```
# Note to group: Emily loaded HSAUR2, not HSAUR as the Rpubs document state.
# They look the same
library(HSAUR2)

## Loading required package:  tools

library(ggplot2)
theme_set(theme_bw())
attach(BtheB)
```

Background/Context

The data we are using is from a clinical trial called "Beat the Blues". The Beat the Blues computer program was designed to deliver cognitive behavioural therapy to depressed patients. This dataset is from the R package "HSAUR2". The data contains 100 observations each relating to a study subject. The study measured 8 variables: drug, length, treatment, bdi.pre, bdi.2m, bdi.3m, bdi.5m and bdi.8m. The variable "drug" refers to if the subject took anti-depressants, "length" is the length of the subjects current episode of depression (either less than 6 months or more than six months), "treatment" is the treatment the subject was placed on (either treatment as usual (TAU), or Beat the Blues (BtheB)). The bdi variables refer to the Beck Depression Inventory at certain points during the study. The variables are titled 3m, 5m, and 8m to refer to 1, 3, and 6 months followup *relative to the 2-month post-treatment visit.*

Data Preview

BtheB Column Label	**Coding Manual**
drug	Did the patient take anti-depressant drugs (No or Yes)
length	Length of current episode (greater than or less than 6 months)
treatment	Type of treatment received
treatment (TAU)	Treatment as usual
treatment (BtheB)	Beat the Blues Cognitive Treatment
bdi.pre	Beck Depression Inventory II before treatment
bdi.2m	Beck Depression Inventory II after two months
bdi.4m	Beck Depression Inventory II after four months
bdi.6m	Beck Depression Inventory II after six months
bdi.8m	Beck Depression Inventory II after eight months

Intepreting bdi:

The Beck Depression Inventory (BDI, BDI-II), created by Dr. Aaron T. Beck, is a 21-question multiple-choice self-report inventory. Testing the patient's thoughts in relation to a level of depression. It signified a paradigm shift in the cognitive approach to depression and therapy.

bdi Scale	Indication
0 to 9	Minimal Depression
10 to 18	Mild Depression
19 to 29	Moderate Depression
30 to 63	Severe Depression

```
summary(BtheB[, c(1:4, 6, 8)])
```

```
##   drug    length   treatment     bdi.pre          bdi.3m          bdi.8m
##  No :56   <6m:49   TAU  :48   Min.   : 2.0   Min.   : 0.0   Min.   : 0.0
##  Yes:44   >6m:51   BtheB:52   1st Qu.:15.0   1st Qu.: 6.0   1st Qu.: 3.0
##                               Median :22.0   Median :13.0   Median :10.5
##                               Mean   :23.3   Mean   :14.8   Mean   :11.1
##                               3rd Qu.:30.2   3rd Qu.:20.0   3rd Qu.:15.2
##                               Max.   :49.0   Max.   :53.0   Max.   :40.0
##                                              NA's   :27     NA's   :48
```

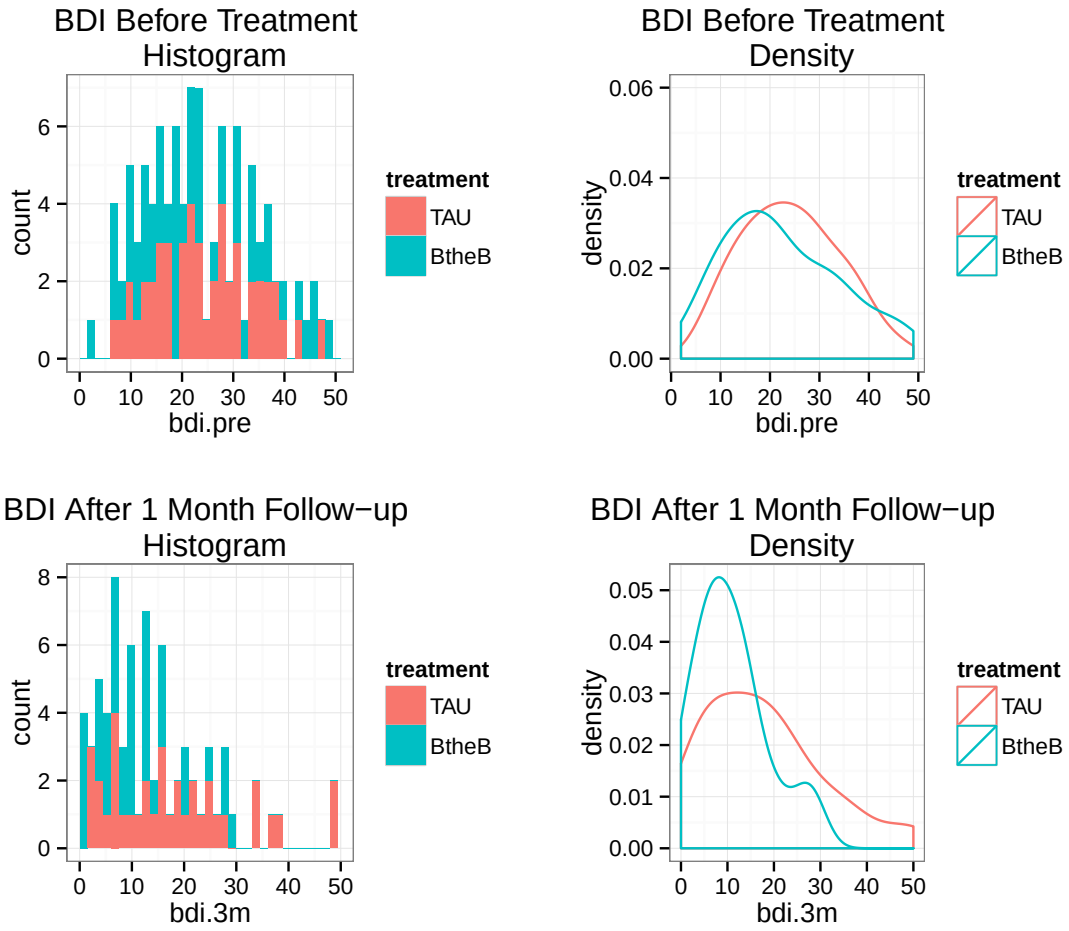
Variable Descriptions

The variables we are interested in are "bdi.pre" and "bdi.3m", Beck Depression Inventory II before treatment and after six months follow-up.

```
# Subsetting the data
BtheB.tau = BtheB[treatment == "TAU", ]
BtheB.btb = BtheB[treatment == "BtheB", ]
```

```
h1 = qplot(bdi.pre, fill = treatment, main = "BDI Before Treatment \n Histogram",
    binwidth = 1.5)
h2 = qplot(bdi.3m, fill = treatment, main = "BDI After 1 Month Follow-up \n Histogram",
    binwidth = 1.5, xlim = c(0, 50))

p1 = qplot(bdi.pre, geom = "density", color = treatment, main = "BDI Before Treatment \n De
```

```
    ylim = c(0, 0.06))
p2 = qplot(bdi.3m, geom = "density", color = treatment, main = "BDI After 1 Month Follow-up
    xlim = c(0, 50))

multiplot(h1, h2, p1, p2, cols = 2)

## Loading required package:  grid
```



Before treatment, both groups report BDIs that range between 0 and 50. The distribution is similar for both, as shown above. At this point we have not differentiated between those who are on drug treatment and those who are not.

At the one month followup date (or 3 months since the initial data were taken), we see shifts in the distribution of the BDI for both groups. The BtheB treatment group reported no BDI levels above about 35, while the TAU group reported BDI across the spectrum

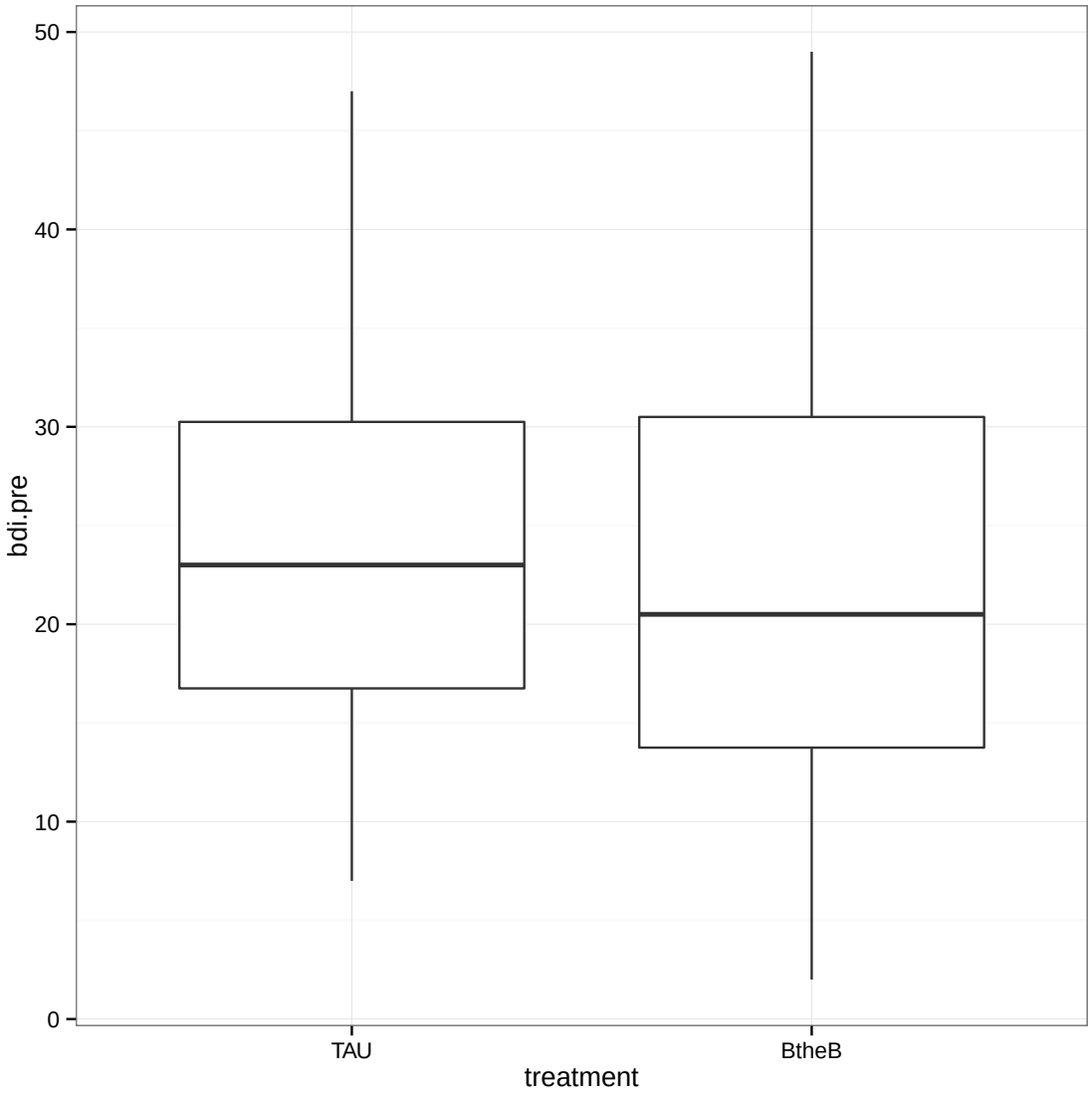
from 0-50. The density figures also show a shift to the left, or to lower BDI levels, for both treatment groups. The density graph in particular illustrates a shift in the mean BDI level from about 20-25 pre-treatment to about 8 3-months post-treatment. After 1 month follow-up, it seems more BtheB subjects score lower on the BDI scale than the TAU group (which means the subjects are less depressed). This suggests the BtheB treatment works better than treatment as usual.

Hypothesis

1. Patients receiving either treatment will experience a decline in depression symptoms as measured by the Beck Depression Inventory.
2. Patients receiving BtheB treatment will experience a greater decline in depression as measured by the Beck Depression Inventory compared to those receiving Treatment As Usual.

Missing Data

```
# JC

qplot(treatment, bdi.pre, geom = "boxplot", data = BtheB)
```

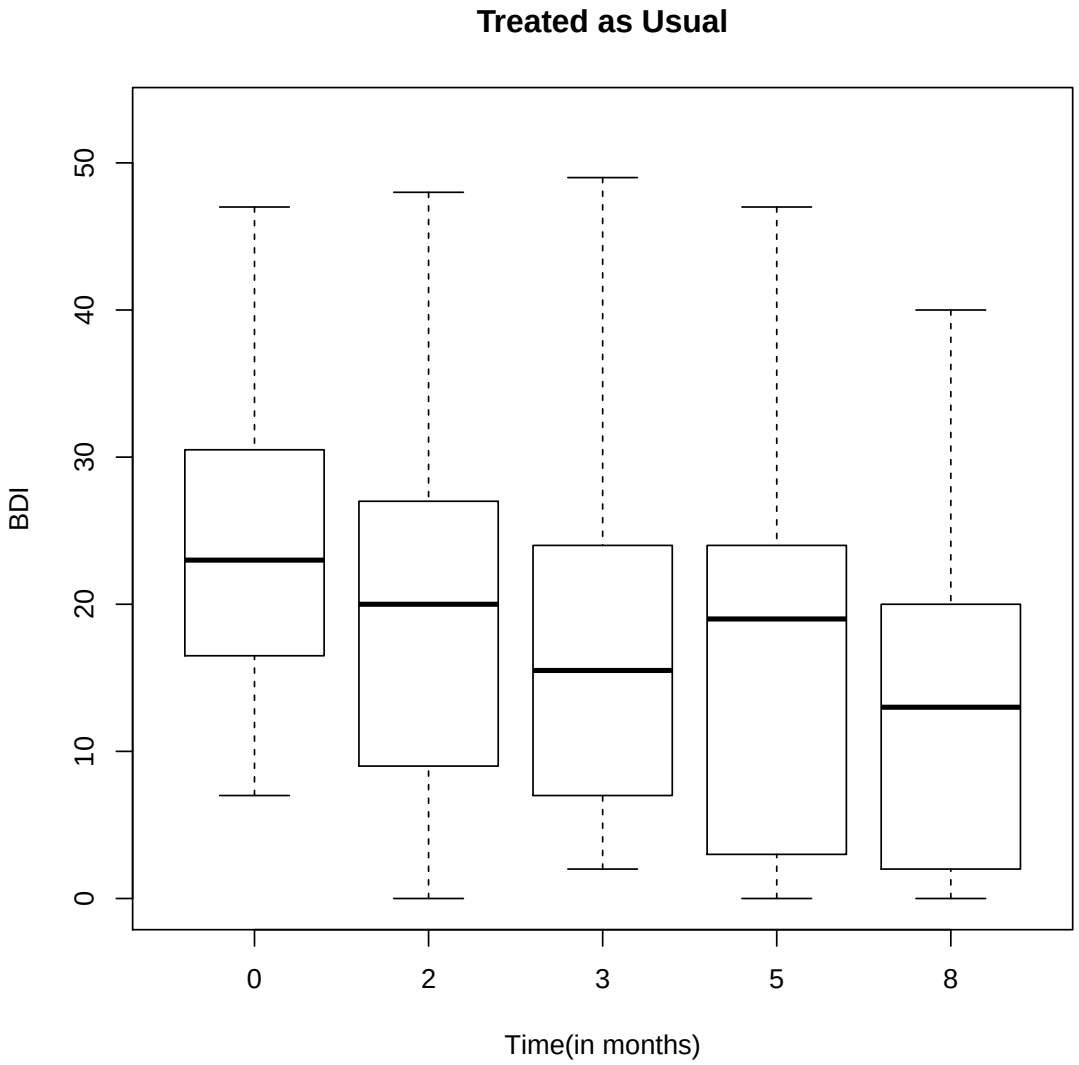



```
# http://people.umass.edu/nick/SADRIworkshop2012/session4_basicStats.html
# adopted from tutorial with() is a generic function that evaluates expr in
# a local environment constructed from data. The environment has the
# caller's environment as its parent. This is useful for simplifying calls
# to modeling functions.
(fm1 <- with(BtheB, lm(bdi.pre ~ treatment)))

##
```

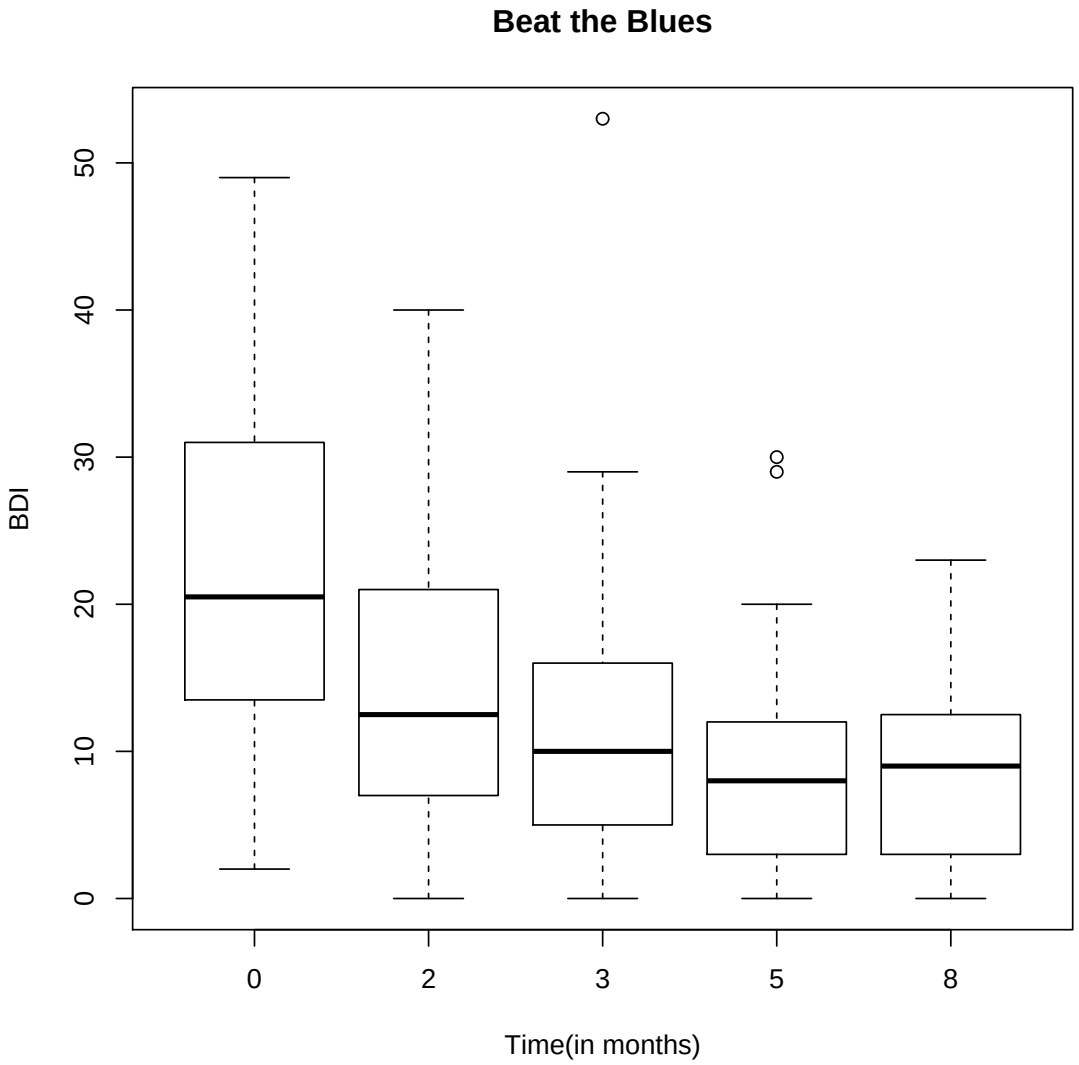
```
## Call:
## lm(formula = bdi.pre ~ treatment)
##
## Coefficients:
##    (Intercept)  treatmentBtheB
##          24.19           -1.65

#

# Missing data classification:

bdi.last <- rep(NA, nrow(BtheB))
trt.dur <- rep(NA, nrow(BtheB))
n.missing <- rep(NA, nrow(BtheB))
durs <- c(1, 2, 3, 5, 8)
for (i in 1:nrow(BtheB)) {
    first.na.idx <- max(which(!is.na(BtheB[i, 4:8])))
    n.missing[i] <- 5 - first.na.idx
    trt.dur[i] <- durs[first.na.idx]
    bdi.last[i] <- BtheB[i, 3 + first.na.idx]
}
bdi.change.rate <- (bdi.last - BtheB$bdi.pre)/trt.dur


fm2 <- with(BtheB, lm(bdi.change.rate ~ treatment))

fm3 <- with(BtheB, lm(bdi.change.rate ~ bdi.pre + treatment + drug + length))

fm4 <- with(BtheB, lm(bdi.change.rate ~ bdi.pre * treatment + drug + length))

library(visreg)

## Error:  there is no package called 'visreg'

# A function for visualizing regression models quickly and easily. Default
# plots contain a confidence band, prediction line, and partial residuals.
# Factors, transformations, conditioning, interactions, and a variety of
# other options are supported.
visreg(fm4, "bdi.pre", by = "treatment")

## Error:  could not find function "visreg"
```

```
# graph data with boxplot of each measures for each treatment group longform
# code An introduction to Applied Multivariate Analysis with R (Everitt,
# Hothorn)

# code contains missing values
BtheB$subject <- factor(rownames(BtheB))

# nrow and ncol return the number of rows or columns present in x. NCOL and
# NROW do the same treating a vector as 1-column matrix.
nobs <- nrow(BtheB)

# reshapes a data frame between wide format with repeated measurements in
# separate columns of the same record and long format with the repeated
# measurements in separate records.
BtheB_long <- reshape(BtheB, idvar = "subject", varying = c("bdi.2m", "bdi.3m",
    "bdi.5m", "bdi.8m"), direction = "long")

# rep replicates the values in x. It is a generic function, and the
# (internal) default method is described here.
BtheB_long$time <- rep(c(2, 3, 5, 8), rep(nobs, 4))

# lme() function removes missing values, does not remove participants with
# at least one missing value This generic function fits a linear
# mixed-effects model in the formulation described in Laird and Ware (1982)
# but allowing for nested random effects. The within-group errors are
# allowed to be correlated and/or have unequal variances.  might have to
# install this package install.packages(lme4)

# requires loading library(lme4) library(nlme)

library(nlme)
library(lme4)

## Error:  there is no package called 'lme4'

# So-called mixed-effect models (or just mixed
# models)includeadditionalrandom-effectterms,andare often appropriate for
# representing clustered, and therefore dependent, dataarising,for example,
# when data are gathered overtime on the same individuals

BtheB_lme1 <- lme(bdi ~ bdi.pre + time + treatment + drug + length, random = ~1 |
```

```
    subject, data = BtheB_long, na.action = na.omit)

BtheB_lme2 <- lme(bdi ~ bdi.pre + time + treatment + drug + length, random = ~time |
    subject, data = BtheB_long, na.action = na.omit)

# anova() Compute analysis of variance (or deviance) tables for one or more
# fitted model objects. These objects represent analysis-of-variance and
# analysis-of-deviance tables. When given a single argument it produces a
# table which tests whether the model terms are significant.When given a
# sequence of objects, anova tests the models against one another in the
# order specified.

anova(BtheB_lme1, BtheB_lme2)

##            Model df  AIC  BIC logLik   Test L.Ratio p-value
## BtheB_lme1     1  8 1883 1912 -933.5
## BtheB_lme2     2 10 1886 1922 -933.2 1 vs 2  0.5665  0.7533

ylim <- range(BtheB[, grep("bdi", names(BtheB))], na.rm = TRUE)
tau <- subset(BtheB, treatment == "TAU")[, grep("bdi", names(BtheB))]
boxplot(tau, main = "Treated as Usual", ylab = "BDI", xlab = "Time(in months)",
    names = c(0, 2, 3, 5, 8), ylim = ylim)
```




```
btheb <- subset(BtheB, treatment == "BtheB")[, grep("bdi", names(BtheB))]
boxplot(btheb, main = "Beat the Blues", ylab = "BDI", xlab = "Time(in months)",
    names = c(0, 2, 3, 5, 8), ylim = ylim)
```




```
# Additional information about the linear mixed-effects fit represented by
# object is extracted and included as components of object. The returned
# object is suitable for printing with the print.summary.lme method.

summary(BtheB_lme1)

## Linear mixed-effects model fit by REML
##  Data: BtheB_long
```

```
##    AIC  BIC logLik
##   1883 1912 -933.5
##
## Random effects:
##  Formula: ~1 | subject
##         (Intercept) Residual
## StdDev:       7.206    5.029
##
## Fixed effects: bdi ~ bdi.pre + time + treatment + drug + length
##                 Value Std.Error  DF t-value p-value
## (Intercept)     5.574    2.2995 182   2.424  0.0163
## bdi.pre         0.640    0.0799  92   8.013  0.0000
## time           -0.702    0.1469 182  -4.775  0.0000
## treatmentBtheB -2.315    1.7152  92  -1.350  0.1804
## drugYes        -2.816    1.7729  92  -1.588  0.1156
## length>6m       0.179    1.6816  92   0.106  0.9155
##  Correlation:
##                (Intr) bdi.pr time   trtmBB drugYs
## bdi.pre        -0.683
## time           -0.232  0.019
## treatmentBtheB -0.390  0.121  0.017
## drugYes        -0.074 -0.236 -0.022 -0.323
## length>6m      -0.244 -0.241 -0.036  0.002  0.157
##
## Standardized Within-Group Residuals:
##      Min       Q1      Med       Q3      Max
## -2.68699 -0.50847 -0.06085  0.42067  3.81414
##
## Number of Observations: 280
## Number of Groups: 97
```

```
summary(BtheB_lme2)
```

```
## Linear mixed-effects model fit by REML
##  Data: BtheB_long
##    AIC  BIC logLik
##   1886 1922 -933.2
##
## Random effects:
##  Formula: ~time | subject
##  Structure: General positive-definite, Log-Cholesky parametrization
```

```
##             StdDev Corr
## (Intercept) 7.3655 (Intr)
## time        0.4587 -0.191
## Residual    4.8985
##
## Fixed effects: bdi ~ bdi.pre + time + treatment + drug + length
##                 Value Std.Error  DF t-value p-value
## (Intercept)     5.599    2.3032 182   2.431  0.0160
## bdi.pre         0.643    0.0799  92   8.045  0.0000
## time           -0.699    0.1557 182  -4.490  0.0000
## treatmentBtheB -2.367    1.7163  92  -1.379  0.1713
## drugYes        -2.864    1.7739  92  -1.615  0.1098
## length>6m       0.119    1.6820  92   0.071  0.9436
##  Correlation:
##                (Intr) bdi.pr time   trtmBB drugYs
## bdi.pre        -0.682
## time           -0.237  0.019
## treatmentBtheB -0.389  0.121  0.017
## drugYes        -0.075 -0.235 -0.021 -0.323
## length>6m      -0.244 -0.239 -0.034  0.002  0.158
##
## Standardized Within-Group Residuals:
##     Min      Q1     Med      Q3     Max
## -2.2583 -0.4986 -0.0684  0.3867  3.6998
##
## Number of Observations: 280
## Number of Groups: 97

# simpler random intercept model
```

We chose these two variables to measure the association using a baseline measure and a midpoint measure. Using the 3 month mark (after 1 month follow-up), only 27 subjects having missing data. We originally wished to use the end of the study, however 48 subjects had missing data.

```
sum(is.na(bdi.pre) == TRUE)

## [1] 0

sum(is.na(bdi.3m) == TRUE)

## [1] 27
```

```
sum(is.na(bdi.8m) == TRUE)

## [1] 48
```

Analysis and Results

We plotted linear regressions for the BDI levels pre-treatment and three months/six months post-followup.

```
blm <- lm(bdi.3m ~ bdi.pre)
blm

##
## Call:
## lm(formula = bdi.3m ~ bdi.pre)
##
## Coefficients:
## (Intercept)      bdi.pre
##      0.0645       0.6369

blm1 <- lm(bdi.3m ~ bdi.pre, data = BtheB.tau)
blm1

##
## Call:
## lm(formula = bdi.3m ~ bdi.pre, data = BtheB.tau)
##
## Coefficients:
## (Intercept)      bdi.pre
##      -1.814        0.823

blm2 <- lm(bdi.3m ~ bdi.pre, data = BtheB.btb)
blm2

##
## Call:
## lm(formula = bdi.3m ~ bdi.pre, data = BtheB.btb)
##
## Coefficients:
## (Intercept)      bdi.pre
##       0.995        0.487
```

```
qplot(BtheB$bdi.pre, BtheB$bdi.3m, data = BtheB, color = treatment) + geom_abline(intercept
    slope = 0.8231, colour = "red") + geom_abline(intercept = 0.9949, slope = 0.4871,
    colour = "blue") + geom_abline(intercept = 0.06449, slope = 0.63686, colour = "black")

## Warning:  Removed 27 rows containing missing values (geom_point).
```