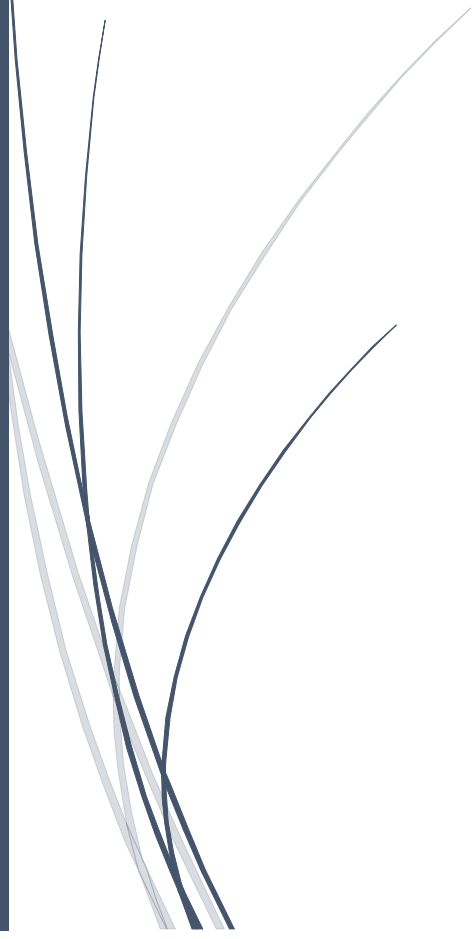
6/8/2020

Clinical Biostatistics

(BCA_CLB – Assignment 3)



Student ID: 6707

(Unikey-skau6707)

Report 1

Introduction

Tension-headache is the most widespread headache disorder which develops a constant, tight, heavy or pressing sensation on or around the head or tautness and tenderness of scalp, neck and shoulder muscles. There are various medical treatments available including physiotherapy to reduce this chronic condition, however have temporary effect. It has been found that acupuncture treatment is effective for tension headache[1,2,3]. In particular, the data obtained on 18 subjects from a controlled cross over trial[4] has been analysed to find the effect of traditional Chinese acupuncture versus placebo acupuncture on chronic tension headache and the report outlines the results of the analysis.

Methods

The effect of Chinese acupuncture versus placebo treatment was evaluated in 18 patients using 2x2 cross over study design. The subjects were randomly allocated to two groups where subjects in group 1 were subjected to Chinese acupuncture during the first period of treatment and placebo acupuncture during the second period of treatment and subjects in group 2 were subjected to these two treatments in reverse order.
Initially, the data was investigated using graphs such as - Subject profile plots which show each subject's change in response over the treatment periods, period-by-period plots for each group (i.e. each sequence) and group-by-period plots which show the mean responses for each group and period.
Further, significance tests were performed to test for equality of carry over effects, treatment effects and period effects using ANOVA[4] method. To test for equal carry over effects sample estimates of the sum of the subjects' responses in period 1 and period 2 between the two groups are compared. And to test for equal treatments effects sample estimates of the differences between subjects' responses in period 1 and period 2 across the two groups are compared whereas to test for equal period effects, a sample estimate of the period 1 response minus the period 2 response for subjects in group 1, with a sample estimate of the period 2 response minus the period 1 response for subjects in group 2 are compared.
Finally, the model assumptions were checked to inspect within-subject residuals and between subject residuals are normally distributed with zero mean and equal variance.
The analysis was conducted in R (R Core Team, 2019) and figures were produced using ggplot function from the package ggplot2 (H. Wickham)[5] and ANOVA analysis was conducted using lmer function from lme4 package[6]. The p-value for significance tests and plots to check model assumptions were obtained using Anova function and ggplot function respectively from car package[7].

Descriptive statistics

The period index (PI - i.e. pain) during first period among subjects in group 1 range from 0 to 43 whereas in second period the PI varies from 0 to 51 with average PI of 26.44 and 30.11 for period 1 and 2 respectively. On the other hand, subjects in group 2 were recorded with PI varying between 21 to 47 in first period and 14 to 43 in second period. And the mean PI for period 1 and 2 is 35.22 and 30.89 respectively for group 2. (as shown in Table 1). On comparing these figures, it seems that Chinese acupuncture was successful at reducing pain, however a statistical analysis is required to support this evidence which is discussed further.

		Min.	Max.	Mean	Median	Std. dev.
Group 1	Period 1	0	43	26.44	32	14.74
(Chinese acupuncture followed by Placebo acupuncture)	Period 2	0	51	30.11	40	16.76
Group 2	Period 1	21	47	35.22	33	10.11
(Placebo acupuncture followed by Chinese acupuncture)	Period 2	14	43	30.89	32	9.96

Table 1: Descriptive statistics for Period Index (i.e. pain) in each of the two periods for two groups receiving either traditional Chinese acupuncture or placebo acupuncture during the first period of treatment and vice versa during the second.

Statistical analysis

A preliminary test was performed using graphical methods to find the effect of treatment

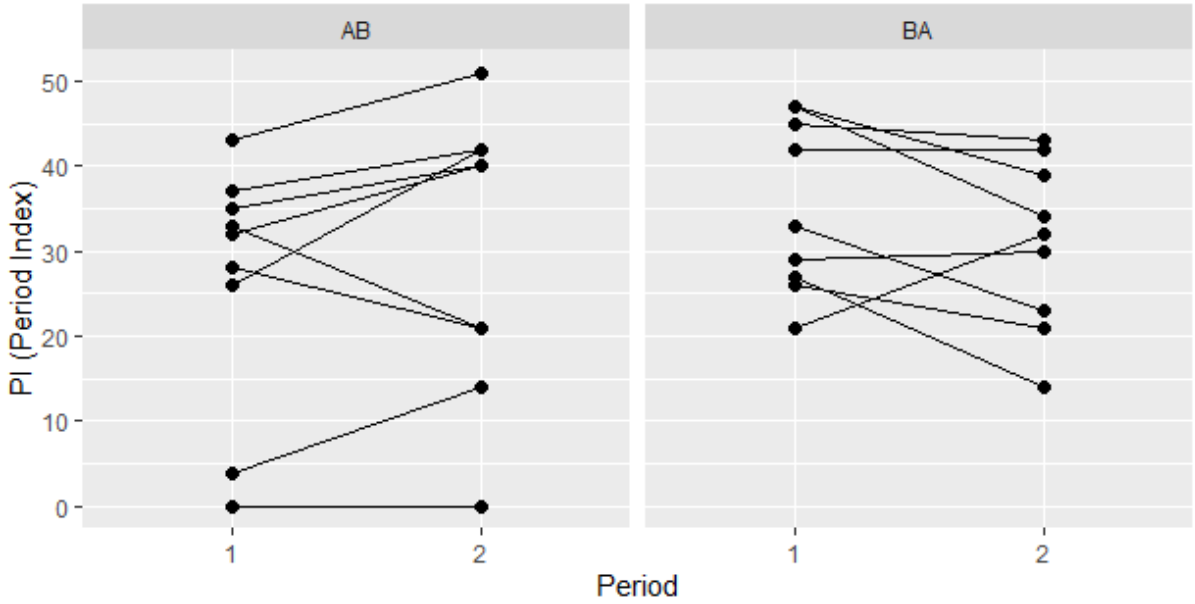


Figure 1: Subject-Profile plot depicts PI recorded for each period in two groups (where AB - Chinese acupuncture followed by Placebo acupuncture and BA - Placebo followed by Chinese acupuncture).

From Figure 1, it is clear that in group 1, most subjects have lower PI (i.e. less pain) when subjected to Chinese acupuncture (i.e. period 1 response) compared to Placebo acupuncture (i.e. period 2 response). There are two exceptions in group 1 – two subjects were found to have high PI when received Chinese acupuncture. In group 2, subjects have higher PI on Placebo treatment (i.e. period 1 response) than period 2 response. Again, there is one exception in group 2, a subject was recorded with more pain when received Chinese acupuncture.

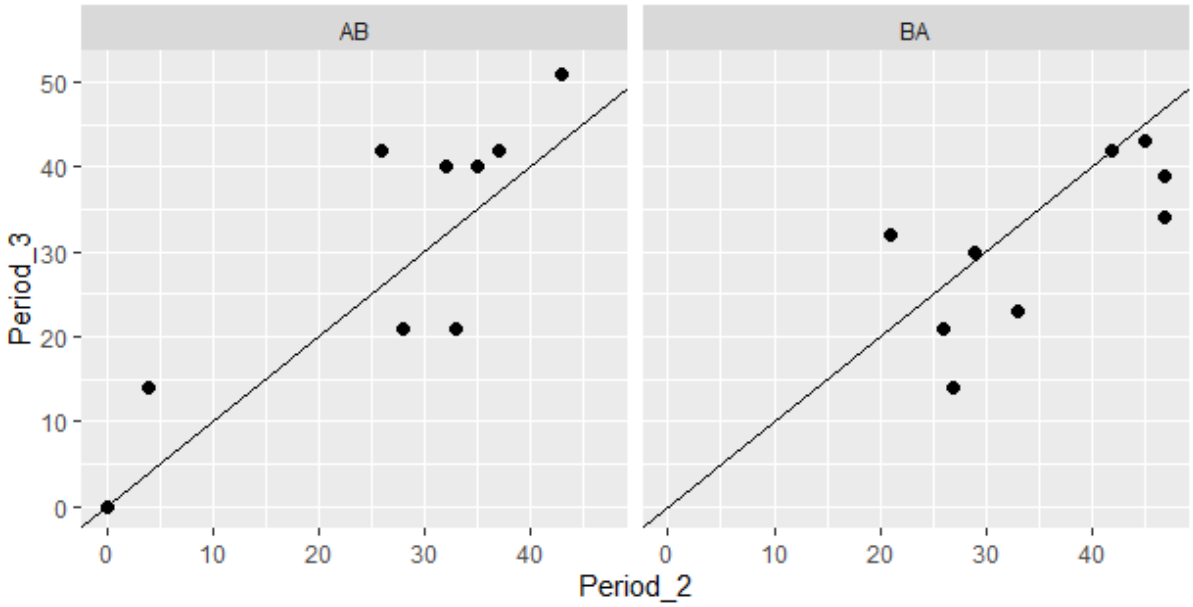


Figure 2: Period by period plot for PI for each group.

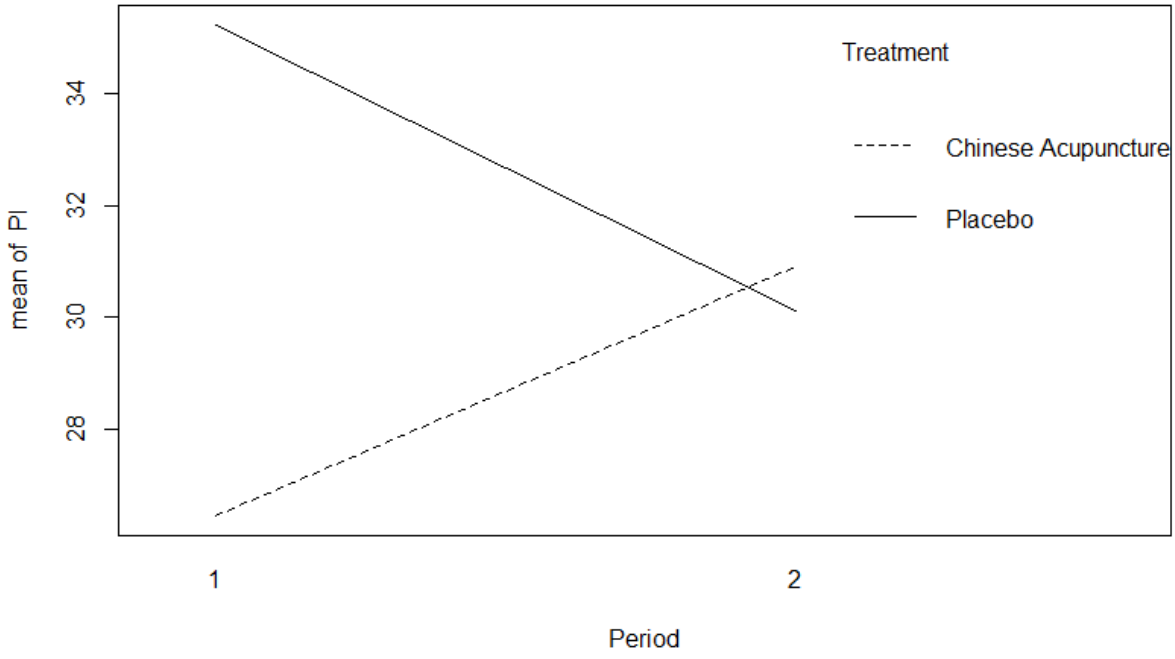


Figure 3: Group by Period plot of PI.

Period vs Period plot are not so informative but it can be seen from interaction plot of mean PI, that in the first period subjects receiving Chinese acupuncture treatment on average, have less pain than those on placebo. However, there is little difference in the second period. Altogether, from the graphs above, it appears that placebo is associated with higher PI – more pain.

Results for hypothesis testing are as follows:

- Test for equal carry over effects- we conclude from results of Anova analysis that there is insufficient evidence to reject our null hypothesis of no difference in carry-over effects (as p-value from Anova analysis is 0.4316 > 0.1 (i.e. at 10% significance level)
- Test for equal period effects - Assuming no difference in carry over effects, hypothesis testing was performed to check for difference of period effects using Anova analysis, again there is no evidence to reject our null hypothesis of no difference in period effects. (as p-value from the test is 0.8661 > 0.05 at 5% significance level)
- Test for equal Treatment effects - Finally, we test that two treatment are equal using similar procedure as above assuming equal carry over effects and the result from Anova analysis suggests that there is no difference in two acupuncture treatments as p-value for the test is 0.0565 > 0.05 at 5% significance level)

	p-value	95% Confidence interval
Carry-over effect	0.4316	(-17.33, 7.78)
Treatment effect	0.0565	(-8.12, 0.12)
Period effect	0.8661	(-3.79, 4.46)

Table 2: p-value and 95% confidence interval from Anova analysis testing for equal carry over effect and equal treatment and period effects.

Both assumptions of the model seem to be verified as within-subject residuals and between-subject residuals follow normal distribution approximately as is evident from the graphs in figure 4 and 5.

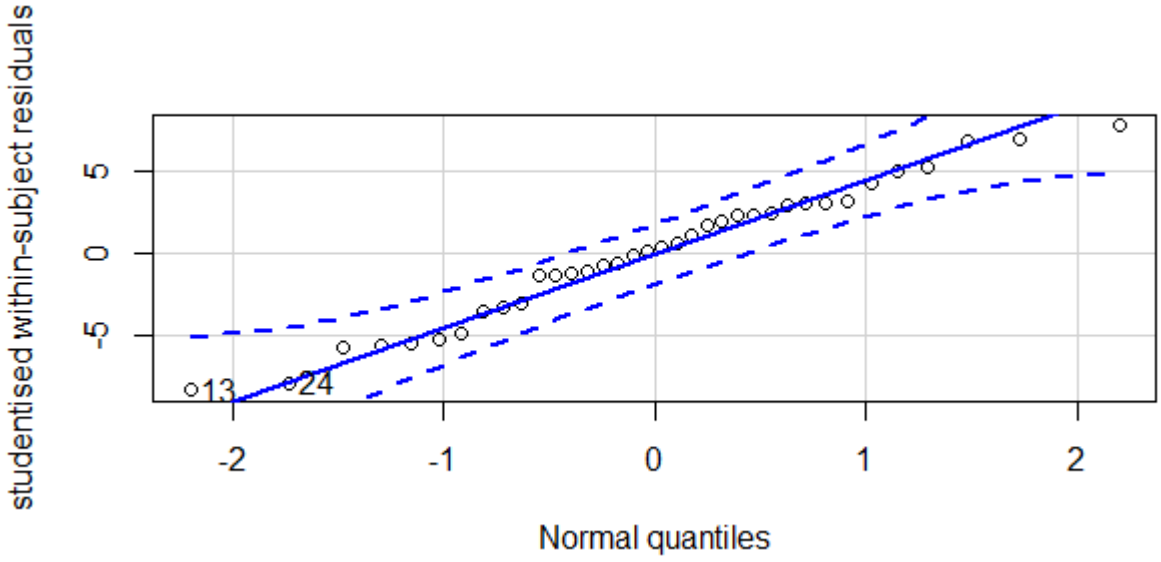


Figure 4: QQplot of Within-subject residuals

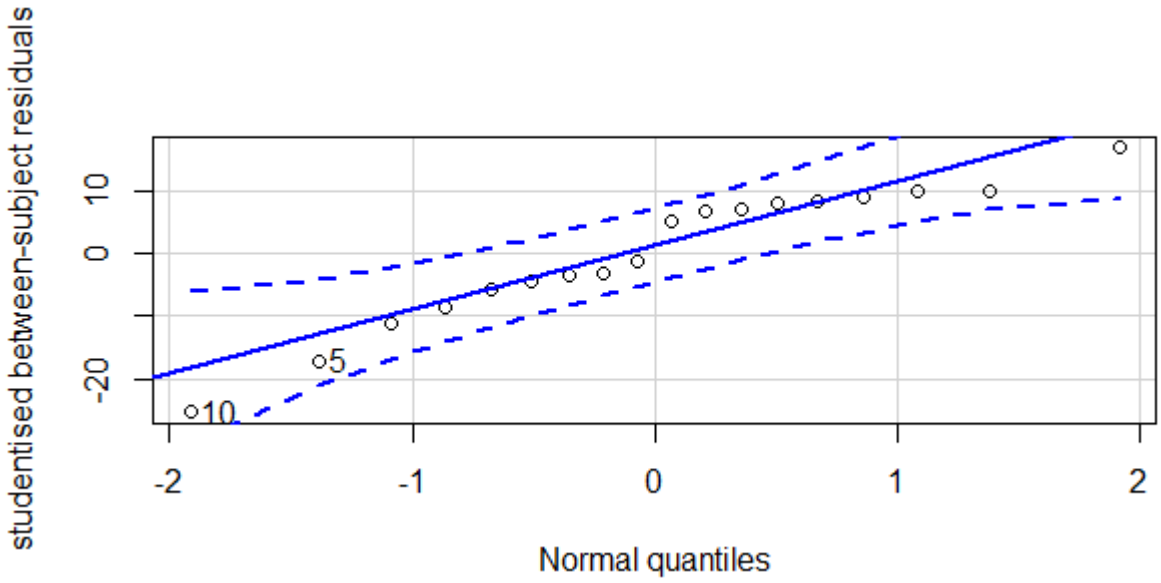


Figure 5: QQplot of Between-subject residuals

Discussion

In conclusion, there is no difference between traditional Chinese acupuncture and Placebo acupuncture as the significance test favours null hypothesis of equal treatment effects (as p-value = 0.06). This result is also evident from the confidence interval (-8.12, 0.12) for difference in treatment effects as it includes null value. This result is contradictory to the result obtained by analysis that was undertaken by Hansen et al[8]. However, no carry-over effect or period effect was apparent as suggested by Hansen et al. In terms of study design, as mentioned in published reading by Hansen et al , the trial consists of five periods where run-in measurements were also obtained including responses from first and second period and we could obtain better results by including baseline measurements for PI only if responses are available for wash-out period. Another factor is that the trial included 25 subjects at beginning, but seven subjects were excluded after trial as three of them suffered from pain other than tension headache, three were unable to attend the treatment and one was hyper-sensitive to treatment, if more data is available then we may get different results.

Report 2

Introduction

Stuttering is a speech disorder that involves frequent and significant problems with normal fluency and flow of speech. The Lidcombe Program (LP) is a behavioural treatment for young children who stutter and its standard format is to deliver the program to individual children. However, if same program is delivered to groups of children then it would be more efficient at reducing stuttering and beneficial at reducing clinical hours. In order to test, whether treatment offered in groups of children is as effective as standard treatment, a non-inferiority trial was conducted on data that contains severity rating for 54 children recorded prior to treatment and at 9 months of post-randomisation for both treatments. This report outlines the analysis performed in detail below.

Methods

The design was a non-inferiority randomised controlled trial. 54 children were randomised for treatment to one of the two treatments that is group or individually delivered LP. Severity rating outcome assessments were conducted pre-randomization and at 9 months.

The experimental hypothesis was that the group treatment was not inferior to the standard treatment. Analysis of covariance was used to test the hypothesis or in other words to test that LP deliver to groups of children is as effective as standard individual treatment. The analysis was performed using aov function from stats package[9] in R software. The conclusion for the test was based on confidence interval, although we obtain p-value from the test.

Descriptive statistics

The severity rating varies from 1 to 9 before delivering LP to individual children with mean 4.1 (SD. = 2.7) which reduces after 9 months of treatment to 1.8 on average, varying between 1 to 5. On the other hand, the range of severity rating recorded prior to group treatment is 2-9 with mean SR of 4.7 (SD. = 2.2) and the LP delivered in group effectively reduces stuttering at 9 months as SR varies between 1-4 with mean 1.7 (SD. = 0.7). The results are summarised in table 1 below and in figure 1.

	Assessment	Min.	Max.	Mean	Median	Std. deviation
Individual Treatment	Prior randomisation	1	9	4.1	3	2.7
	At 9 months	1	5	1.8	1	1.1
Group treatment	Prior randomisation	2	9	4.7	5	2.2
	At 9 months	1	4	1.7	2	0.7

Table 1: Descriptive statistics for Severity rating assessed prior to randomisation and at 9 months post randomisation.

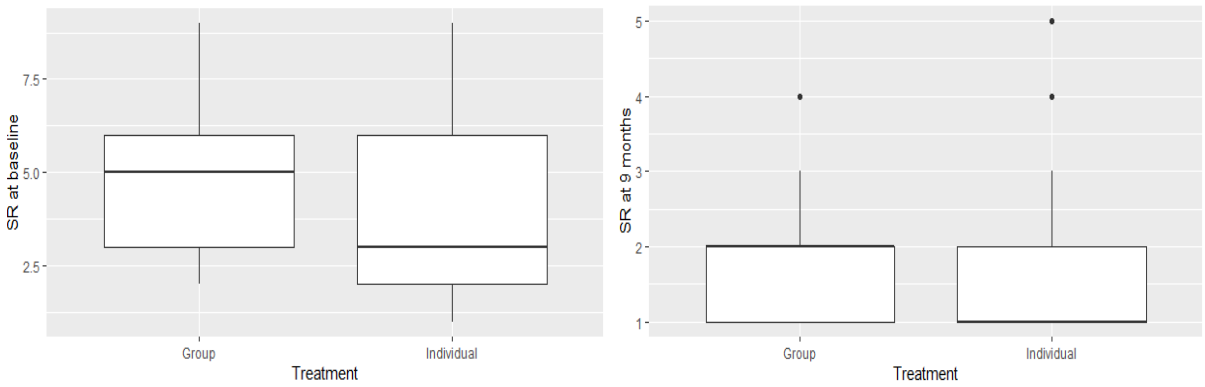


Figure 1: Boxplot for SR prior to randomisation and post randomisation.

Statistical Analysis

Analysis of covariance showed no difference in both treatments. It can be concluded from the results of the estimate -0.14 and confidence interval (-0.68, 0.40) that the group treatment is non-inferior to individual treatment as lower limit of confidence interval for treatment effect does not cross non-inferiority margin.

Discussion

We conclude that both treatments are effective at reducing stuttering. Statistical analysis results suggest that group treatment is not worse than individual treatment. On average the severity rating decreased from 4.7 (Std. dev. = 2.2) to 1.7 (Std. dev. = 0.7) for children under group treatment. The LP delivered to groups of children is not only cost-efficient but also reduces waiting hours when visiting specialists stuttering clinics. The analysis was performed using 47 records where 21 records are from individual treatment and 26 records from group treatment whereas initially the data was obtained for 54 children but the severity rating was missing at 9 months for 7 observations (6 from individual and 1 from group treatment). We could have obtained better results if there is no missing data.

Solution 3

We are given power = 90% (i.e. 0.9) so $\beta = 0.1$ and one-sided type I error as 2.5% (i.e. $\alpha = 0.025$). And expected sensitivity of standard-dose CT for a specific tumor is 90% (i.e. $p = 0.9$). Further, patients are to be randomly allocated to either group at a 1:1 ratio means $n_1 = n_2 = n$.

The sample size using above information for each value of δ is computed in R software given next to part c)

Using formula: $n = \frac{2pq\left(z_{1-\alpha}+z_{1-\beta}\right)^2}{\delta^2}$

a) For non-inferiority margin $\delta = 0.1$ (i.e. 10%)

We require 190 patients in each group to achieve 90% power for a 2.5% one-sided type I error. Thus, a total sample size of 380 patients is required.

b) For non-inferiority margin $\delta = 0.2$ (i.e. 20%)

We require 48 patients in each group to achieve 90% power for a 2.5% one-sided type I error. Thus, a total sample size of 96 patients is required.

c) For non-inferiority margin $\delta = 0.15$ (i.e. 15%)

We require 85 patients in each group to achieve 90% power for a 2.5% one-sided type I error. Thus, a total sample size of 170 patients is required.

Code for Sample size calculation:

```
> ## For delta = 10%
> ceiling(2 * 0.9 * (1-0.9) *((qnorm(0.025) + qnorm(0.1))/0.1)^2)
[1] 190
> ## For delta = 20%
> ceiling(2 * 0.9 * (1-0.9) *((qnorm(0.025) + qnorm(0.1))/0.2)^2)
[1] 48
> ## For delta = 15 %
> ceiling(2 * 0.9 * (1-0.9) *((qnorm(0.025) + qnorm(0.1))/0.15)^2)
[1] 85
```

References:

1. Essentials of Chinese acupuncture. Beijing: Foreign Languages Press, 1980
2. Bischko JJ. Acupuncture in headache. Res Clin Stud Headache 1978; 5:72-85
3. Jensen LB, Tallgren A, Troest T, Jensen SB. Effect of acupuncture on myogenic headache. Scand J Dent Res 1977; 85:456-70
4. Jones B and Kenward M (2003) Design and Analysis of Cross-over Trials, 2nd edition, Chapman and Hall, London.
5. H. Wickham. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2016.
6. Douglas Bates, Martin Maechler, Ben Bolker, Steve Walker (2015). Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software, 67(1), 1-48. doi:10.18637/jss. v067.i01
7. John Fox and Sanford Weisberg (2019). An {R} Companion to Applied Regre ssion, Third Edition. Thousand Oaks CA: Sage. URL: https://socialsciences.m cmaster.ca/jfox/Books/Companion/
8. Table1 and 2 from Hansen PE, Hansen JH. Acupuncture treatment of chronic t ension headache-a controlled cross-over trial. Cephalalgia 1985;5:137-42. Osl o. ISSN 0333-1024
9. R Core Team (2019). R: A language and environment for statistical computin g. R Foundation for Statistical Computing, Vienna, Austria. URL https://ww w.R-project.org/.

APPENDIX

Code for Question 1, 2 and 3

```
###################### QUESTION 1 ######################

suppressMessages(library(tidyverse)) # suppress startup messages

acup<- read_csv("acupuncture.csv", col_types = cols())
acup<- mutate(acup, Group = factor(Group), Patient = factor(Patient))

## Descriptive statistics
summary(acup[c(1:9),c(1:5)]);summary(acup[c(10:18),c(1:5)])
sqrt(var(acup[c(1:9),c(3:4)]));sqrt(var(acup[c(10:18),c(3:4)]))
## 1.2 Stack data for plots ---------------------------------------

treats <- c("Chinese Acupuncture","Placebo")
## stack data
acup_stacked <- acup %>% select( -c("Period_1"))%>% gather("Period",
"PI",Period_2, Period_3,
    factor_key = TRUE) %>% mutate(Period = as.numeric(Period),
    Treatment = factor(rep(c(treats, rev(treats)), each = 9),
    levels = treats))

str(acup_stacked)
summary(acup_stacked)

xtabs(~Period+Treatment, data = acup_stacked)

##Subject Profile plots
library(ggplot2)
ggplot(acup_stacked, aes(factor(Period), PI, group = Patient)) +
  geom_line() + facet_wrap(acup_stacked$Group) + geom_point(size=2) +
  xlab("Period") + ylab("PI (Period Index)")

## Period-by-period plots
ggplot(acup, aes(Period_2, Period_3)) + geom_point(size=2) +
facet_wrap(acup$Group) + geom_abline(aes(intercept = 0, slope = 1))

##Period vs Period plot not so informative but can see from interaction plot
##of mean PI, that in the first period subjects receiving acupuncture treatment,
##on average, have less pain than those on placebo.
##However, there is little difference in the second period.
with(acup_stacked, interaction.plot(Period, Treatment, PI))

##Altogether, from the graphs above, it appears that
##placebo is associated with higher PI – more pain. We will need to analyse
##the data further to be sure of what is happening in the data.

acup <- mutate(acup, P2plusP3 = Period_2 + Period_3,
               P2minusP3 = Period_2 - Period_3, P3minusP2 = Period_3 - Period_2,
               cross_diff = (Group == "AB")*P2minusP3 + (Group == "BA")*P3minusP2)
```

```
## t-test for carry-over effects
(tt1 <- t.test(P2plusP3 ~ Group, data = acup, var.equal = TRUE))
## Thus no carry over effects as p-value is 0.4316 > 0.05

## t-test for treatment effects
(tt2 <- t.test(P2minusP3 ~ Group, data = acup, var.equal = TRUE))
## NO treatment effect as p-value is 0.05647 > 0.05

## t-test for period effects
(tt3 <- t.test(cross_diff ~ Group, data = acup, var.equal = TRUE))
## strong evidence for no period effects as p-value is 0.8661 > 0.05

library(knitr)
ttests1 <- t(sapply(list(tt1, tt2, tt3),
                    function(x) c(x$statistic, x$parameter, x$p.value)))

row.names(ttests1) <- paste(c("Carry-over", "Treatment", "Period"), "effects")
colnames(ttests1) <- c("t value", "df", "p value")
kable(ttests1, digits = c(4,0,4), caption = "Results of t-tests.")

##So we have no evidence of a difference in carry-over effects.
##Therefore, we assume they are equal and proceed to test for a difference
##in treatment and period effects. We see that there is no evidence of a
##difference in period effects as well as difference in treatment
##effects. We will later calculate a point estimate and C.I. for this difference.


############## ANOVA ############

addmargins(xtabs(~Group+Patient, data = acup_stacked))
acup.aov <- aov(PI ~ Group+factor(Period)+Treatment+Error(Patient),
                data = acup_stacked)
summary(acup.aov)

############ LME4 PACKAGE ########
## Alternatively, using newer 'lme4' library
suppressMessages(library(lme4))
acup.lmer <- lmer(PI~ Group + factor(Period)+Treatment + (1|Patient),
                  data = acup_stacked)
summary(acup.lmer)
anova(acup.lmer)
car::Anova(acup.lmer, test.statistic = "F") # produces F tests

############ nlme PACKAGE ########
suppressMessages(library(nlme))
acup.lme <- lme(PI ~ Group + factor(Period)+Treatment,
                random = ~1|Patient, data = acup_stacked)
summary(acup.lme)
anova(acup.lme)
car::Anova(acup.lme, test.statistic = "F")    # produces Chi-squared tests

## same results as obtained for t-test

## Using t-tests confidence intervals
tt1$conf.int/2  #for carry over effect
```

```
tt2$conf.int/2  #for treatment effect
tt3$conf.int/2  #for period effect
## OR we can use alternate method to get CI using lme4 & nlme package
confint(acup.lmer)
intervals(acup.lme)
############# MODEL CHECKING ############
plot(acup.lmer)

## within subject residuals
suppressMessages(library(car))
qqPlot(residuals(acup.lmer),xlab="Normal quantiles",ylab = "studentised within-
subject residuals")

##These appear to be normally distributed, as desired, and there is no obvious
##relationship between the residuals and the fitted values.

## Between subject residuals
qqPlot(ranef(acup.lmer)$Patient[[1]],xlab="Normal quantiles",ylab = "studentised
between-subject residuals")

##These residuals also look normally distributed, as desired.



###################### QUESTION 2 ####################

### Reading the dataset "lp_singlevgrp.csv"

suppressMessages(library(tidyverse)) # suppress startup messages
stutter<- read_csv("Lp_singlevgrp.csv", col_types = cols())

## deleting records with missing values
stutter_lp <- na.omit(stutter)

#descriptive statistics
summary(stutter_lp[c(1:21),]);summary(stutter_lp[c(22:47),])
sqrt(var(stutter_lp[c(1:21),]));sqrt(var(stutter_lp[c(22:47),]));

## declaring grp as categorical variable and renaming it in new column group
stutter <- mutate(stutter, grp = factor(grp))
stutter_lp$grp[stutter_lp$grp == 0] <- "Individual"
stutter_lp$grp[stutter_lp$grp == 1] <- "Group"

ggplot(stutter_lp, aes(x=grp, y=dp1)) + geom_boxplot()+ xlab("Treatment") +
ylab("SR at baseline")
ggplot(stutter_lp, aes(x=grp, y=dp4)) + geom_boxplot()+ xlab("Treatment") +
ylab("SR at 9 months")

## ANOVA
stutter.aov <- aov(dp4 ~ dp1+grp, data = stutter_lp)
summary(stutter.aov)
confint(stutter.aov)
coefficients(stutter.aov)

## USING alternative method
```

```
library(rstatix)
s.aov <- stutter_lp %>% anova_test(dp4 ~ dp1 + grp)
get_anova_table(s.aov)


###################### QUESTION 3 ######################

## For delta = 10%
ceiling(2 * 0.9 * (1-0.9) *((qnorm(0.025) + qnorm(0.1))/0.1)^2)
library(TrialSize)
TwoSampleProportion.Equivalence(alpha=0.025, beta=0.1, p1=0.9,
                        p2=0.9, k=1, delta=0, margin=0.1)
epiR::epi.ssninfb(treat=0.90, cont=0.90, delta=-0.1, n=NA,
            power=0.9, alpha=0.05)

## For delta = 20%
ceiling(2 * 0.9 * (1-0.9) *((qnorm(0.025) + qnorm(0.1))/0.2)^2)
TwoSampleProportion.Equivalence(alpha=0.025, beta=0.1, p1=0.9,
                        p2=0.9, k=1, delta=0, margin=0.2)
epiR::epi.ssninfb(treat=0.90, cont=0.90, delta=-0.2, n=NA,
            power=0.9, alpha=0.025)


## For delta = 15 %
ceiling(2 * 0.9 * (1-0.9) *((qnorm(0.025) + qnorm(0.1))/0.15)^2)
TwoSampleProportion.Equivalence(alpha=0.025, beta=0.1, p1=0.9,
                        p2=0.9, k=1, delta=0, margin=0.15)
epiR::epi.ssninfb(treat=0.90, cont=0.90, delta=-0.15, n=NA,
            power=0.9, alpha=0.025)
```