Erik Parner
6 October 2016

Solution to exercise 3-3

Question 1

The data for each of the Captopril and Placebo group consists of paired continuous data. Validation of the analysis of the difference

dif = After 1 week - Baseline

by a paired t-test is performed by a Bland-Altman plot and QQ-plot (Figure 1). The reduction in systolic blood pressure in the Captopril group was 12.7 (95% CI: 5.8;19.6), which is statistical significant (p-value=0.003). The reduction in systolic blood pressure in the Placebo group was 4.7 (95% CI: -2.6;12.0), which is not statistical significant (p-values=0.17).

The authors did not compare the two treatment groups, but only analyzed the two groups separately. A statistical non-significant finding can be due to either no true difference between baseline and after 1 week, but it can also be due to low statistical power, i.e. a too small sample to statistical identify the true difference. Since the Placebo group is smaller than the Captopril group the power to detect a significant different will be smaller in the Placebo group.

Comments: The Bland-Altman plots and QQ-plots are needed to validate the assumption of paired t-test, i.e. that the difference (diff=After 1 week – Baseline) follows approximately a normal distribution with the same mean and same standard deviation.

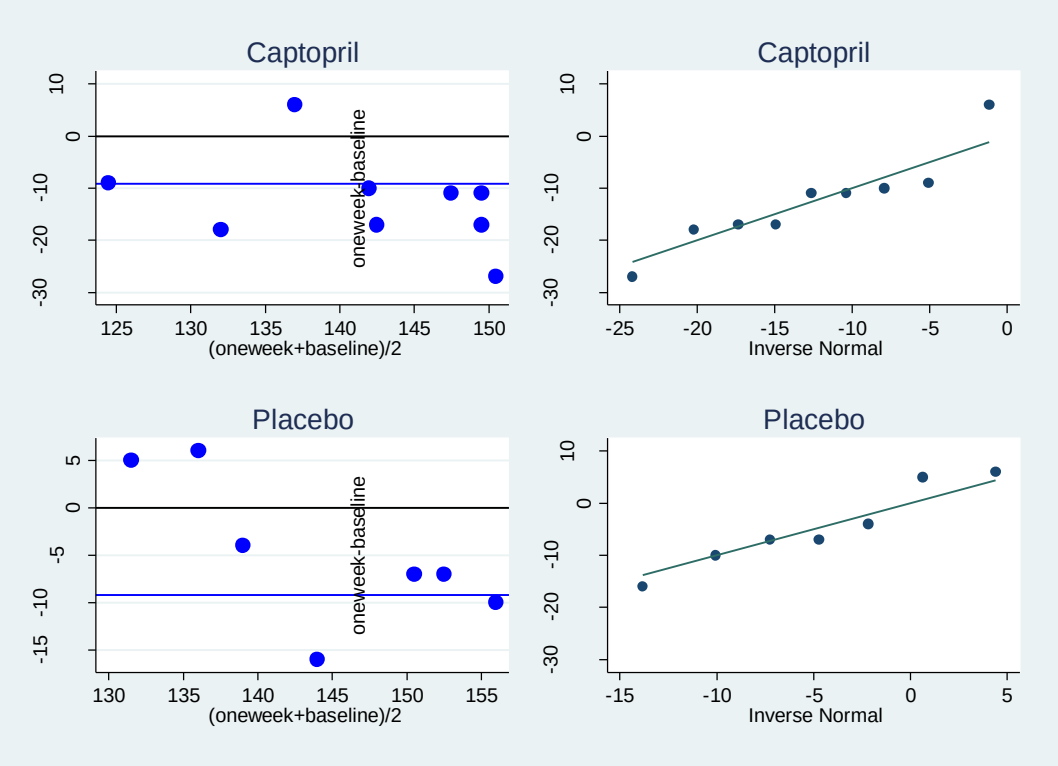

Figure 1. Bland-Altman plots and QQ-plots of the difference for the Captopril respectively the Placebo group.

Question 2

The assumption of the same standard deviation of the reduction in systolic blood pressure in the Captopril and Placebo group was evaluated using the F-test (p=0.78). A two-sample t-test showed that the Captopril group had a 8.0 (95% CI: -1.3;17.2) larger reduction in systolic blood pressure as compared to the Placebo groups, the difference was however not statistical significant (p=0.09).

Comment: For comparison we also applied the Wilcoxon rank-sum (Mann-Whitney) test. Even though the p-values of the t-test and the Wilcoxon rank-sum test are similar, only the latter is statistical significant. In general the non-parametric test will have slightly lower statistical power than the t-test, resulting in general in a slightly higher p-values. This example is however opposite. Both tests are valid. We should of course present the test specified in the study protocol.

Do-file

```
*********************************************************************
* Exercise3-3.do
* Task: Solution to exercise 3-3.
* Erik Parner: 6-10-2016.
*********************************************************************

clear
cd "D:\Teaching\BasicBiostat\Exercises"

capture log close
graph drop _all
log using exercise3-4.log,text replace

use captopril.dta, clear
graph drop _all

* Q1
gene dif=oneweek-baseline
gene ave=(oneweek+baseline)/2
label var dif "oneweek-baseline"
label var ave "(oneweek+baseline)/2"

* Bland-Altman plot and QQ-plot in the Captopril group.
sum dif
local mean=r(mean)
scatter dif ave if group==1,mco(blue) msi(large)  /// ///
    , yline(0,lco(black)) yline(`mean',lco(blue)) ///
    name(graph1) xtit("(oneweek+baseline)/2") ytit("oneweek-baseline")
title("Captopril")
qnorm dif if group==1 , name(graph2) title("Captopril")

* Bland-Altman plot and QQ-plot in the Placebo group.
sum dif
local mean=r(mean)
scatter dif ave if group==2,mco(blue) msi(large)  /// ///
    , yline(0,lco(black)) yline(`mean',lco(blue)) ///
```

```
    name(graph3) xtit("(oneweek+baseline)/2") ytit("oneweek-baseline")
title("Placebo")
qnorm dif if group==2 , name(graph4) title("Placebo")

graph combine graph1 graph2 graph3 graph4
graph drop _all

ttest baseline=oneweek if group==1
ttest baseline=oneweek if group==2

* Q2
sdtest dif,by(group)
ttest dif,by(group)

* Just for comparison. The t-test analysis is fine.
ranksum dif, by(group)

* Close the log file.
log close
```