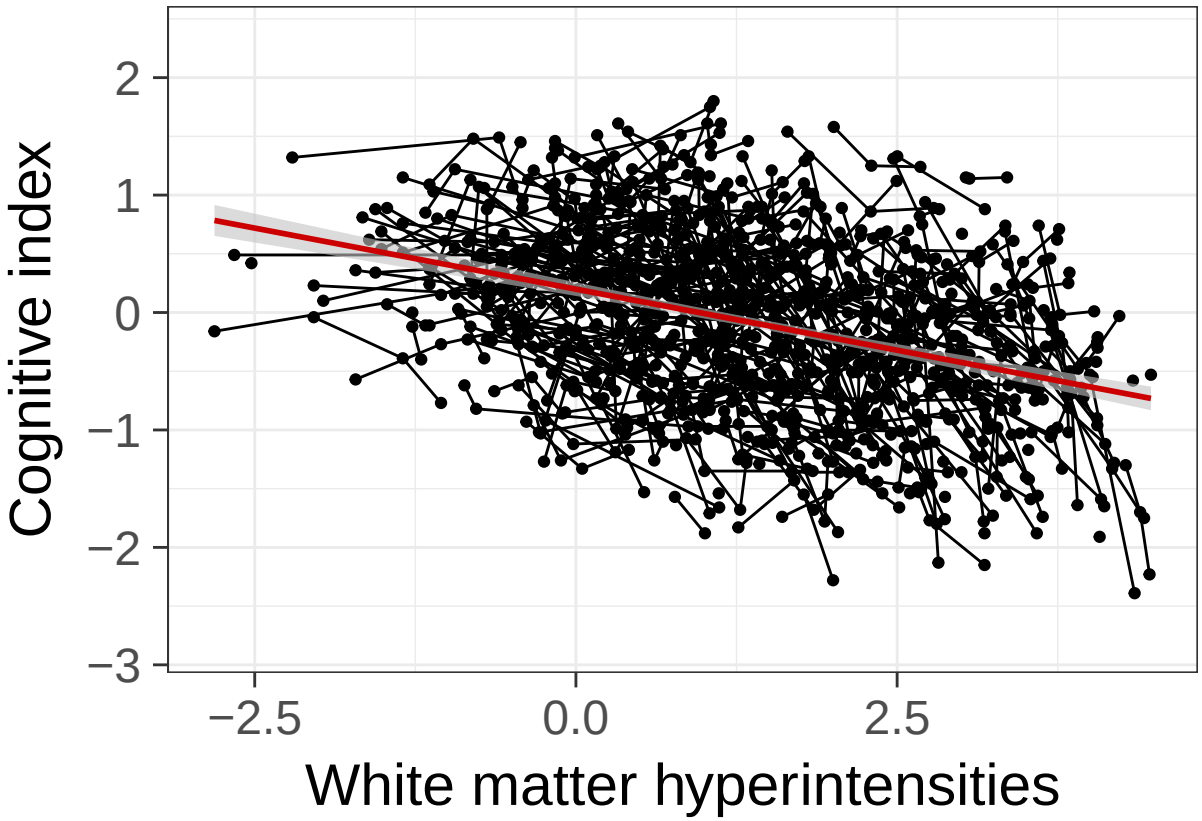

Histogram of df$age
Frequency
0
50
100
150
200
250
50
60
70
80
90
df$age

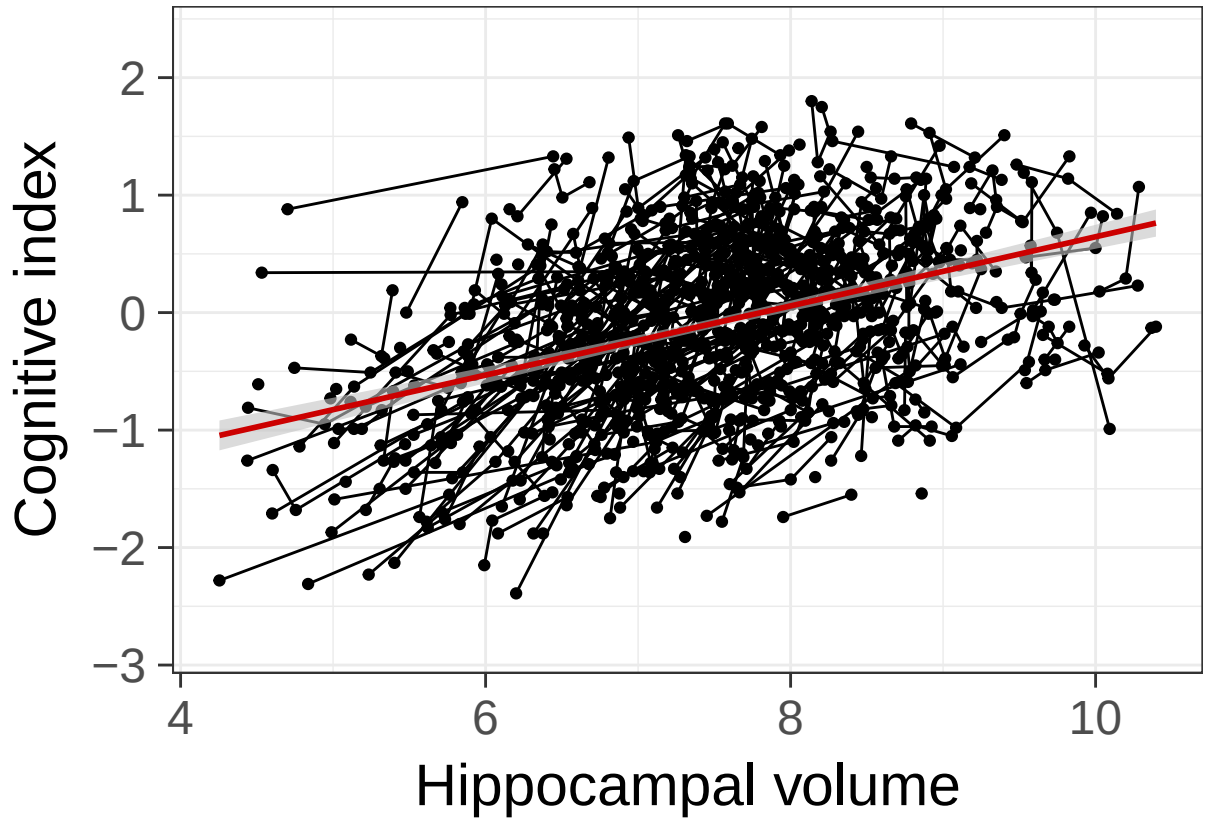


And then let's make sure that the relationship between IVs and cognitive decline is (roughly) linear. It would be easiest to do this with visualizations. So let's define a few functions for plotting:

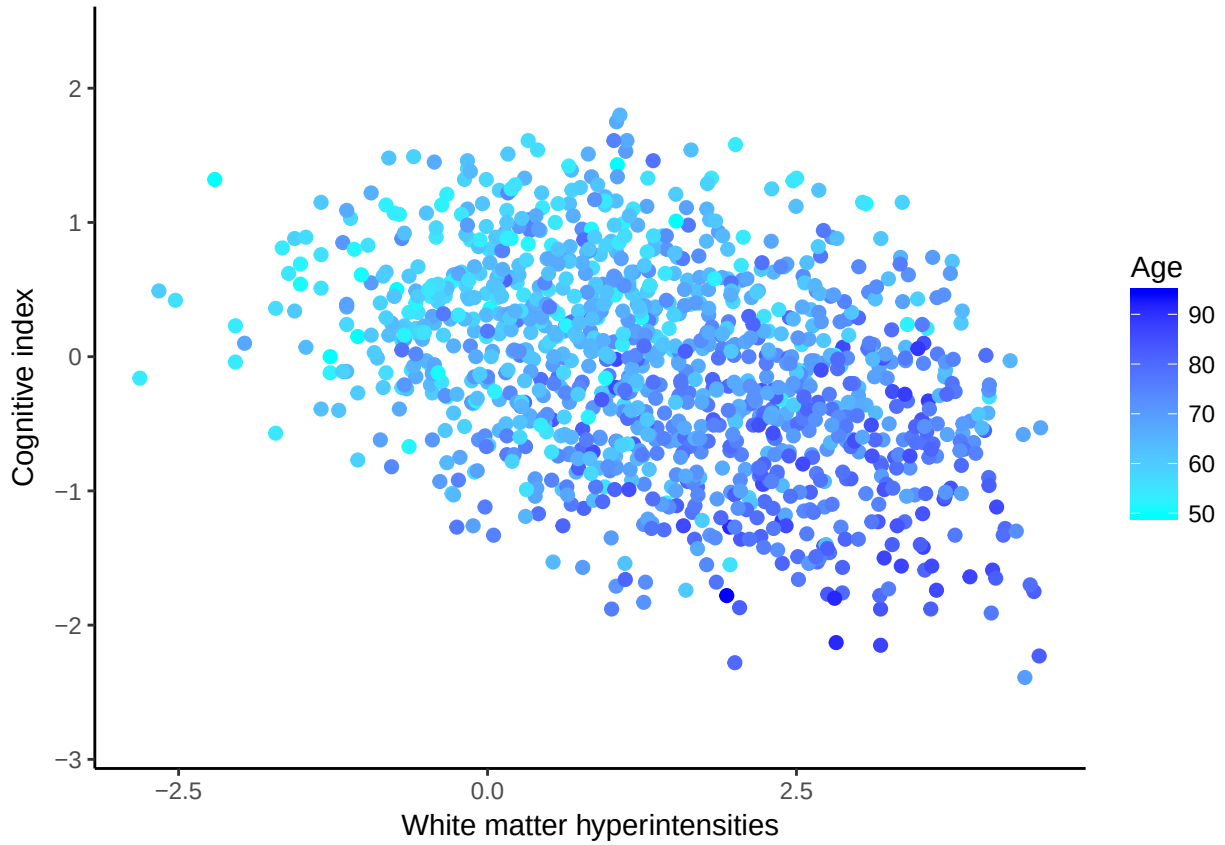
```
library(ggplot2)

spaghettiPlot <- function(x, y, group, xLab, yLab) {
    ggplot(df, aes(x, y, group=group)) +
    geom_line() + geom_point() +
    stat_smooth(aes(group=1), method=lm, colour="#cc0000", fill="#c3c3c3", alpha=0.6) +
    theme_bw(base_size=22) +
    labs(x=xLab, y=yLab)
}

scatterPlot <- function(x, y, m, xLab, yLab, mLab) {
  ggplot(df, aes(x, y)) + geom_point(aes(color=m), size=2) +
  scale_color_gradient(low="cyan", high="blue") +
  labs(x=xLab, y=yLab, colour=mLab) +
  theme_classic()
}
```

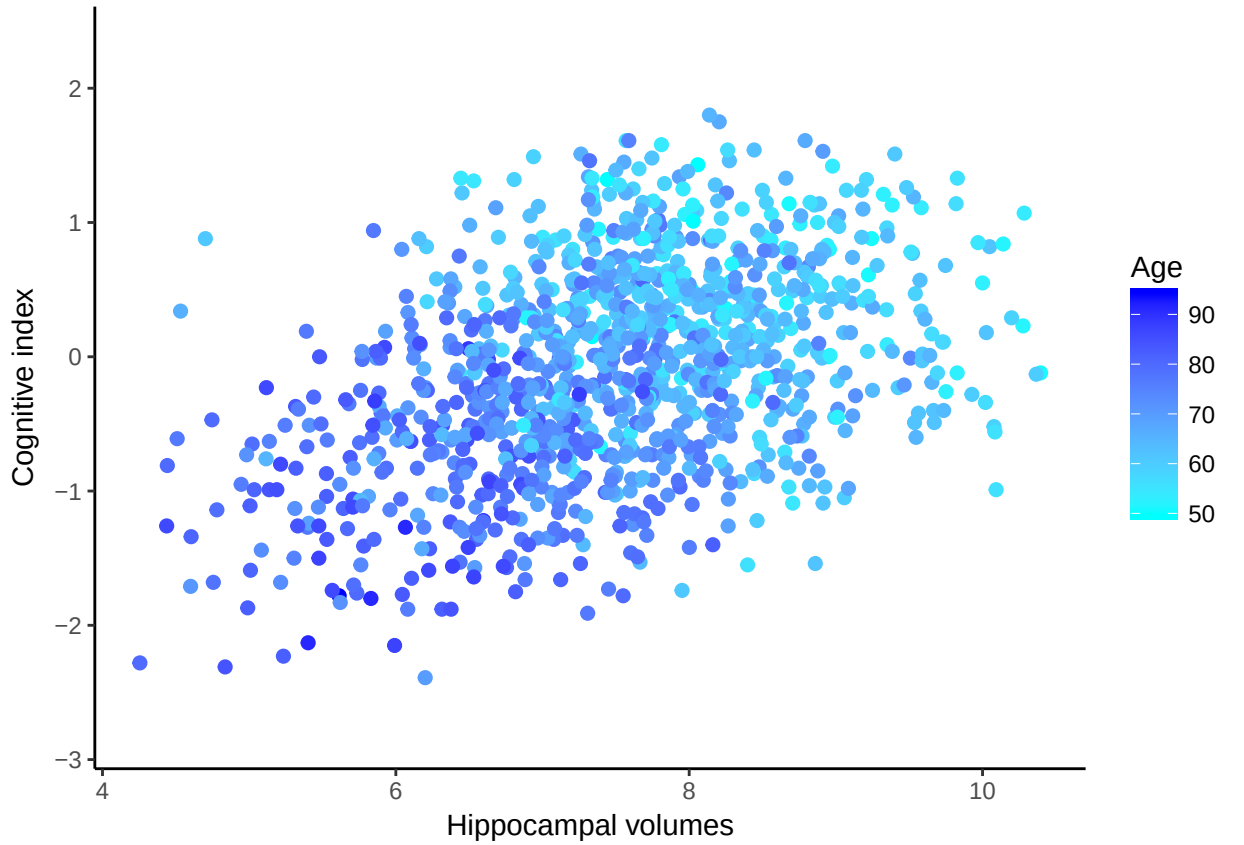
Our spaghetti plots tell us that the relationships between log-normalized white matter hyperintensities, hippocampal volumes, and the cognitive index are pretty linear over time:


Cognitive index
2
1
0
−1
−2
−3
−2.5
0.0
2.5
White matter hyperintensities



We know, however, that age is a critical determinant of all our IVs and our DV. Our scatterplot will plot the relationship between the IV and DV while colour coding according to age:


Cognitive index
2
1
0
−1
−2
−3
−2.5
0.0
2.5
White matter hyperintensities
Age
90
80
70
60
50



We can clearly see here an effect of age. In order to describe the complex relationships between age, white matter hyperintensities, hippocampal volumes, and cognitive function, we used linear mixed effects regression (LMER). This allowed us to describe fixed effects (i.e., WMH, HV, age) in tandem with subject-specific random effects.

Age, in the context of longitudinal studies, can be decomposed into two components: age at baseline and follow-up time (Morrell, Brant & Ferrucci, 2009). Age at baseline represents cross-sectional differences between subjects, while follow-up time represents longitudinal differences within individuals.

Before examining the effects of WMH and HV, let us first examine age-related cognitive decline using this framework. Two fixed effects were specified: one for baseline age, and another for time between visits. These two represent population effects. The coefficient for age represents the relationship between cognition and age, while the coefficient for time represents how cognition changes as time progresses. We also included two random effects: one for the intercept and one for the slope of time. This is because we expect individual differences in changes in cognition.

Here, the random intercept corresponds to individual variability in cognition at each timepoint, while the random slope corresponds to individual differences in the rate of change between timepoints. Because all subjects have a roughly equivalent time between assessments (5 years then 4 years), we do not expect much variability (measured by the variance and standard deviation of the random effects. That is precisely what we see:

```
library(lmerTest)
fit <- lmer(cognitiveindex ~ 1 + age06 + time + (1 + time|subject), data = df)
summary(fit)
```

```
## Linear mixed model fit by REML t-tests use Satterthwaite approximations
##   to degrees of freedom [lmerMod]
```

```
## Formula: cognitiveindex ~ 1 + age06 + time + (1 + time | subject)
##    Data: df
##
## REML criterion at convergence: 1738
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -3.10907 -0.40684  0.02916  0.42221  2.49370
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  subject  (Intercept) 0.320179 0.56584
##           time        0.001978 0.04447  -0.10
##  Residual             0.067896 0.26057
## Number of obs: 1240, groups:  subject, 503
##
## Fixed effects:
##               Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)  -0.041978   0.027608 485.200000   -1.52    0.129
## age06        -0.430413   0.026750 505.000000  -16.09   <2e-16 ***
## time         -0.048590   0.003107 372.100000  -15.64   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##       (Intr) age06
## age06  0.001
## time  -0.258  0.037
```

Both fixed effects are significant. Note the interpretation of the slope and intercept of the random effects of time. The variance is a measure of ... intercept significant, slope not (bc rate of change is constant)

for baseline age, another for time between visits, and a quadratic term for age. As previous research has shown a quadratic relationship between age and WMH (van Leijsen et al., 2017), we wanted to see if this also applied to age and cognition.

```
library(lmerTest)

fit <- lmer(cognitiveindex ~ age06 + time + agesq + (1+time|subject), data=df)
fit2 <- lmer(cognitiveindex ~ age06 + time + agesq + (time||subject), data=df)

anova(fit)
```

```
## Analysis of Variance Table of type III  with  Satterthwaite
## approximation for degrees of freedom
##        Sum Sq Mean Sq NumDF  DenDF F.value Pr(>F)
## age06 17.4703 17.4703     1 502.81 257.291 <2e-16 ***
## time  16.6060 16.6060     1 372.04 244.561 <2e-16 ***
## agesq  0.0073  0.0073     1 505.95   0.108  0.743
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
anova(fit, fit2)
```

```
## refitting model(s) with ML (instead of REML)
```

```
## Data: df
```

```
## Models:
## ..1: cognitiveindex ~ age06 + time + agesq + ((1 | subject) + (0 +
## ..1:     time | subject))
## object: cognitiveindex ~ age06 + time + agesq + (1 + time | subject)
##        Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## ..1     7 1732.5 1768.4 -859.25   1718.5
## object  8 1733.3 1774.3 -858.66   1717.3 1.1731      1     0.2788
```

```
#lmeTime1 <- lmer(cognitiveindex ~ age + time + timesq + (1 + time|rundmcs), data=df)

#lmeWMH <- lmer(cognitiveindex ~ logwmh + age + (1|rundmcs), data=df)
#lmeHV <- lmer(cognitiveindex ~ hv + age + (1|rundmcs), data=df)
#summary(lmeWMH)
#summary(lmeHV)
#summary(lmeTime1)
```

Does this mean WMH are synonymous with age-related cognitive decline? It's possible that the non-log transformed version is significant because it doesn't accurately describe the relationship