Analysis of the paper "Response of digestive organs of Hypsiboas albopunctatus to benzo[alpha]pyrene"

Lara Z. Fanali, Bruno Valverde, Lilian Franco-Belussi, Diogo B. Provete, Classius de Oliveira

2017-02-18

Contents

R version

```
R.Version()$version.string; R.Version()$platform
```

```
## [1] "R version 3.3.2 (2016-10-31)"
```

```
## [1] "x86_64-apple-darwin13.4.0"
```

Loading packages

```
library(plyr)
packageVersion("plyr")
```

```
## [1] '1.8.4'
```

```
library(tables)
packageVersion("tables")
```

```
## [1] '0.8'
```

```
library(lme4)
packageVersion("lme4")
```

```
## [1] '1.1.12'
```

```
library(sjPlot)
packageVersion("sjPlot")
```

```
## [1] '2.2.1'
```

```
library(lsmeans)
packageVersion("lsmeans")
```

```
## [1] '2.25'
```

```
library(sciplot)
packageVersion("sciplot")
```

```
## [1] '1.1.0'
```

```
library(gridExtra)
packageVersion("gridExtra")
```

```
## [1] '2.2.1'
```

```
library(Rmisc)
packageVersion("Rmisc")
```

```
## [1] '1.5'
```

Data input and handling

Importing and visualizing the data

```
liver<-read.table("liver_data.txt", h=TRUE)
head(liver);str(liver)
```

```
##   animal photo time treatment     area
## 1      1     1   3d     1cont 4576.531
## 2      1     2   3d     1cont 2686.123
## 3      1     3   3d     1cont 1998.655
## 4      1     4   3d     1cont 5063.698
## 5      1     5   3d     1cont 2661.902
## 6      1     6   3d     1cont 3197.442
```

```
## 'data.frame':    900 obs. of  5 variables:
##  $ animal   : int  1 1 1 1 1 1 1 1 1 1 ...
##  $ photo    : int  1 2 3 4 5 6 7 8 9 10 ...
##  $ time     : Factor w/ 2 levels "3d","7d": 1 1 1 1 1 1 1 1 1 1 ...
##  $ treatment: Factor w/ 3 levels "1cont","3mg",..: 1 1 1 1 1 1 1 1 1 1 ...
##  $ area     : num  4577 2686 1999 5064 2662 ...
```

```
intestine<-read.table("intestine_data.txt", h=TRUE)
head(intestine);str(intestine)
```

```
##   animal photo time treatment thickness
## 1      1     1   3d     1cont  34.15129
## 2      1     2   3d     1cont  34.74980
## 3      1     3   3d     1cont  35.77058
## 4      1     4   3d     1cont  35.55163
## 5      1     5   3d     1cont  35.70421
## 6      1     6   3d     1cont  34.86159
```

```
## 'data.frame':    750 obs. of  5 variables:
##  $ animal   : int  1 1 1 1 1 1 1 1 1 1 ...
##  $ photo    : int  1 2 3 4 5 6 7 8 9 10 ...
##  $ time     : Factor w/ 2 levels "3d","7d": 1 1 1 1 1 1 1 1 1 1 ...
##  $ treatment: Factor w/ 3 levels "1cont","3mg",..: 1 1 1 1 1 1 1 1 1 1 ...
##  $ thickness: num  34.2 34.7 35.8 35.6 35.7 ...
```

Converting `animal` and `photo` to factor to allow nesting in the model:

```
liver$animal<-factor(liver$animal)
liver$photo<-factor(liver$photo)
intestine$animal<-factor(intestine$animal)
intestine$foto<-factor(intestine$photo)
```

The tables below show the number of pictures (25) in each treatment (rows) in each of the five or six animals (columns), for the two exposure times:

For intestine thickness

```
tabular(treatment~time*animal, data=intestine)
```

```
##
##            time
##            3d                7d
##            animal            animal
##  treatment 1      2  3  4  5  1      2  3  4  5
##  1cont     25     25 25 25 25 25     25 25 25 25
##  3mg       25     25 25 25 25 25     25 25 25 25
##  7mg       25     25 25 25 25 25     25 25 25 25
```

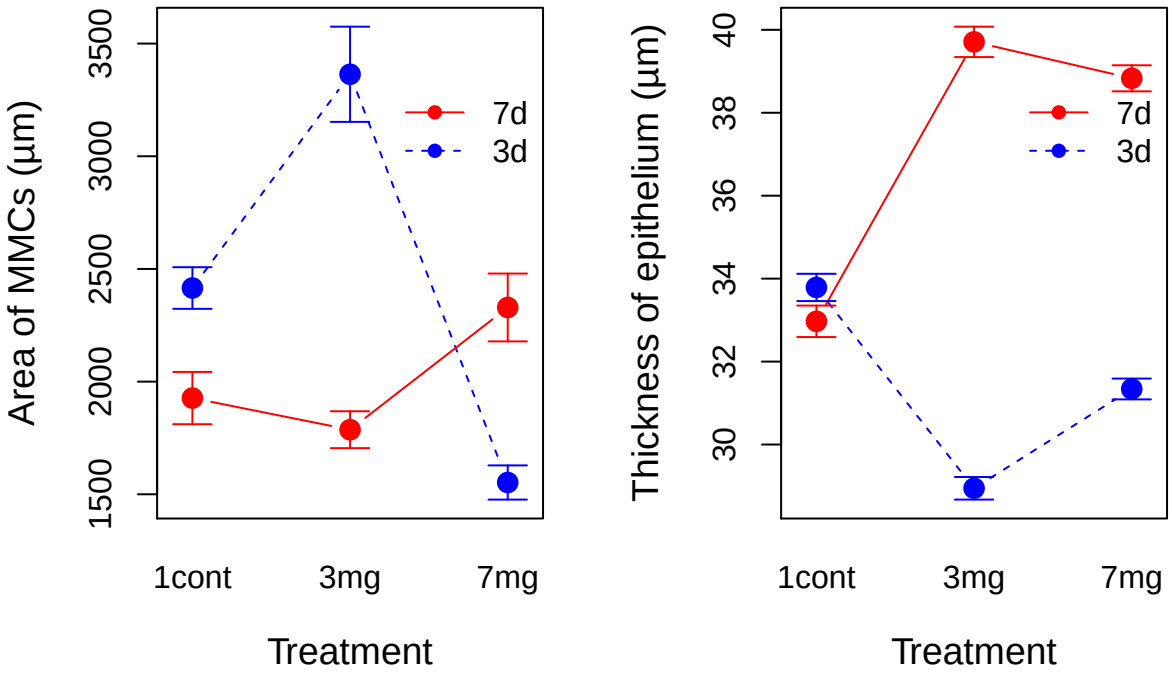
For melanomacrophage area in the liver

```
tabular(treatment~time*animal, data=liver)
```

```
##
##            time
##            3d                   7d
##            animal               animal
##  treatment 1      2  3  4  5  6  1      2  3  4  5  6
##  1cont     25     25 25 25 25 25 25     25 25 25 25 25
##  3mg       25     25 25 25 25 25 25     25 25 25 25 25
##  7mg       25     25 25 25 25 25 25     25 25 25 25 25
```

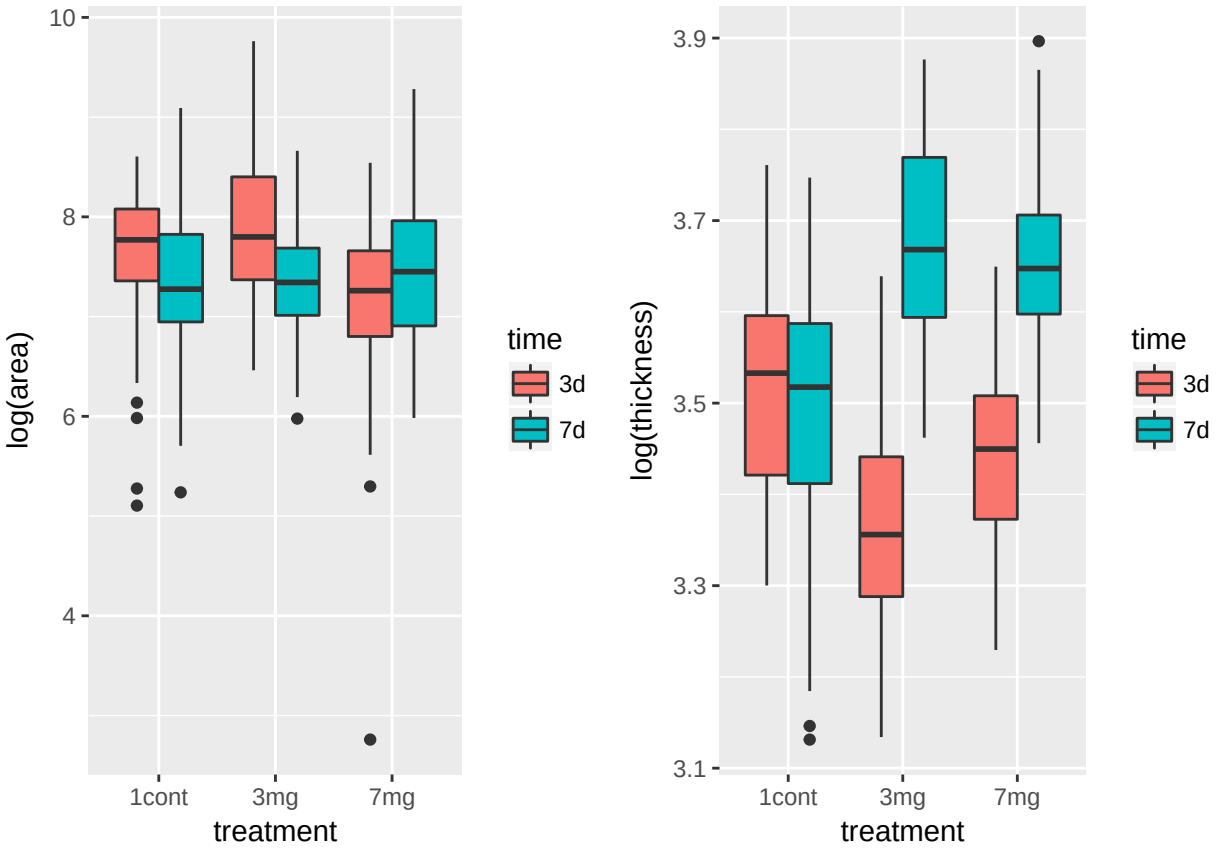
Data exploration

```
par(mfrow=c(1,2))
lineplot.CI(treatment,area, group = time, data = liver, cex = 1.5, xlab = "Treatment", ylab = "Area of

lineplot.CI(treatment,thickness, group = time, data = intestine, cex = 1.5, xlab = "Treatment", ylab =
```



These plots show that there's interaction in the two response variables, because the two lines representing the factors cross at some point. Thus, we need to include that in the models.

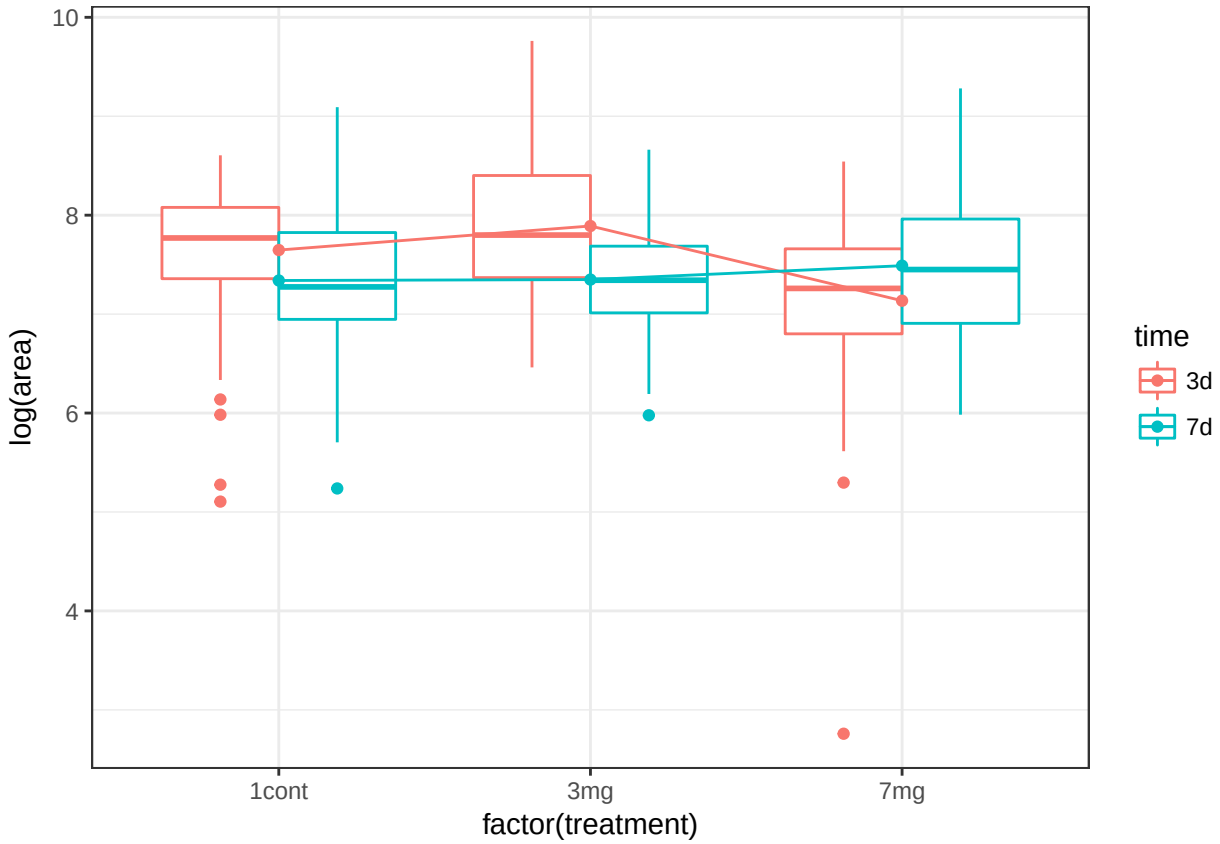
```
p<-ggplot(liver, aes(treatment,log(area))) +
    geom_boxplot(aes(fill = time))
q<-ggplot(intestine, aes(treatment,log(thickness))) +
    geom_boxplot(aes(fill = time))
grid.arrange(p,q, ncol=2)
```



Other types of graphs:

1. Boxplot showing interaction

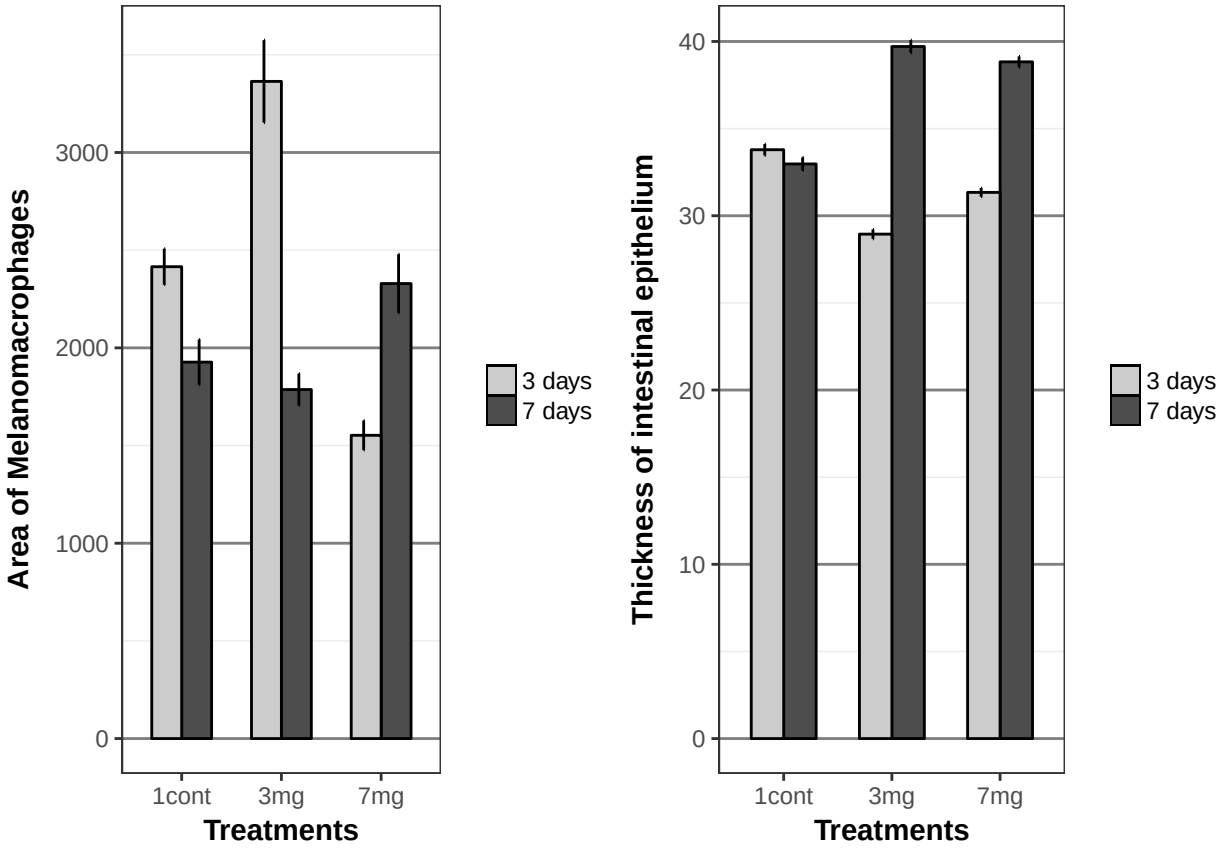
```
toothInt <- ddply(liver,.(treatment,time),summarise, val = mean(log(area)))

ggplot(liver, aes(x = factor(treatment), y = log(area), colour = time)) +
    geom_boxplot() +
    geom_point(data = toothInt, aes(y = val)) +
    geom_line(data = toothInt, aes(y = val, group = time)) +
    theme_bw()
```



2. Bar chart with the two factors:

```
sum = summarySE(liver, measurevar="area", groupvars=c("time","treatment"))
p1 <- ggplot(sum, aes(x = treatment, y = area,
                          fill = time, width = 0.6,ymax=area+se, ymin=area-se)) +
  geom_bar(stat = "identity", position = "dodge", width = 0.7) +
  geom_bar(stat = "identity", position = "dodge", colour = "black", width = 0.7,  show.legend = FALSE) +
  scale_fill_manual(name = "Time" ,
                values = c('grey80', 'grey30'),
                labels = c("3 days",
                           "7 days")) +
    geom_errorbar(position=position_dodge(width=0.7),
                      width=0.0, size=0.5, color="black") +
  labs(x = "Treatments", y = "Area of Melanomacrophages") +
  theme_bw() +
    theme(panel.grid.major.x = element_blank(),
              panel.grid.major.y = element_line(colour = "grey50"),
              plot.title = element_text(size = rel(1.5),
              face = "bold", vjust = 1.5),
              axis.title = element_text(face = "bold"),
              legend.position = "right",
              legend.title = element_blank(),
              legend.key.size = unit(0.4, "cm"),
              legend.key = element_rect(fill = "black"),
              axis.title.y = element_text(vjust= 1.8),
              axis.title.x = element_text(vjust= -0.5)
            )
```

```
sum1 = summarySE(intestine, measurevar="thickness", groupvars=c("time","treatment"))
q1 <- ggplot(sum1, aes(x = treatment, y = thickness,
                           fill = time, width = 0.6,ymax=thickness+se, ymin=thickness-se)) +
  geom_bar(stat = "identity", position = "dodge", width = 0.7) +
  geom_bar(stat = "identity", position = "dodge", colour = "black", width = 0.7,  show.legend = FALSE) +
  scale_fill_manual(name = "Time" ,
                values = c('grey80', 'grey30'),
                labels = c("3 days",
                           "7 days")) +
    geom_errorbar(position=position_dodge(width=0.7),
                      width=0.0, size=0.5, color="black") +
  labs(x = "Treatments", y = "Thickness of intestinal epithelium") +
  theme_bw() +
    theme(panel.grid.major.x = element_blank(),
              panel.grid.major.y = element_line(colour = "grey50"),
              plot.title = element_text(size = rel(1.5),
              face = "bold", vjust = 1.5),
              axis.title = element_text(face = "bold"),
              legend.position = "right",
              legend.title = element_blank(),
              legend.key.size = unit(0.4, "cm"),
              legend.key = element_rect(fill = "black"),
              axis.title.y = element_text(vjust= 1.8),
              axis.title.x = element_text(vjust= -0.5)
            )
grid.arrange(p1,q1,ncol=2)
```



Modelling

Now, we want to build one linear model for each response variable: `area` and `thickness`, with the two factors: 1) `Time` - with two levels, and 2) `Treatment` with three levels. Our experiment has a 3x2 factorial design, fully crossed. We have multiple measurements (pseudoreplicates) per animal (true replicate), i.e., clustered data. Our sampling units are pictures of histological sections, which are nested within animals. The concentration of the contaminant was applied to the animal. To model the effects of time and treatment, it's easier to build a simple linear mixed-effects model in which we include a random intercept to `animal`, since we're not interested in differences among animals:

Model for intestinal epithelium thickness

```
#see page 703 of Crawley (2012) the R book
mod1<-lmer(thickness~treatment*time+(1|animal), data = intestine)
summary(mod1)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: thickness ~ treatment * time + (1 | animal)
##    Data: intestine
##
## REML criterion at convergence: 4019
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.6195 -0.7499  0.0522  0.7510  2.6809
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  animal   (Intercept)  0.7756  0.8807
##  Residual             12.3346  3.5121
## Number of obs: 750, groups:  animal, 5
##
## Fixed effects:
##                      Estimate Std. Error t value
## (Intercept)          33.7878     0.5038   67.07
## treatment3mg         -4.8434     0.4442  -10.90
## treatment7mg         -2.4484     0.4442   -5.51
## time7d               -0.8170     0.4442   -1.84
## treatment3mg:time7d  11.5815     0.6283   18.43
## treatment7mg:time7d   8.3080     0.6283   13.22
##
## Correlation of Fixed Effects:
##             (Intr) trtmn3 trtmn7 time7d trt3:7
## treatmnt3mg -0.441
## treatmnt7mg -0.441  0.500
## time7d      -0.441  0.500  0.500
## trtmnt3mg:7  0.312 -0.707 -0.354 -0.707
## trtmnt7mg:7  0.312 -0.354 -0.707 -0.707  0.500
```

```
MuMIn::r.squaredGLMM(mod1)#marginal and conditional R^2
```

```
##       R2m       R2c
## 0.5318621 0.5595573
```

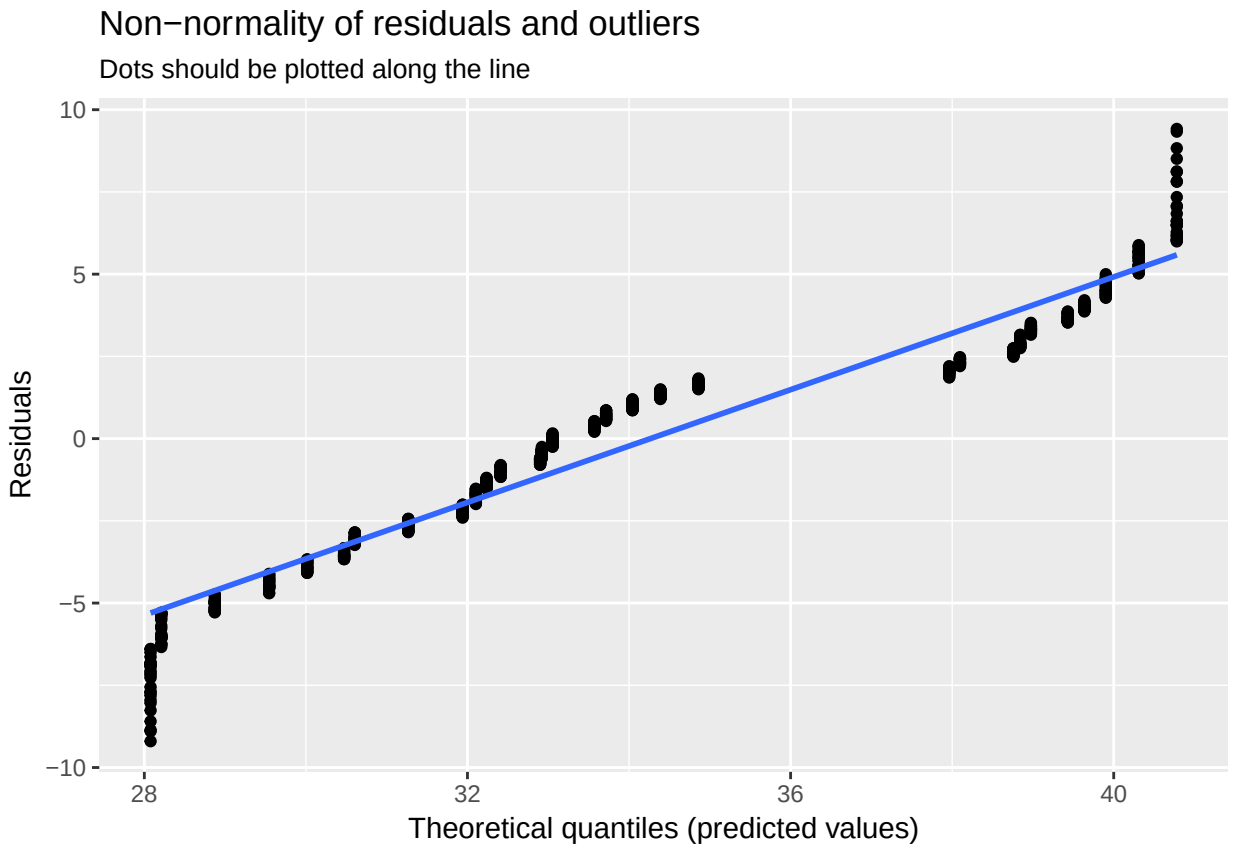
Effective sample size

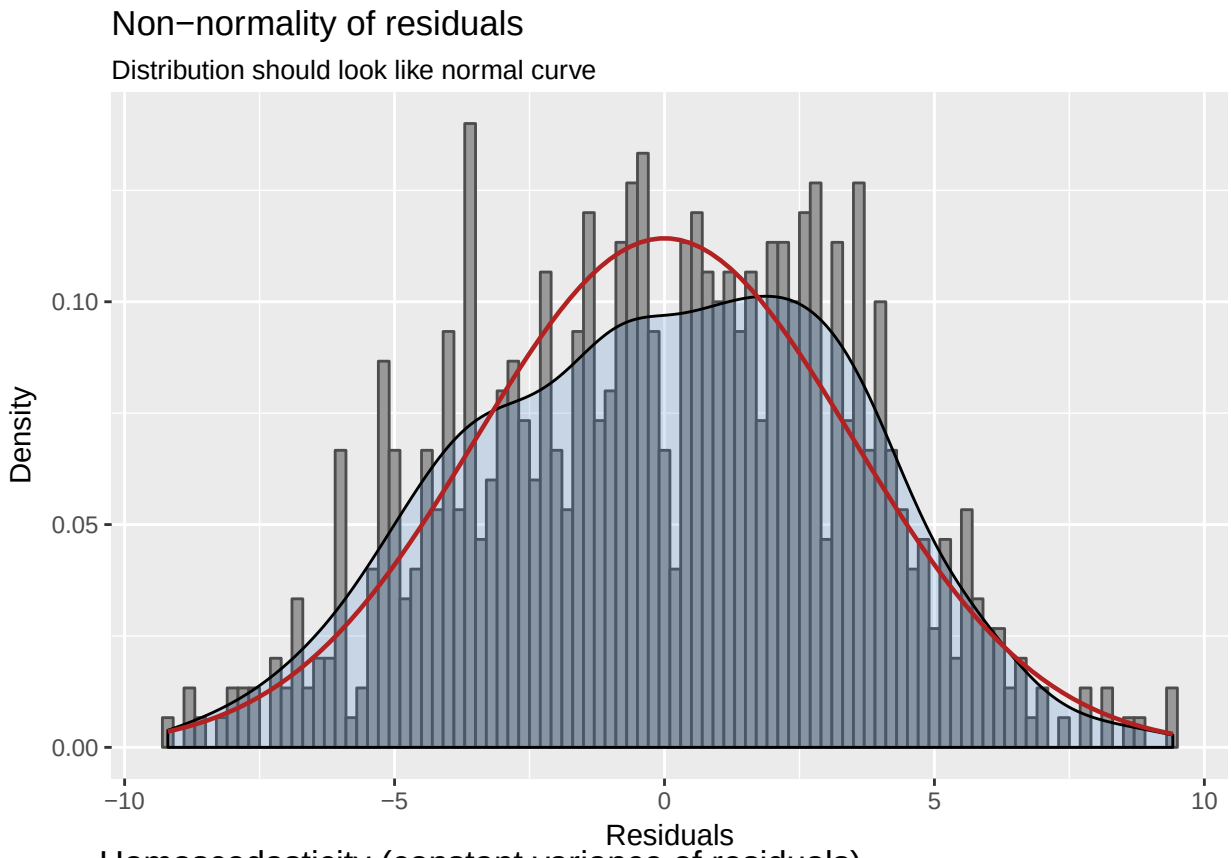
```
25*5/((24*0.059)+1)
```

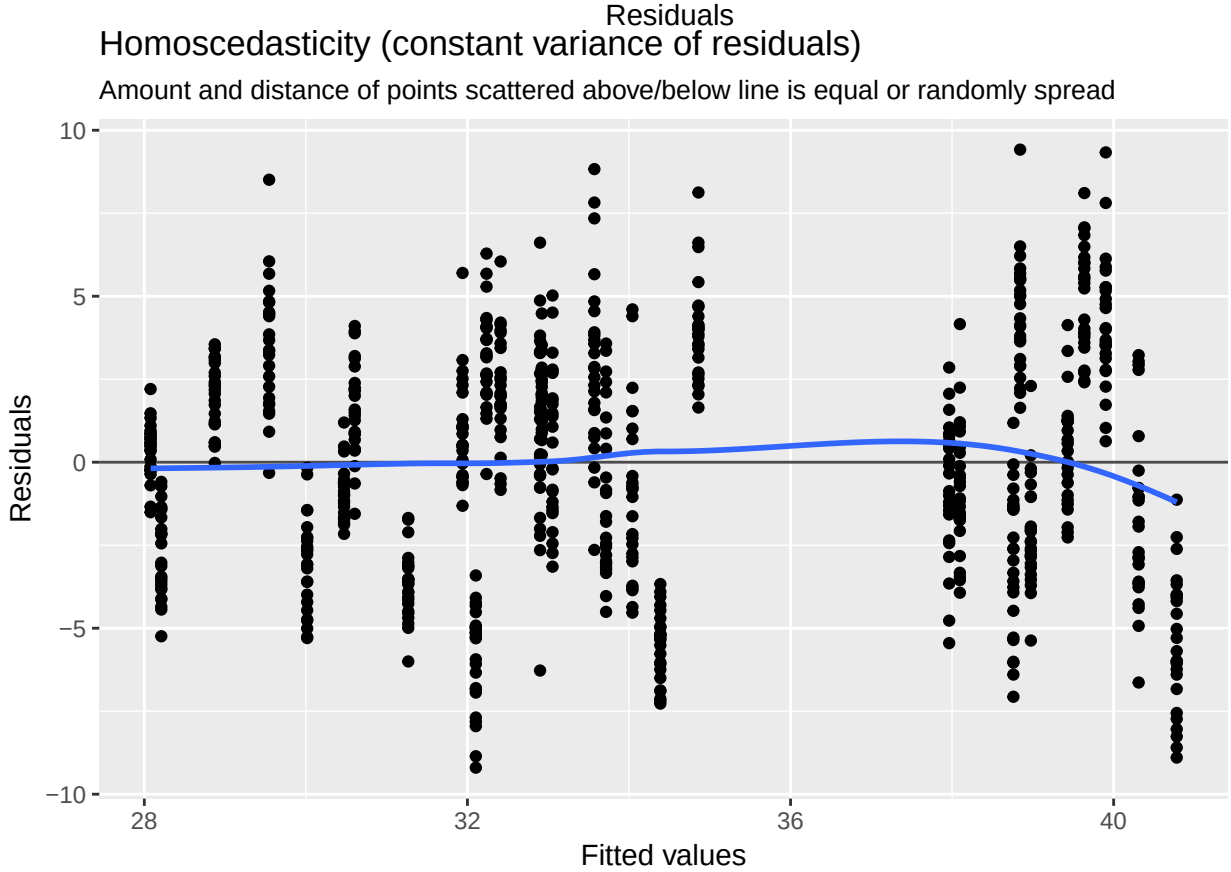
```
## [1] 51.73841
```

Testing model assumptions

```
sjp.lmer(mod1, type = "ma")
```




Non−normality of residuals
Distribution should look like normal curve
Density
0.10
0.05
0.00
−10
−5
0
5
10
Residuals


Homoscedasticity (constant variance of residuals)
Amount and distance of points scattered above/below line is equal or randomly spread
Residuals
10
5
0
−5
−10
28
32
36
40
Fitted values

Residuals are normally distributed and we don't have problems with heteroceidasticity either.

Making contrast to test the effects of treatment and time on thickness of intestinal epithelium:

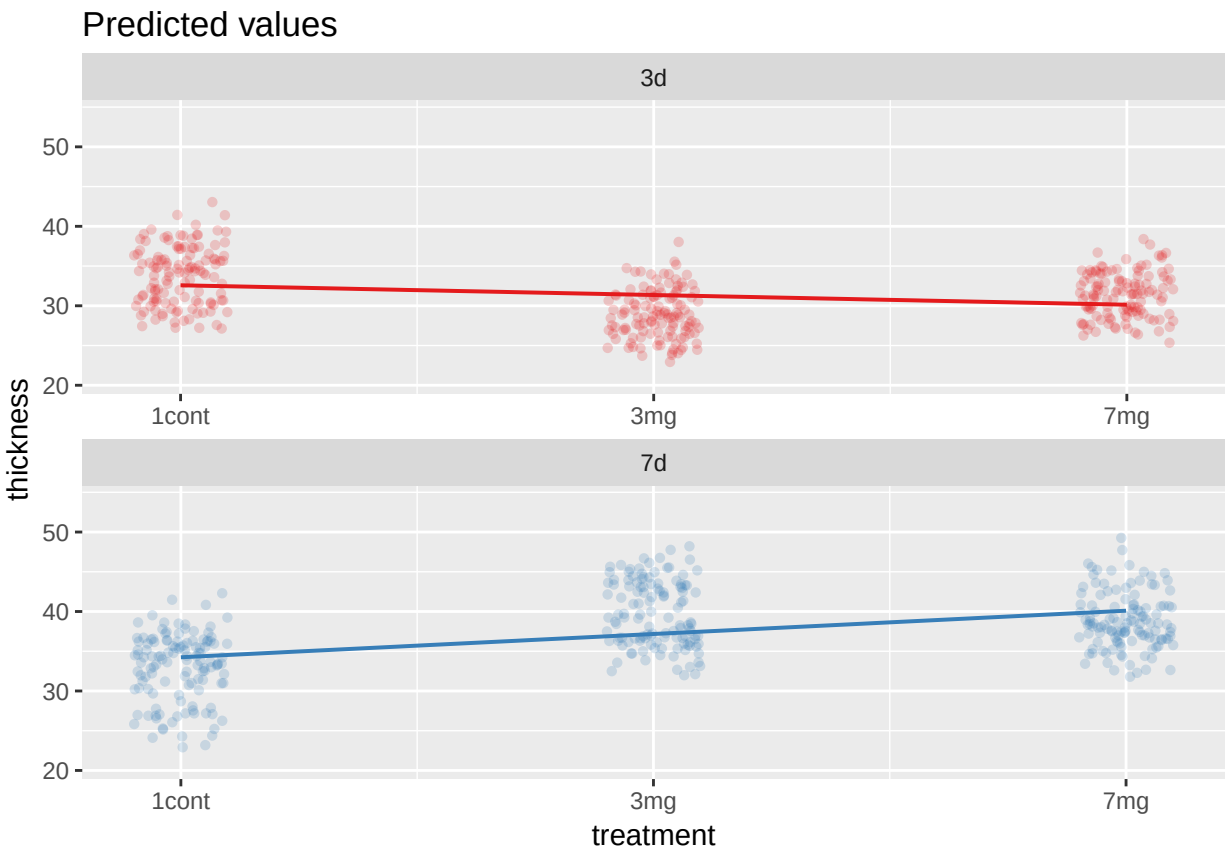
```
pairs(lsmeans(mod1,  ~ treatment|time))#contrasts between levels of factors
```

```
## Loading required namespace: lmerTest
```

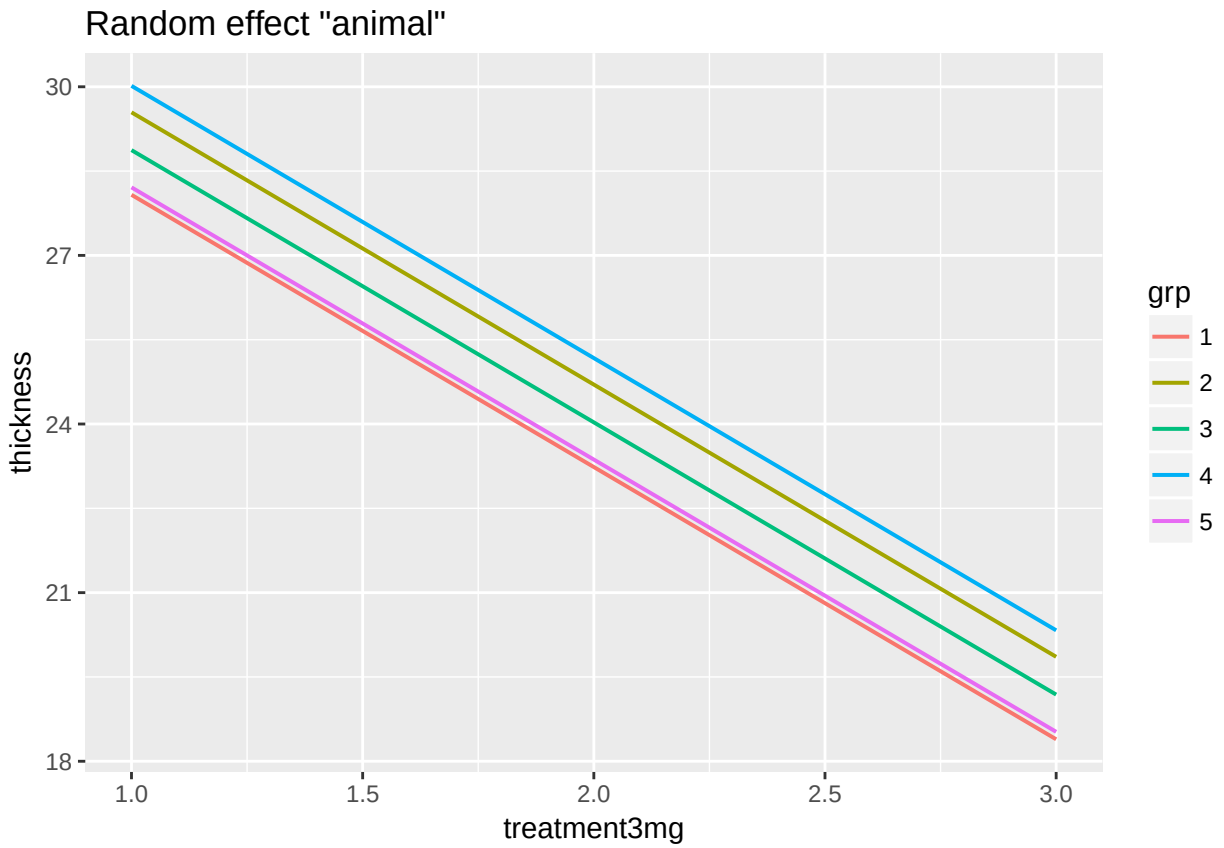
```
## time = 3d:
##  contrast     estimate        SE  df t.ratio p.value
##  1cont - 3mg  4.843370 0.4442454 740  10.902  <.0001
##  1cont - 7mg  2.448382 0.4442454 740   5.511  <.0001
##  3mg - 7mg   -2.394988 0.4442454 740  -5.391  <.0001
##
## time = 7d:
##  contrast     estimate        SE  df t.ratio p.value
##  1cont - 3mg -6.738147 0.4442454 740 -15.168  <.0001
##  1cont - 7mg -5.859582 0.4442454 740 -13.190  <.0001
##  3mg - 7mg    0.878565 0.4442454 740   1.978  0.1184
##
## P value adjustment: tukey method for comparing a family of 3 estimates
```

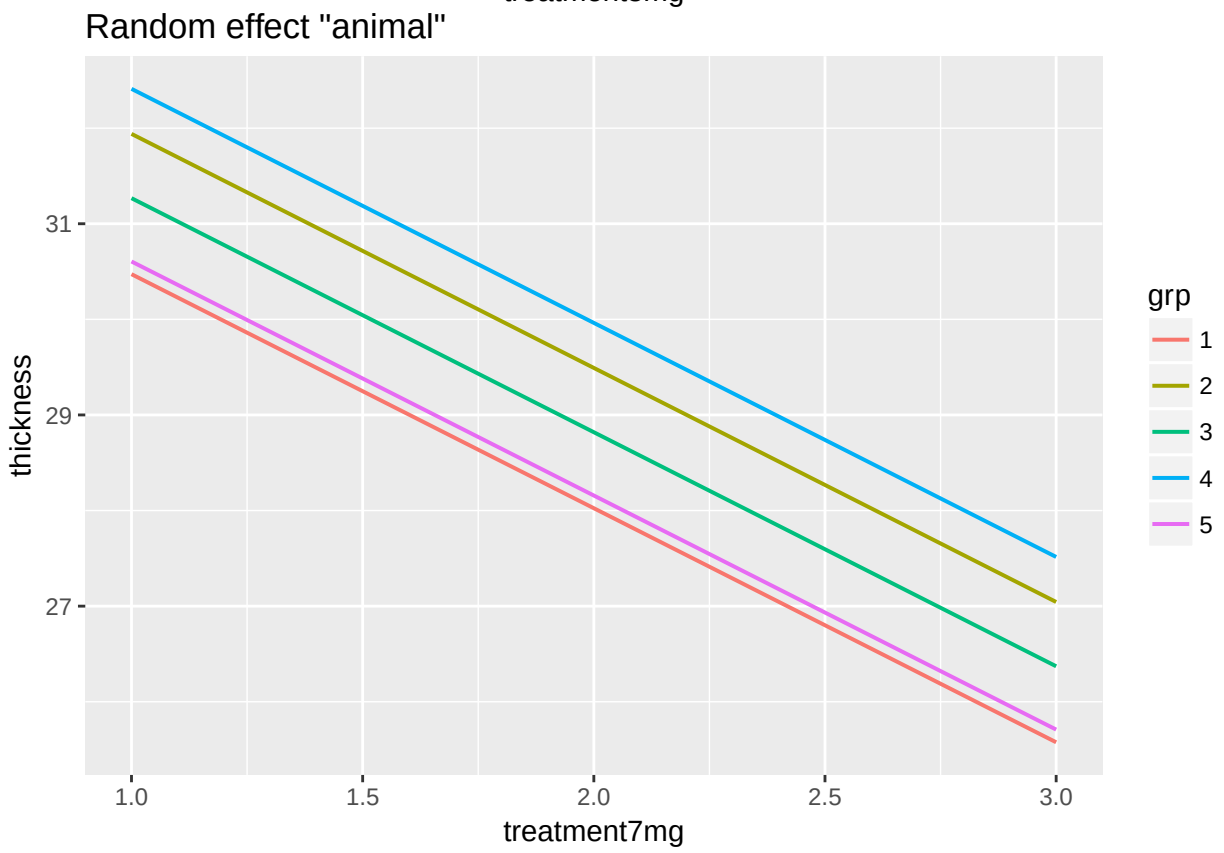
There's difference between all levels of the two factors, but between 3mg-7mg.

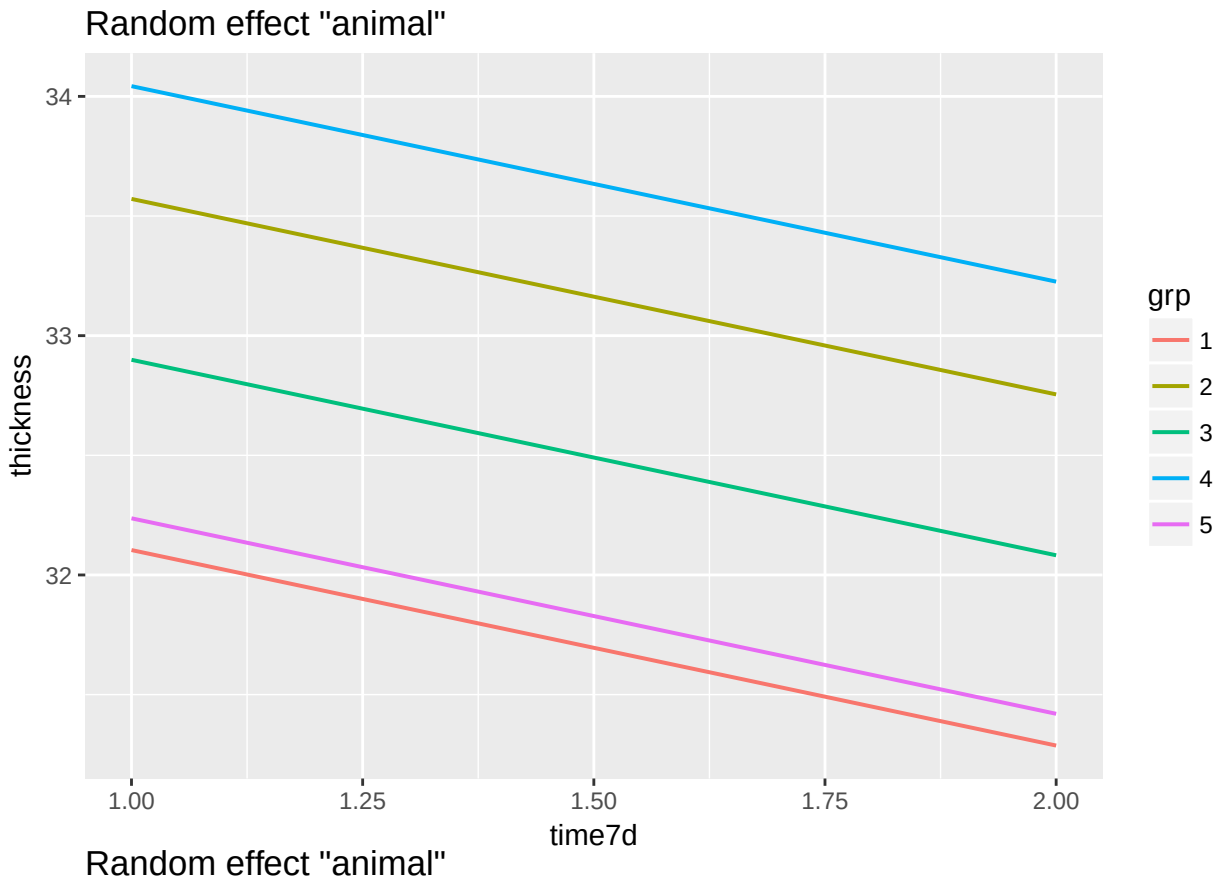
```
sjp.lmer(mod1, type = "pred", vars = c("treatment", "time"))
```

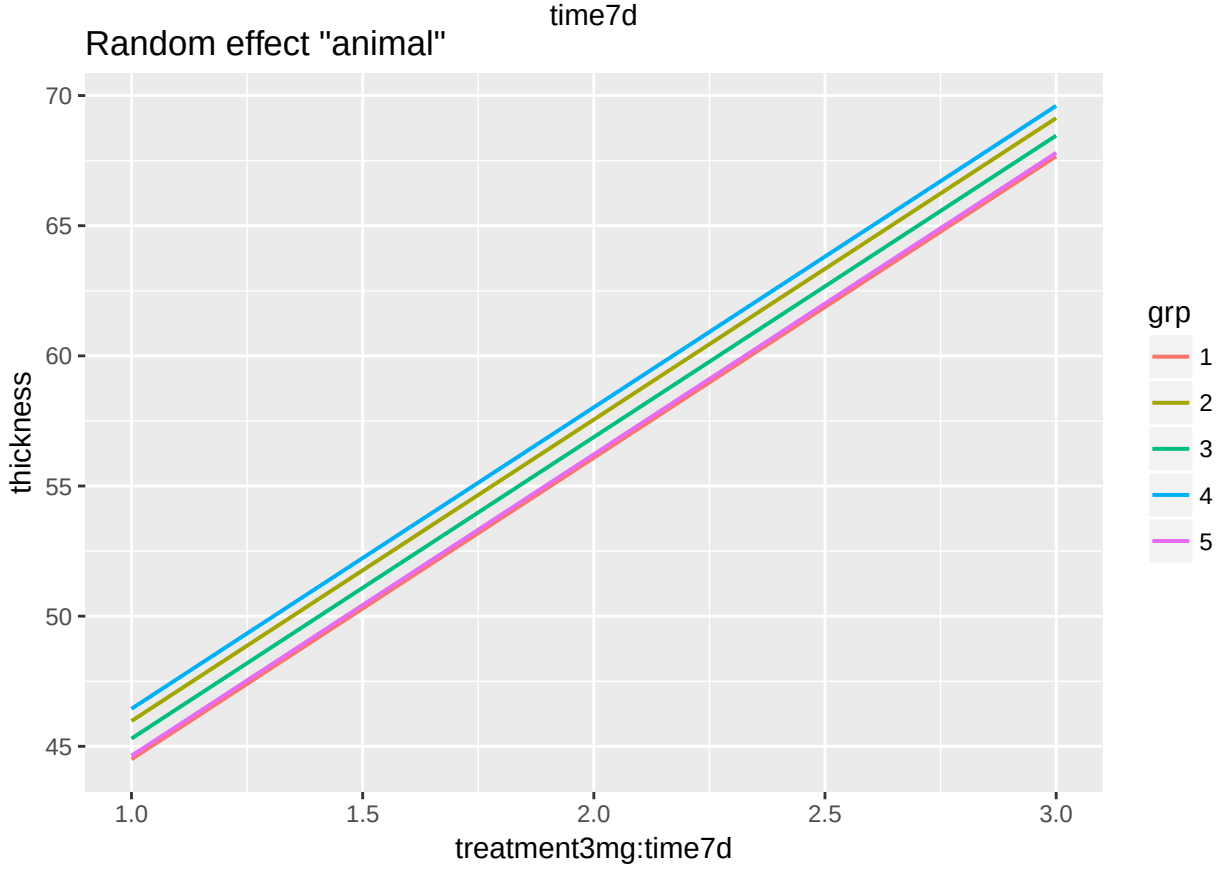


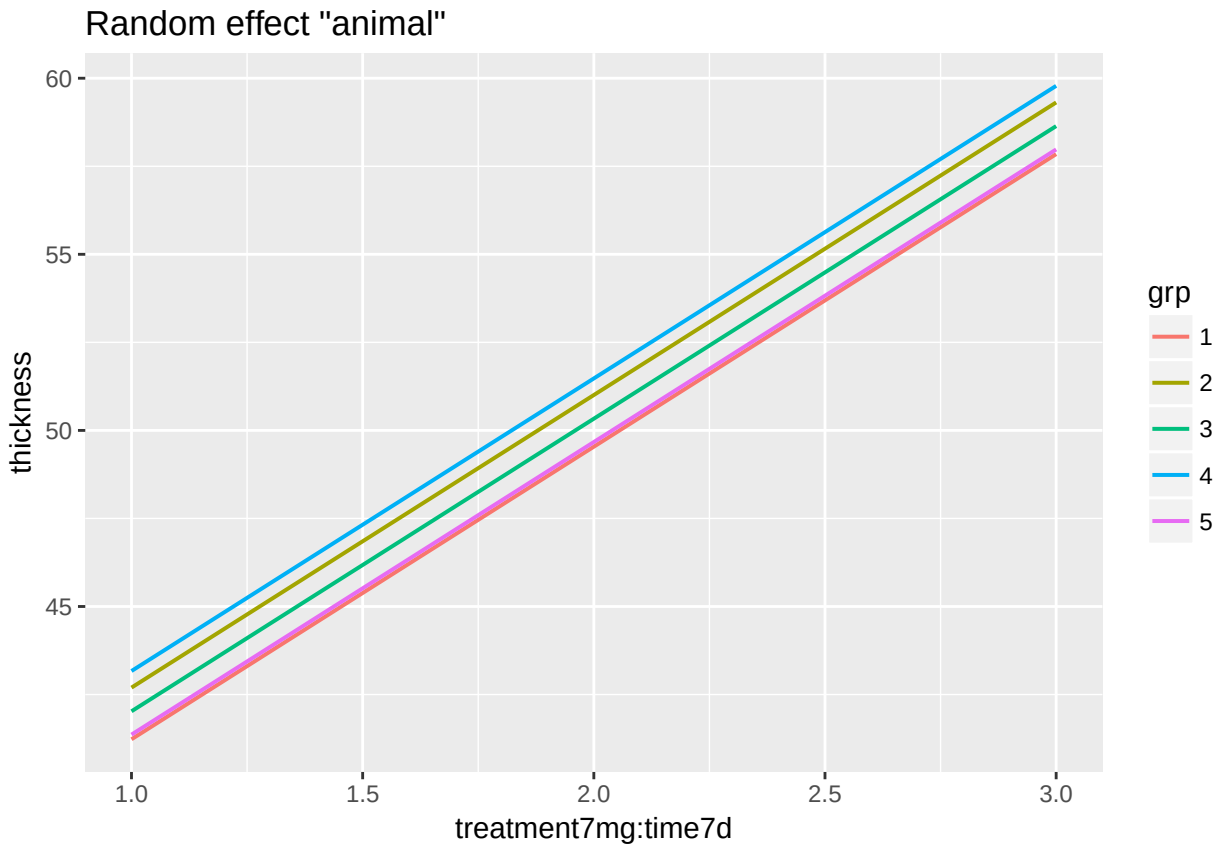

```
sjp.lmer(mod1, type = "ri.slope")
```


Random effect "animal"
30
27
24
21
18
thickness
1.0
1.5
2.0
2.5
3.0
treatment3mg
grp
1
2
3
4
5


Random effect "animal"
31
29
27
thickness
1.0
1.5
2.0
2.5
3.0
treatment7mg
grp
1
2
3
4
5


Random effect "animal"
thickness
34
33
32
1.00
1.25
1.50
1.75
2.00
time7d
grp
1
2
3
4
5


Random effect "animal"
thickness
70
65
60
55
50
45
1.0
1.5
2.0
2.5
3.0
treatment3mg:time7d
grp
1
2
3
4
5



Model for Melanomacrophage area in the liver

```
range(liver$area)#very wide variation, to reduce it we will log-transform the data
```

```
## [1]    15.78076 17350.14750
```

```
range(log(liver$area))
```

```
## [1] 2.758791 9.761356
```

```
mod2<-lmer(log(area)~treatment*time+(1|animal), data = liver)
summary(mod2)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: log(area) ~ treatment * time + (1 | animal)
##    Data: liver
##
## REML criterion at convergence: 1636.1
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -6.7232 -0.6019  0.0206  0.6720  3.4067
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  animal   (Intercept) 0.1099   0.3314
##  Residual             0.3454   0.5877
```

```
## Number of obs: 900, groups:  animal, 6
##
## Fixed effects:
##                      Estimate Std. Error t value
## (Intercept)           7.64745    0.14357   53.27
## treatment3mg          0.24305    0.06786    3.58
## treatment7mg         -0.51197    0.06786   -7.54
## time7d               -0.30766    0.06786   -4.53
## treatment3mg:time7d -0.23376    0.09597   -2.44
## treatment7mg:time7d  0.66145    0.09597    6.89
##
## Correlation of Fixed Effects:
##             (Intr) trtmn3 trtmn7 time7d trt3:7
## treatmnt3mg -0.236
## treatmnt7mg -0.236  0.500
## time7d      -0.236  0.500  0.500
## trtmnt3mg:7  0.167 -0.707 -0.354 -0.707
## trtmnt7mg:7  0.167 -0.354 -0.707 -0.707  0.500
```

```
MuMIn::r.squaredGLMM(mod2)#marginal and conditional R^2
```

```
##       R2m       R2c
## 0.1142692 0.3280146
```

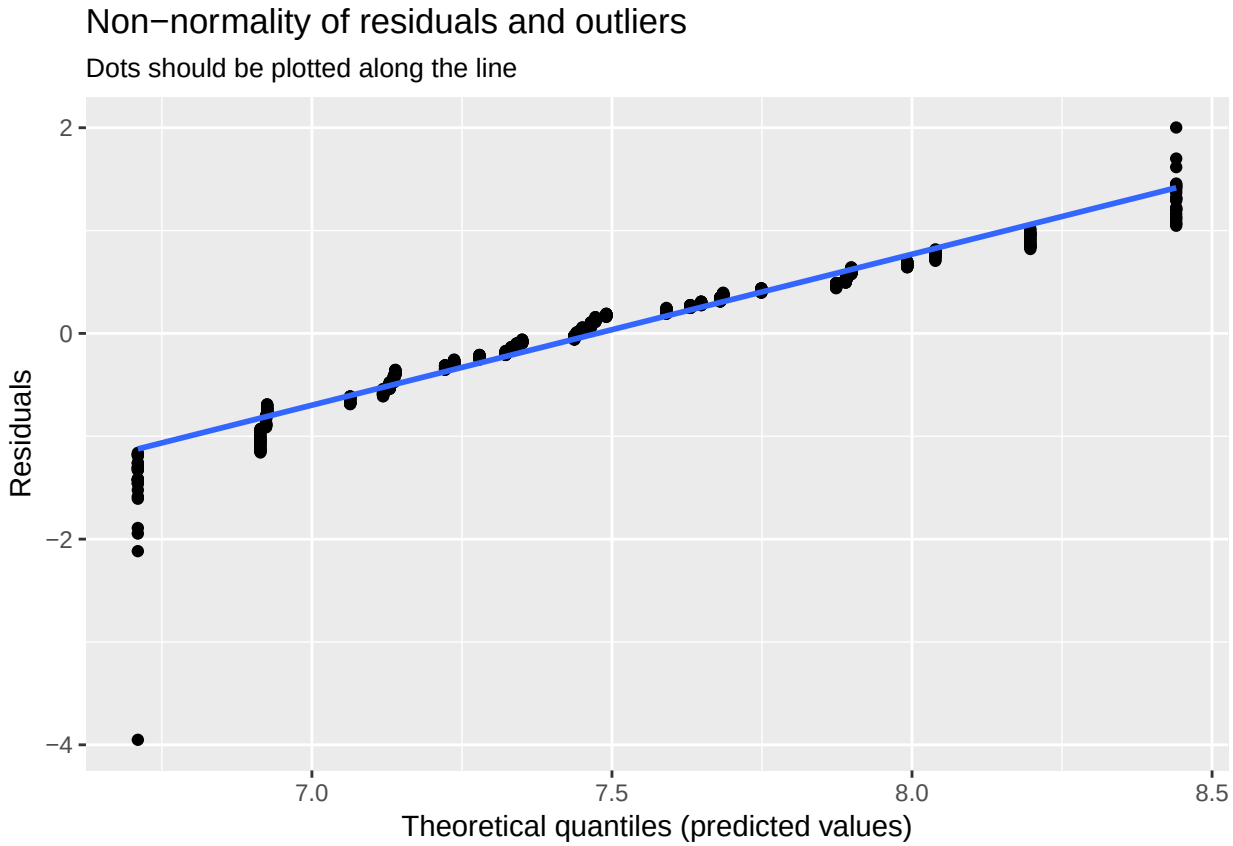
Effective sample size

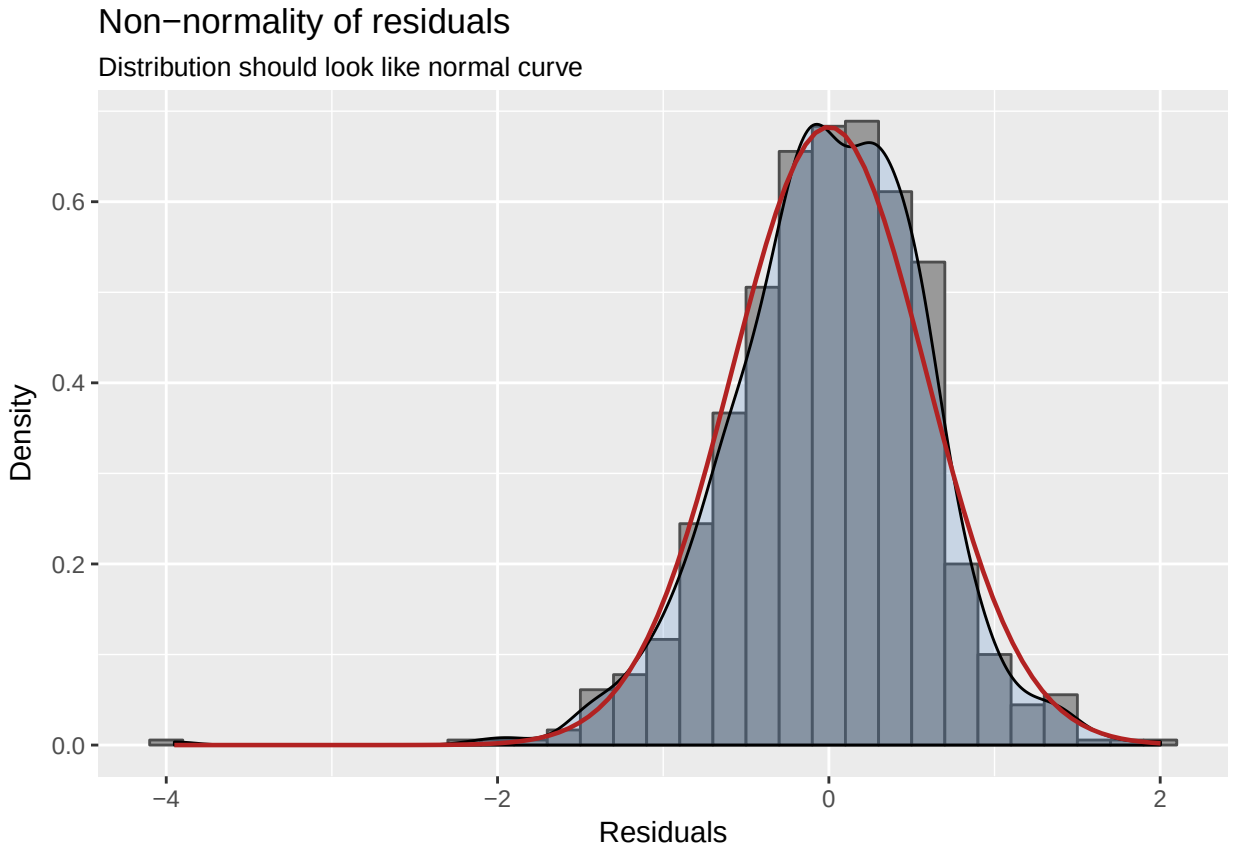
```
25*6/((24*0.241)+1)
```

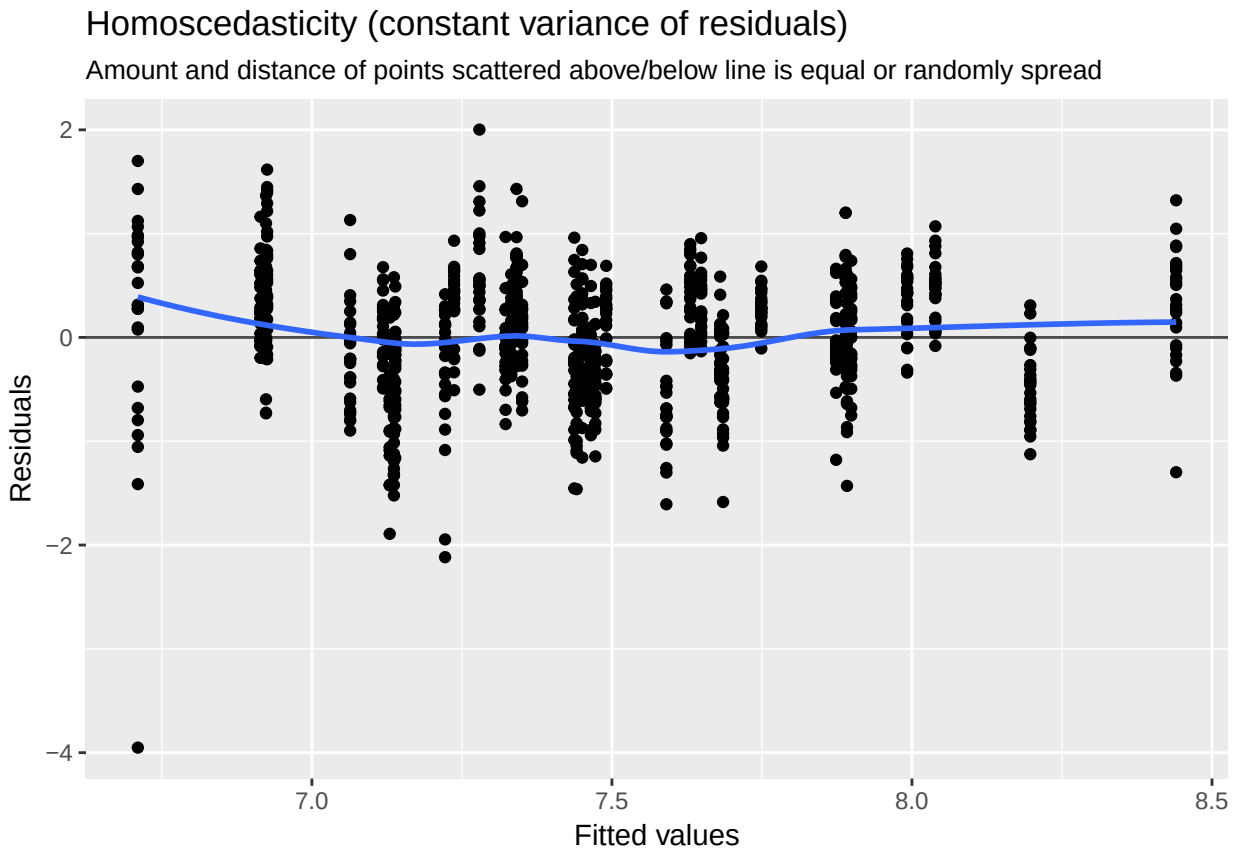
```
## [1] 22.11085
```

testing model assumptions

```
sjp.lmer(mod2, type = "ma")
```


Non−normality of residuals and outliers
Dots should be plotted along the line
Residuals
2
0
−2
−4
7.0
7.5
8.0
8.5
Theoretical quantiles (predicted values)


Non−normality of residuals
Distribution should look like normal curve
Density
0.6
0.4
0.2
0.0
−4
−2
0
2
Residuals



Residuals are normally distributed and we don't have problems with heteroceidasticity either.

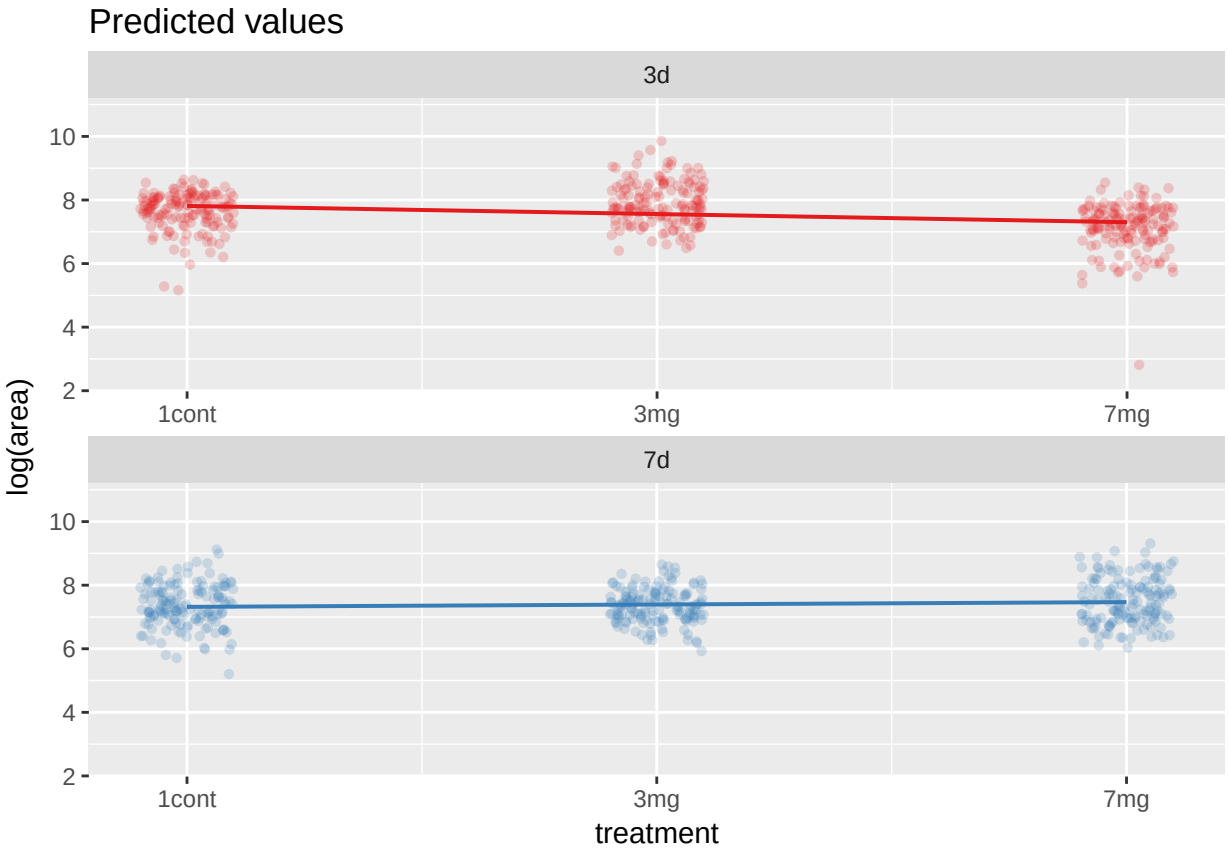
Now we can explore the results of the model. Making contrast to test the effects of treatment and time on area of melanomacrophages:

```
pairs(lsmeans(mod2,  ~ treatment|time))#contrasts between levels of factors
```

```
## time = 3d:
##  contrast        estimate         SE  df t.ratio p.value
##  1cont - 3mg -0.243051550 0.06786072 889  -3.582  0.0010
##  1cont - 7mg  0.511967229 0.06786072 889   7.544  <.0001
##  3mg - 7mg    0.755018779 0.06786072 889  11.126  <.0001
##
## time = 7d:
##  contrast        estimate         SE  df t.ratio p.value
##  1cont - 3mg -0.009287057 0.06786072 889  -0.137  0.9897
##  1cont - 7mg -0.149486293 0.06786072 889  -2.203  0.0712
##  3mg - 7mg   -0.140199236 0.06786072 889  -2.066  0.0976
##
## Results are given on the log (not the response) scale.
## P value adjustment: tukey method for comparing a family of 3 estimates
```

There's only difference between the levels of treatment for 72h.

```
sjp.lmer(mod2, type = "pred", vars = c("treatment", "time"))
```



Summary table with models result for fixed and random factors (to show the table change: no.output=FALSE):

```
sjt.lmer(mod1, mod2,show.header = TRUE,show.est = FALSE, string.std = "Standardized coefficient", string
```

Correlation between small intestine and liver

Are the responses of intestine and liver correlated?

```
#excluding the animal in the liver dataset
#not used in the intestine analysis
newliver<-liver[!(liver$animal==6 ),]
cor.test(newliver$area, intestine$thickness, alternative = "l")#one-tailed hypothesis test (negative as
```

```
##
##  Pearson's product-moment correlation
##
## data:  newliver$area and intestine$thickness
## t = -1.2311, df = 748, p-value = 0.1093
## alternative hypothesis: true correlation is less than 0
## 95 percent confidence interval:
##  -1.0000000  0.0151838
## sample estimates:
##         cor
## -0.04496674
```