A case study in non-inferiority margin selection in a two-arm trial

by

Xiao Liang

B.S., Zhejiang Normal University, 2009

A REPORT

submitted in partial fulfillment of the requirements for the degree

MASTER OF SCIENCE

Department of Statistics
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2017

Approved by:

Major Professor
Christopher I. Vahl

Copyright

Abstract

Non-inferiority trials have been widely used in many medical areas. The goal of a non-inferiority trial is to show that a new test therapy is either better or not too much worse than the active control rather than showing the test therapy is superior to a negative control (i.e. placebo). The appeal of a non-inferiority trial is that it is often unethical to give some patients a treatment with no therapeutic benefit. When designing a non-inferiority trial, the issues of assay sensitivity, sample size, constancy condition, and a suitable non-inferiority margin need to be considered. A poor choice of the non-inferiority margin is a major reason that many non-inferiority trials fail. A numerical example is presented to show how to estimate the non-inferiority margin without historical data.

Keywords: non-inferiority trial, non-inferiority margin, superiority margin

Table of Contents

List of Figures

List of Tables

Acknowledgements

I would like to thank everyone (faculty, family, and classmates) for supporting my master's report. I would like to thank my major professor, Dr. Christopher Vahl, who introduced me to the topic of my master's report. In addition, I would like to thank Dr. Wei-Wen Hsu and Dr. Abigail Jager for their time and comments on my master's report. Furthermore, I would like to thank the Department of Statistics at Kansas State University for providing me with the opportunity to achieve my Master's degree.

Chapter 1 - What is a Non-inferiority Trial?

Introduction

In recent years, testing for non-inferiority has become one of the hottest topics in the area of clinical trials. There was only one non-inferiority trial published between 1989 and 1998. In contrast, there were 582 published non-inferiority trials between 2000 and 2009 (Hurley, McKibbin, Moroney, & Suda, 2011). Among the 43 new drug applications approved by the FDA between 2002 and 2009, two thirds of them provided a statistical analysis with tests for non-inferiority (Hurley et al., 2011). Non-inferiority trials have been widely used to test drugs and vaccines used to treat or prevent many serious and life-threatening medical conditions such as blood clots, cancer, heart and kidney failure as well as infections and infectious diseases (Rothmann, Wiens, & Chan, 2012).

Before talking about the issues and statistical explanations of non-inferiority trials, we would like to introduce how the idea of non-inferiority generated. When we discuss the hypothesis tests, the most common one will be a superiority test which is the easiest way to evaluate the effect of a test treatment compared to a control treatment in clinical trial. The control treatment can be a placebo control, which has no medical benefit, or an active control, which is a tested therapy. In general, the purpose of the superiority test is to test whether the treatment is better than the control treatment. Let $M_1 > 0$ represent the smallest benefit of the treatment over a placebo control that is acceptable. When researchers compare a new treatment to a placebo control or an active control, they try to make a statement based on the result of the data analysis. Shortly speaking, the statement 'superiority' means that the test treatment is better than the control. In statistics, we would say that 'superiority' means that the data analysis tells

us to reject the null hypothesis that the distributions of the treatment and the placebo (or the active control) are same (Schumi & Wittes, 2011).

Now, let us use a figure (Figure 1.1) to explain the relationship among the superiority margin, confidence interval, and the idea of superiority. Suppose we have treatments A and B, and we are testing whether treatment A is superior to treatment B. If the confidence interval of the treatment difference is on the right side of superiority margin M_1, then we can claim that treatment A is superior to treatment B. Otherwise, we cannot state that treatment A is superior to treatment B.

Another common statistical test is the equivalence test. Instead of testing whether the test treatment is better than the control treatment, an equivalence test tries to determine whether the effect of the test treatment is the same as the control treatment for practical purposes. Shortly speaking, 'equivalence' means the absolute difference between the treatment and the control is smaller than a fixed margin (Schumi & Wittes, 2011).

Now, let us use another figure (Figure 1.2) to explain the relationship among equivalence margin, confidence interval, and the idea of equivalence. Suppose we have treatments A and B, and we are testing whether treatment A and treatment B are equivalent to each other. If the confidence interval for the treatment difference is between $-M_3$ and M_3 (with $M_3 > 0$), then we can say that treatment A is equivalent to treatment B. Otherwise, we cannot state that treatment A is equivalent to treatment B.

Now, suppose we have to test a new treatment. Superiority testing, which in this case compares the new treatment to a placebo treatment, will be considered to be unethical, especially for testing drugs used to treat some medical conditions, such as cancer or sepsis, that are known to result in serious harm or death to patients receiving the placebo treatment. Instead of using a

placebo as the control, a drug that has already been shown to be superior to a placebo, an active control, is used to judge the efficacy of the test drug. If the new treatment is not superior to the active control, we cannot make conclusion that the new treatment has no medical benefit. Then we use equivalence test to examine between the new treatment and the active control. If it turns out that they are not equivalent, we still cannot state that the new treatment has no medical benefit. The new treatment may be slightly less effective, but it may be a benefit to patients. Then we need a new hypothesis test to compare the new treatment to the active control. This is the idea that leads to a non-inferiority test.

Even though ethical purposes could be the most important reason for non-inferiority trials, there are several situations when researchers may consider not performing a non-inferiority trial according to Freedman (1987).

1. Researchers should not try to perform a non-inferiority trial, when there is new evidence implies that there is uncertain side effect in the standard treatment or active control.
2. There is no standard therapy can deal with some kind of disease. In other words, we cannot find an active control agent.
3. The effect of standard therapy is not significant.

Now, let's introduce the definition of non-inferiority trials. If the effect of a new treatment is either better than or not too much worse than the effect of a proven treatment for a certain disease, the new treatment's effect is said to be non-inferior to the effect of the proven treatment. It is common to say that the new treatment is non-inferior to the proven treatment.

Let us use figure (Figure 1.3) to explain the relationship among non-inferiority margin, confidence interval, and the idea of non-inferiority. Suppose we have treatments A and B, and

we are testing whether treatment A is non-inferior to treatment B. In this trial, we have two groups of patients who are going to be randomly assigned treatments. In addition, a response variable will be measured on each patient. In order to judge the efficacy of each treatment, we are going to make the comparison, which can be called the treatment difference, between the mean of the response variable of the two groups of patients. If the confidence interval of the treatment difference is on the right side of non-inferiority margin $-M_2$ (with $M_2 > 0$), then we can say that treatment A is non-inferior to treatment B. Otherwise, we cannot state that treatment A is non-inferior to treatment B.

The definition of non-inferiority sounds simple. However, there is a hazard associated with non-inferiority trials in that making conclusions in some non-inferiority trials is often complicated. For example, suppose A, B, C are three different drugs. Let us make some assumptions:

1. Drug A is a proven drug for disease X.
2. Drug B is non-inferior to drug A.
3. Drug C is non-inferior to drug B.

Then, we can make some conclusions. According to the assumption 1 and 2, we can say that drug B is better than placebo. However, we cannot make too strong of a claim by adding in assumption 3. Although, drug C is non-inferior to drug B, we cannot be sure that drug C is better than the placebo. Making things worse, we cannot conclude that drug C is non-inferior to drug A, either.

We can note that selection of this margin is a big challenge and an active area of research. If the non-inferiority margin is too small, it will be difficult to demonstrate usefulness of the new drug, which may be able to treat some diseases. On the other hand, if the non-inferiority margin

is too large, it may lead to the approval of some treatments, which have only slight or no benefit to patients. In this report, we are focusing on the selection of a non-inferiority margin.

Figures

Figure 1.1. Relationship between the superiority margin and confidence interval

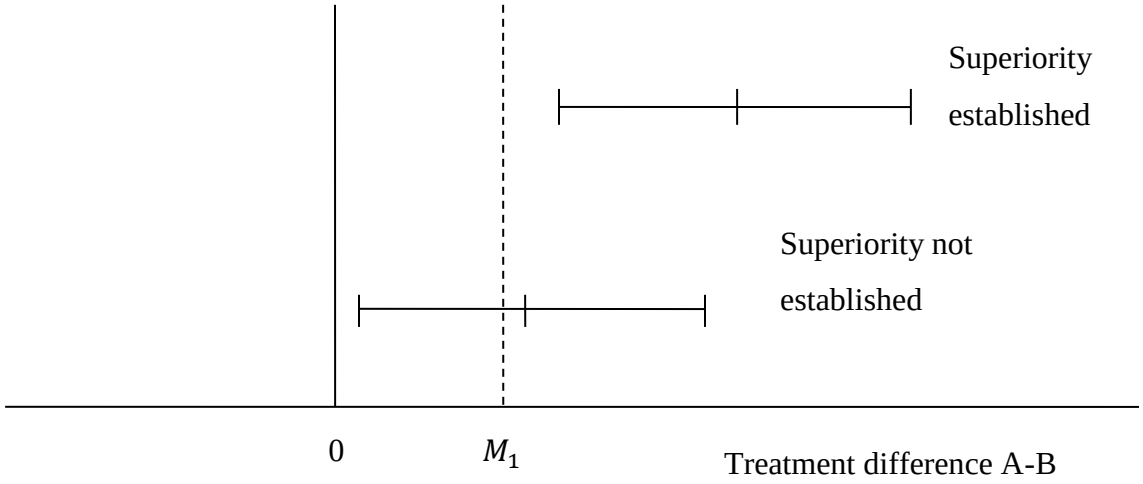


Figure 1.2. Relationship between the equivalence margin and confidence interval

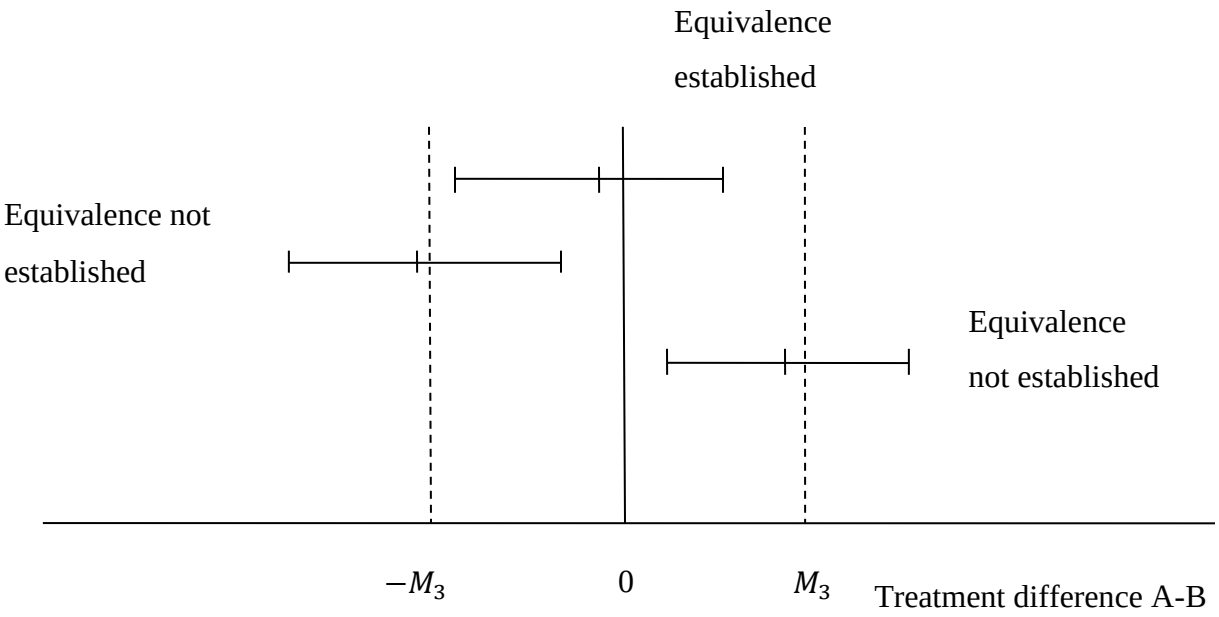
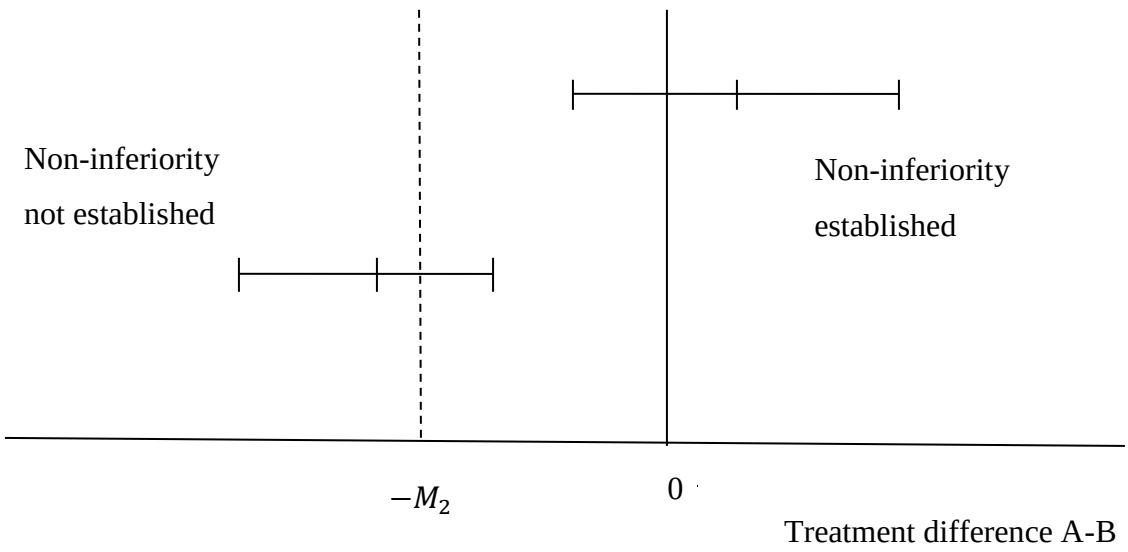


Figure 1.3. Relationship between the non-inferiority margin and confidence interval



Chapter 2 - Non-inferiority Trials Considerations

Assay Sensitivity

"Assay sensitivity is a study property defined as the ability to distinguish an effective treatment from a less effective or ineffective treatment" (FDA, 2010). Assay sensitivity is therefore a very important issue in the research of non-inferiority trials, since the goal of non-inferiority trial is to compare different treatments. If a non-inferiority trial lacks assay sensitivity, then any statistical analysis of this trial will have no practical meaning.

According to the FDA, evidence of assay sensitivity is demonstrated in two ways:

1) Historical evidence of effectiveness of the active control. In other words, we can easily distinguish the proved active control from placebo.
2) Design of the proposed non-inferiority study.

Sample Size

Sample size is another issue that needs to be considered in the study of non-inferiority trials. Generally, the sample sizes of non-inferiority trials, active-controlled superiority trials, and placebo-controlled trials satisfy the following two relationships:

1) The sample sizes of non-inferiority trials are smaller than the sample sizes of active-controlled superiority trials.
2) The sample sizes of placebo-controlled trials are smaller than the sample sizes of non-inferiority trials (Snapinn, 2000).

In addition, according to the FDA, the sample size of a non-inferiority trial is related to many factors, including the estimated success rate of both the active control and the new treatment, the margin of difference, and the variance. If we do not change all of the factors

except for the margin of difference, we can see the following pattern: the sample size increases as the margin of difference decreases; the sample size decreases as the margin of difference increases.

Constancy Condition

Constancy condition is an issue we need to consider when we are dealing with the non-inferiority trials without placebo group. Constancy condition means that we can hold the effect of the active control over time in non-inferiority trials (Chow & Shao, 2005). In practice, even though we have the active control, the effect of the active control in the current patients might be not as good as the effect in the historical patients. Without the placebo group, we are not able to test the effect of the active control. As a result, lack constancy condition will result in a useless statistical analysis.

Selecting Margins

General Method

In Chapter 1, we give the definition of non-inferior as phrased below below by Rothmann, et al(2012):

> "If the effect of an experimental therapy on an endpoint is either better than or not too much worse than the effect of a control therapy on that same endpoint, the experimental therapy's effect is said to be non-inferior to the effect of the control therapy. It is common to say that the experimental therapy is non-inferior to the control therapy."

It is clear that we need to explicitly define how much is 'not too much' if we want to analyze non-inferiority trials. In other words, we need to know how to estimate the non-inferiority margin.

First, let us make some assumptions.

y_{Ai} : Response for active control agent for i^{th} subject,

$y_{Ai} \sim N(\mu_A, \sigma_A^2), i = 1, 2, \cdots, n_A$,

y_{Ti} : Response for test therapy for i^{th} subject,

$y_{Ti} \sim N(\mu_T, \sigma_T^2), i = 1, 2, \cdots, n_T$,

y_{Pi} : Response for placebo for i^{th} subject,

$y_{Pi} \sim N(\mu_P, \sigma_P^2), i = 1, 2, \cdots, n_P$,

M_1: Superiority margin ($M_1 > 0$), and

$-M_2$: Non-inferiority margin ($M_2 > 0$).

The hypotheses for non-inferiority trial:

$$H_0: \mu_T - \mu_A \leq -M_2$$

$$H_1: \mu_T - \mu_A > -M_2$$

According to the International Conference on Harmonization Guideline E10 (ICH, 2001), the non-inferiority margin $-M_2$ should be chosen to satisfy at least the following two criteria:

Criterion 1: We want the ability to claim that the test therapy is non-inferior to the active control agent and is superior to the placebo (even though the placebo is not considered in the active control trial).

Criterion 2: The non-inferiority margin should be suitably conservative and variability should be taken into account.

If $-M_2$ is a fixed number, then we are able to test the hypotheses by using some classic statistical techniques. Unfortunately, in reality, usually, $-M_2$ is not pre-specified (Chow & Shao, 2005).

In addition, it is almost impossible to find a fixed $-M_2$ under criterion 1. Let $M_1 > 0$ be a superiority margin associated with a placebo-controlled trial, which is a superiority trial that examines a test therapy over a placebo control. We are able to assume that $\mu_A - \mu_P > M_1$, because the historical evidence has proved that the active control is superior to the placebo.

However, when the test therapy is non-inferior to the active control, for a fixed M_2, we are not able to say that the test therapy in superior to the placebo unless $M_2 = 0$. Statistically speaking, if $\mu_A - \mu_P > M_1$ and $\mu_T - \mu_A > -M_2$, then

$$\mu_T - \mu_P = \mu_T - \mu_A + \mu_A - \mu_P > M_1 - M_2$$

So, we cannot make statement that $\mu_T - \mu_P > M_1$ unless $M_2 = 0$. Also, if $\mu_T - \mu_P > 0$, then we can assume that $\mu_T - \mu_P > M_1 - M_2 > 0$. Therefore, $M_1 > M_2$. In addition, we have $\mu_A - \mu_P > M_1 > M_2$. We can assume the following equation based on the above result: $M_2 = \gamma(\mu_A - \mu_P)$ where $0 \leq \gamma \leq 1$. In addition, γ is a fix number. Therefore, the non-inferiority margin M_2 is associated with the active control effect $\mu_A - \mu_P$.

Next, we are going to determine a non-inferiority margin under criterion 1. Let $M_1 > 0$ be a superiority margin associated with a placebo control submitted to the trial. Since we have $\mu_A - \mu_P > M_1 > M_2$, then we can assume that M_2 and the superiority margin M_1 have the following relationship: $M_2 = rM_1$ where r is a pre-specified number, which is between 0 and 1. Suppose that the test therapy is non-inferior to the active control agent and superior to the placebo, then we have the following set of inequalities:

$$\begin{cases} \mu_T - \mu_A > -M_2 \\ \mu_T - \mu_P > M_1 \end{cases}$$

If $\mu_T - \mu_A$ reaches the lowest value $-M_2$, then we have $\mu_T = \mu_A - M_2$.Thus,

$$\mu_T - \mu_P = \mu_A - M_2 - \mu_P > M_1$$

so that $\mu_A - \mu_P - M_1 > M_2$. In this case, if M_2 reaches the highest value, then

$$M_2 = \mu_A - \mu_P - M_1.$$

Therefore, we have

$$\begin{cases} M_2 = \mu_A - \mu_P - M_1 \\ M_1 = \dfrac{M_2}{r} \end{cases} \Longrightarrow M_2 = \mu_A - \mu_P - \frac{M_2}{r}$$

$$\Rightarrow rM_2 = r(\mu_A - \mu_P) - M_2$$

$$\Rightarrow (1+r)M_2 = r(\mu_A - \mu_P)$$

$$\Rightarrow M_2 = \frac{r}{1+r}(\mu_A - \mu_P).$$

Since we assume $M_2 = \gamma(\mu_A - \mu_P)$, then we have $\gamma = \frac{r}{1+r}$. As we know, $r \leq 1$, then

$$1 - \gamma = \frac{1}{1+r} \Rightarrow \frac{1}{1-\gamma} = 1 + r \Rightarrow \frac{\gamma}{1-\gamma} = r \leq 1 \Rightarrow \gamma \leq 1 - \gamma \Rightarrow 2\gamma \leq 1 \Rightarrow \gamma \leq \frac{1}{2}$$

In addition, we need to take variability into account. Suppose that we have the same assumption in the beginning of this section.

y_{Ai} : Response for active control agent for i^{th} subject,

$y_{Ai} \sim N(\mu_A, \sigma_A^2), i = 1, 2, \cdots, n_A$,

y_{Ti} : Response for test therapy for i^{th} subject,

$y_{Ti} \sim N(\mu_T, \sigma_T^2), i = 1, 2, \cdots, n_T$,

y_{Pi} : Response for placebo for i^{th} subject,

$y_{Pi} \sim N(\mu_P, \sigma_P^2), i = 1, 2, \cdots, n_P$,

Then

$\overline{y_{T\cdot}} = \frac{1}{n_T}\sum_1^{n_T} y_{Ti}$,

$\overline{y_{P\cdot}} = \frac{1}{n_P}\sum_1^{n_P} y_{Pi\cdot}$,,

$E(\overline{y_{T\cdot}} - \overline{Y_{P\cdot}}) = \mu_T - \mu_P$, and

$VAR_{T-P}(\overline{y_{T\cdot}} - \overline{Y_{P\cdot}}) = \frac{\sigma_T^2}{n_T} + \frac{\sigma_P^2}{n_P}$,

So that

$$\widehat{\mu_T} \sim N\left(\mu_T, \frac{\sigma_T^2}{n_T}\right) \text{ and } \widehat{\mu_P} \sim N\left(\mu_P, \frac{\sigma_P^2}{n_P}\right).$$

Using independence, we have $\left(\widehat{\mu_T} - \widehat{\mu_p}\right) \sim N\left(\mu_T - \mu_p, \frac{\sigma_T^2}{n_T} + \frac{\sigma_P^2}{n_P}\right)$. When $\mu_T = \mu_A - M_2$, according to the properties of standard normal distribution,

$$P\left(\widehat{\mu_T} - \widehat{\mu_p} < M_1\right) = P\left(Z < \frac{M_1 - (\mu_T - \mu_p)}{\sqrt{\frac{\sigma_T^2}{n_T} + \frac{\sigma_P^2}{n_P}}}\right) = P\left(Z < \frac{M_1 - \mu_A + M_2 + \mu_p}{\sqrt{\frac{\sigma_T^2}{n_T} + \frac{\sigma_P^2}{n_P}}}\right)$$

$$= P\left(Z < \frac{M_1 + M_2 - (\mu_A - \mu_p)}{SE_{T-P}}\right).$$

We also have $M_2 = \frac{r}{1+r}(\mu_A - \mu_p)$. Therefore,

$$M_1 + M_2 - (\mu_A - \mu_p) = M_1 + \frac{r}{1+r}(\mu_A - \mu_p) - (\mu_A - \mu_p)$$

$$= M_1 + \frac{r}{1+r}(\mu_A - \mu_p) - \frac{1+r}{1+r}(\mu_A - \mu_p)$$

$$= M_1 - \frac{1}{1+r}(\mu_A - \mu_p).$$

In addition, $M_1 = \frac{M_2}{r}$ and $M_2 = \frac{r}{1+r}(\mu_A - \mu_p)$, so that

$$M_1 - \frac{1}{1+r}(\mu_A - \mu_p) = \frac{M_2}{r} - \frac{1}{1+r}(\mu_A - \mu_p) = \frac{1}{1+r}(\mu_A - \mu_p) - \frac{1}{1+r}(\mu_A - \mu_p) = 0.$$

As a result, $M_1 + M_2 - (\mu_A - \mu_p) = 0$. When $\mu_T = \mu_P - M_2$, according to the definition of standard normal distribution,

$$P\left(\widehat{\mu_T} - \widehat{\mu_p} < M_1\right) \le P\left(Z < \frac{M_1 + M_2 - (\mu_A - \mu_p)}{SE_{T-P}}\right) \le P(Z < 0) = \frac{1}{2}.$$

Now, we have:

$$P\left(\widehat{\mu_T} - \widehat{\mu_p} < M_1\right) \le P\left(Z < \frac{M_1 + M_2 - (\mu_A - \mu_p)}{SE_{T-P}}\right) \le P(Z < 0) = \frac{1}{2}$$

$$\Rightarrow \frac{M_1 + M_2 - (\mu_A - \mu_p)}{SE_{T-P}} = -Z_{1-\varepsilon} \text{ (assume the Z value)}$$

$$\Rightarrow M_1 + M_2 - (\mu_A - \mu_p) = (-Z_{1-\varepsilon})SE_{T-P}$$

$$\Rightarrow M_2 = (\mu_A - \mu_p) - M_1 - Z_{1-\varepsilon}SE_{T-P}$$

Again, using the equation $M_1 = \frac{M_2}{r}$.

$$M_2 = (\mu_A - \mu_p) - M_1 - Z_{1-\varepsilon}SE_{T-P}$$

$$\Rightarrow M_2 = (\mu_A - \mu_p) - \frac{M_2}{r} - Z_{1-\varepsilon}SE_{T-P}$$

$$\Rightarrow \frac{r+1}{r} M_2 = (\mu_A - \mu_p) - Z_{1-\varepsilon}SE_{T-P}$$

$$\Rightarrow M_2 = \frac{r}{r+1}\left[(\mu_A - \mu_p) - Z_{1-\varepsilon}SE_{T-P}\right]$$

In conclusion, the non-inferiority margin $M_2 = \frac{r}{r+1}\left[(\mu_A - \mu_p) - Z_{1-\varepsilon}SE_{T-P}\right]$. Finally,

$$H_0: \mu_T - \mu_P \le M_1$$

$$H_1: \mu_T - \mu_P > M_1.$$

Suppose the type II power $1 - \beta$ is the probability correctly rejecting the null hypothesis when the null hypothesis is not true, then the power of the test is approximately equal to

$$\Phi\left(\frac{\mu_T - \mu_P - M_1}{SE_{T-P}} - Z_{1-\alpha}\right).$$

Suppose $\mu_T = \mu_A - M_2$, then $\frac{\mu_A - \mu_P - M_1 - M_2}{SE_{T-P}} - Z_{1-\alpha} = Z_\beta$ (Chow & Shao, 2005). Therefore,

$$M_2 + M_1 = (\mu_A - \mu_p) - (Z_{1-\alpha} + Z_\beta)SE_{T-P}.$$

Also, $M_1 = \frac{M_2}{r}$, then

$$M_2 + \frac{M_2}{r} = (\mu_A - \mu_p) - (Z_{1-\alpha} + Z_\beta)SE_{T-P} \text{ , then}$$

$M_2 \frac{r+1}{r} = (\mu_A - \mu_p) - (Z_{1-\alpha} + Z_\beta) SE_{T-P}$, then

$M_2 = \frac{r}{1+r}[(\mu_A - \mu_p) - (Z_{1-\alpha} + Z_\beta) SE_{T-P}]$.

Comparing this form to the non-inferiority margin we derived, we can see that

$$Z_{1-\varepsilon} = Z_{1-\alpha} + Z_\beta.$$

Selecting Margin without Historical Data

Although a clinical trial with test therapy, active control, and placebo is usually recommended for a non-inferiority trial (called a three-arm trial), it is very common that researchers have to conduct a clinical trial without placebo when use of the placebo is unethical. It is also highly possible that the trial does not have access to historical data since it does not have placebo group. In general situations, we can estimate the non-inferiority margin $-M_2$ based on the value of the superiority margin M_1 as FDA recommends. However, that process does not work anymore, since the superiority margin M_1 depends on historical data. Here, Chow and Shao (2005) provide an alternative solution to estimate non-inferiority margin.

In this case, we usually know that the active control agent is the proven treatment. In other words, the active control agent is 'better' than the placebo. Applying the concept from Chapter 1, we know that active control agent is superior to placebo. Therefore, we can assume that the power of the level α test showing that the active control agent is superior to placebo by the margin M_1 is at the level η:

$$M_2 = (Z_{1-\alpha} + Z_\eta) SE_{A-P} - Z_{1-\varepsilon} SE_{T-P}.$$

In order to specify the non-inferiority margin, we are required to get some statistical information of the placebo group, such as the population variances.

$$M_2 = \left(Z_{1-\alpha} + Z_{\eta}\right)\sqrt{\frac{{\sigma_A}^2}{n_A} + c^2} - Z_{1-\varepsilon}\sqrt{\frac{{\sigma_T}^2}{n_T} + c^2}$$

We replace *c* to the smaller estimated value of $\frac{\sigma_T}{\sqrt{n_T}}$ and $\frac{\sigma_A}{\sqrt{n_A}}$ (Chow & Shao, 2005). Since $SE_{A-P} = \sqrt{\frac{{\sigma_A}^2}{n_A} + \frac{{\sigma_P}^2}{n_P}}$ and $SE_{T-P} = \sqrt{\frac{{\sigma_T}^2}{n_T} + \frac{{\sigma_P}^2}{n_P}}$, we use c^2 to take place of $\frac{{\sigma_P}^2}{n_P}$. Suppose we have a three arm non-inferiority trial. One group takes new treatment. One group takes proven treatment. One group takes placebo. According to the definition of placebo, $\frac{{\sigma_P}^2}{n_P}$ should be very small but not equal to 0. Therefore, it is logical to use the smaller estimated value of $\frac{\sigma_T}{\sqrt{n_T}}$ and $\frac{\sigma_A}{\sqrt{n_A}}$ to replace $\frac{\sigma_P}{\sqrt{n_P}}$.

Chapter 3 - A Case Study

Statistical Analysis

In this chapter, we estimate a two-arm non-inferiority margin for a trial, which is related to a blood pressure study. The data, from Wellek (2010) is provided in Table 3.1. The calculation in this chapter is processed in SAS. The SAS code and output are provided in Appendix A. This clinical trial only has two treatments: moxonidin and captopril. Moxonidin is a proven treatment for blood pressure. In contrast, captopril is a new treatment. In this trial, researchers measure the reduction of blood pressure (mm Hg) after 4 weeks of treatment. A group of patients takes 0.2 to 0.4 mg moxonidin every day during the treatment period. The other patients take 25 to 50 mg captopril daily over the treatment period. Assume that we do not have historical data for estimating the non-inferiority margin. Also, we know that it is a two-arm non-inferiority trial with active control (moxonidin) and therapy test (captopril).

To estimate the non-inferiority margin, we can use the following result from Chapter 2:

$$M_2 = \left(Z_{1-\alpha} + Z_{\eta}\right)\sqrt{\frac{{\sigma_A}^2}{n_A} + c^2} - Z_{1-\varepsilon}\sqrt{\frac{{\sigma_T}^2}{n_T} + c^2}$$

where c is the smaller estimated value of $\frac{\sigma_T}{\sqrt{n_T}}$ and $\frac{\sigma_A}{\sqrt{n_A}}$. Before we calculate the non-inferiority margin, we need to test if the data fit the assumption of non-inferiority trials. It is clear that the sample size is very small in this case (Table 3.1) and hence, may be an issue for this study.

Looking at the results from SAS output, we can know some basic statistics of the data set. The sample sizes of both group is 12. In the moxonidin group, the mean is 4.175, the variance is 49.704, and the standard deviation is 7.050 (Table 3.2). In the captopril group, the mean is 7.208, the variance is 43.886, and the standard deviation is 6.625 (Table 3.3).

In addition, we need to test the normality of the data. For the moxonidin group, the data does not fail the normality test, since the Shapiro-Wilk statistic is 0.916 with the p-value = 0.254 (Table 3.4). Similarly, the captopril group does not fails the normality test, since the Shapiro-Wilk statistic is 0.933 with p-value = 0.413 (Table 3.5). Also, if we look at the QQ plots (Figures 3.1 and 3.2), we can conclude that the data fits the assumption of normality.

In order to find the estimated non-inferiority margin, we need to give values for $\varepsilon, \alpha, \eta$. Let $\alpha = 0.05$ and $\eta = 0.80$, we get five different estimated margins based on five different values of ε in Table 3.6. It is clear that the value of the estimated margin drops as the value of ε decreases. In this case study, first, we can use *c*, which is the smaller estimated value of $\frac{\sigma_T}{\sqrt{n_T}}$ and $\frac{\sigma_A}{\sqrt{n_A}}$ to test whether captopril is non-inferior to moxonidin.

According to the result from the R program, the 90% confidence interval of the difference between moxonidin group and captopril group is (−1.763, 7.830). We can see that −1.763, which is the lower bound of the 90% CI, is greater than any of the estimated non-inferiority margins $-M_2$. As a result, we can conclude that captopril is non-inferior to moxonidin.

Tables and Figures

Table 3.1. Reduction of Blood Pressure in Moxonidin Group and Captopril Group

Moxonidin		Captopril	
i	X_i	j	Y_j
1	10.3	1	3.3
2	11.3	2	17.7
3	2.0	3	6.7
4	-6.1	4	11.1
5	6.2	5	-5.8
6	6.8	6	6.9
7	3.7	7	5.8
8	-3.3	8	3.0
9	-3.6	9	6.0
10	-3.5	10	3.5
11	13.7	11	18.7
12	12.6	12	9.6

Table 3.2. Basic Statistical Measures of Moxonidin

Basic Statistical Measures			
Location		Variability	
Mean	4.175	Std Deviation	7.050
Median	4.950	Variance	49.704
Mode	.	Range	19.800
		Interquartile Rang	14.200

Table 3.3. Basic Statistical Measures of Captopril

Basic Statistical Measures			
Location		Variability	
Mean	7.208	Std Deviation	6.625
Media	6.350	Variance	43.886
Mode	.	Range	24.500
		Interquartile Range	6.950

Table 3.4. Test for Normality of Moxonidin

Tests for Normality				
Test	Statistic		P Value	
Shapiro-Wilk	W	0.916	Pr<W	0.254
Kolmogorov-Smirnov	D	0.189	Pr>D	>0.150
Cramer-von Mises	W-Sq	0.551	Pr>W-Sq	>0.250
Anderson-Darling	A-Sq	0.396	Pr>A-Sq	>0.250

Table 3.5. Test for Normality of Captopril

Test for Normality				
Test	Statistic		P Value	
Shapiro-Wilk	W	0.934	Pr<W	0.413
Kolmogorov-Smirnov	D	0.185	Pr>D	>0.150
Cramer-von Mises	W-Sq	0.074	Pr>W-Sq	0.235
Anderson-Darling	A-Sq	0.443	Pr>A-Sq	0.242

Table 3.6. Estimated Non-inferiority Margin When α=0.05 and β=0.8

ε	0.25	0.20	0.15	0.10	0.05
$-M_2$	-5.120	-4.668	-4.141	-3.478	-2.495

Figure 3.1. Q-Q Plot for Moxonidin

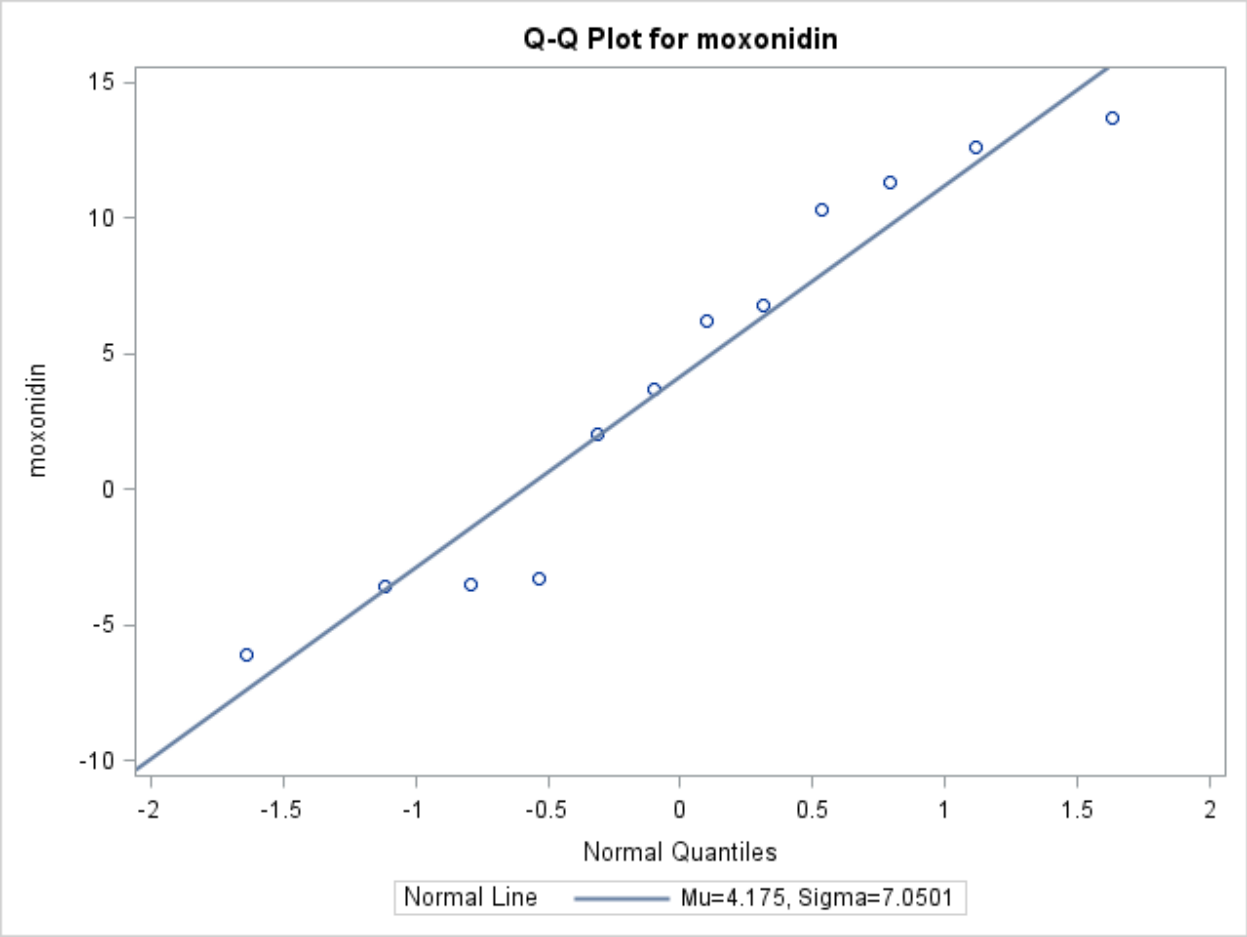
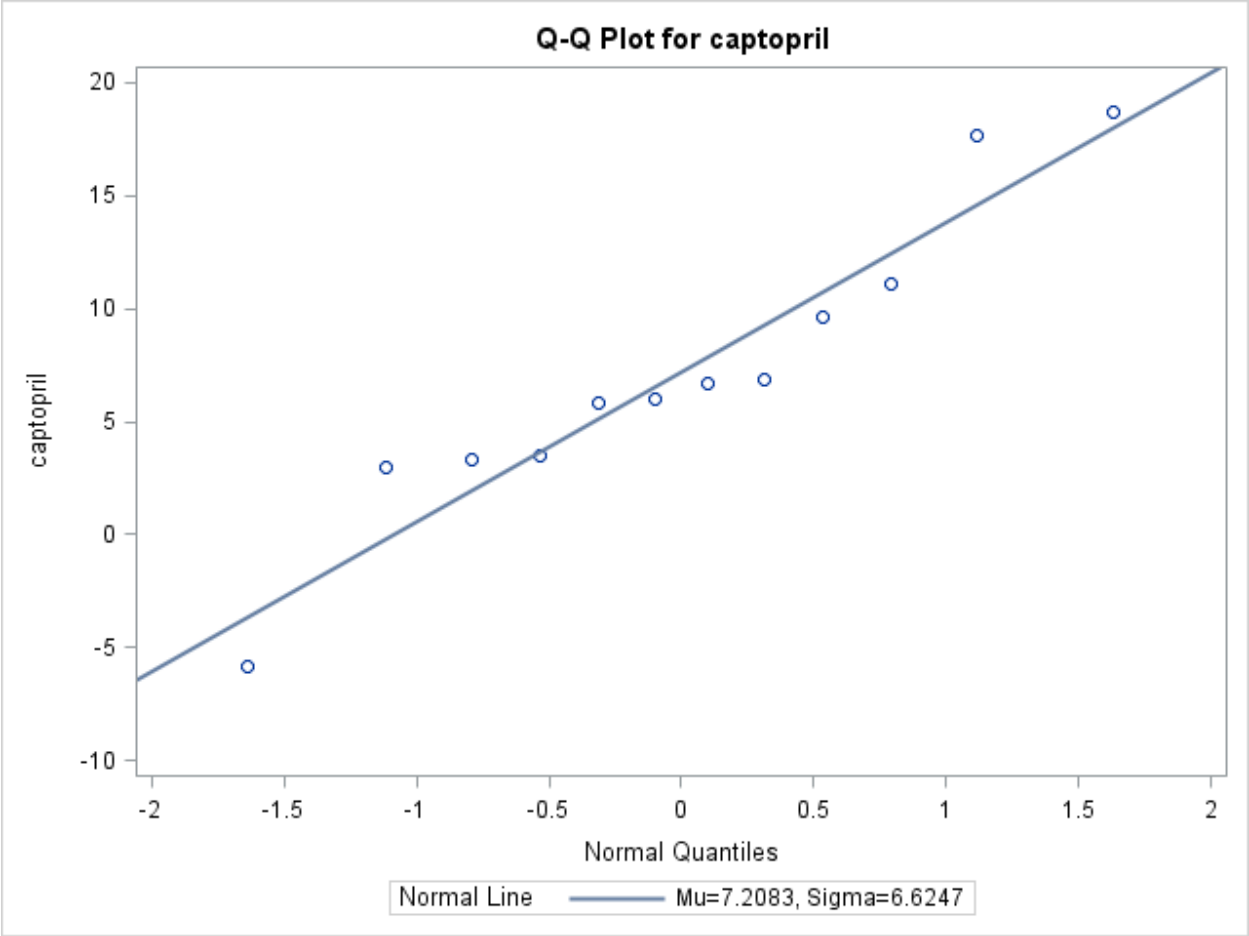


Figure 3.2. Q-Q Plot for Captopril



Chapter 4 - Summary

In this report, we introduced the general idea of non-inferiority trials. There is evidence indicating that non-inferiority trails are becoming popular in the field of clinical research. We also know that non-inferiority trials are different from the other two major types of clinical trials, namely superiority and equivalence trials. It is clear that non-inferiority trials can reduce risk to patients because the patients none have to receive a placebo treatment.

However, there are several challenges in the study of non-inferiority trials. We must take into consideration assay sensitivity, sample size, constancy condition, and non-inferiority margin, although the major focus of this report is the non-inferiority margin. We provided the formula for specifying a non-inferiority margin by using the guideline recommended by the International Conference on Harmonization.

In the end, we used a case study related to blood pressure to implement the theoretical results for selecting the margin of a non-inferiority trial. The results showed that captopril is non-inferior to moxonidin. We also note that, while not included in the actual trial, the standard deviation of the placebo group plays a very important role in specifying the hypotheses of a non-inferiority trial.

References

Chow, S., & Shao, J. (2005). On non-inferiority margin and statistical tests in active control trials. *Statistics in Medicine,* 1101-1113.

EMA. (2005). Committee for medicinal products for human use (CHMP) guideline on the choice of the non-inferiority margin. 1628-1638. Retrieved November 19, 2014, from http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003636.pdf

Freedman, B. (1987). Equipoise and the Ethics of Clinical Research. *New England Journal of Medicine, 317*(3), 141-145.

FDA. (2010). Guidance for industry non inferiority clinical trials (draft guidance). 1-55.

Hurley, A., & McKibbin, T., & Moroney S., & Suda, K. (2011). Pulication of noninferiority clinical trials: change over a 20-years interval. *Pharmacotherapy*, 833-839.

ICH. (2001). Guidance for Industry E 10 Choice of Control Group and Related Issues in Clinical Trials. Retrieved January 1, 2014, from http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm073139.pdf

Rothmann, M.D., Wiens, B.L., & Chan, I.S.F. (2012). *Design and Analysis of Non-Inferiority Trials* (pp. 1-6). Hoboken: Chapman & Hall/CRC.

Schumi, J., & Wittes, J. (2011). Through the looking glass: Understanding non-inferiority. *Trials,* 106-106.

Snapinn, S. (2000). Commentary noninferiority trials. *Curr Control Trials Cardiovasc Med,* 19-21. Retrieved November 19, 2014, from http://cvm.controlled-trials/content/1/1/019

Wellek, S. (2010). Equivalence tests for two unrelated samples. In *Testing statistical hypotheses of equivalence and noninferiority* (2nd ed., p. 123). Boca Raton: CRC Press.

Appendix A – Case Study

SAS and R Code for Case Study

```
data moxonidin;
input moxonidin;
cards;
10.3
11.3
2.0
-6.1
6.2
6.8
3.7
-3.3
-3.6
-3.5
13.7
12.6
;

data captopril;
input captopril;
cards;
3.3
17.7
6.7
11.1
-5.8
6.9
5.8
3.0
6.0
3.5
18.7
9.6
;

proc univariate data=moxonidin normal;
qqplot moxonidin / normal(mu=est sigma=est color=red l=1);
run;

proc univariate data=captopril normal;
qqplot captopril / normal(mu=est sigma=est color=red l=1);
run;
```

```
proc iml;
moxonidin={10.3,11.3,2.0,-6.1,6.2,6.8,3.7,-3.3,-3.6,-3.5,13.7,12.6};
captopril={3.3,17.7,6.7,11.1,-5.8,6.9,5.8,3.0,6.0,3.5,18.7,9.6};
moxonidin_u=mean(moxonidin);
captopril_u=mean(captopril);
u=captopril_u-moxonidin_u;
print u;
moxonidin_var=var(moxonidin);
captopril_var=var(captopril);
c2=min(moxonidin_var/12,captopril_var/12);
c=sqrt(c2);
print c;
margin1=(probit(1-0.05)+probit(0.8))*sqrt((moxonidin_var/12)+c2)-probit(1-
0.25)*sqrt((captopril_var/12)+c2);
margin2=(probit(1-0.05)+probit(0.8))*sqrt((moxonidin_var/12)+c2)-probit(1-
0.2)*sqrt((captopril_var/12)+c2);
margin3=(probit(1-0.05)+probit(0.8))*sqrt((moxonidin_var/12)+c2)-probit(1-
0.15)*sqrt((captopril_var/12)+c2);
margin4=(probit(1-0.05)+probit(0.8))*sqrt((moxonidin_var/12)+c2)-probit(1-
0.1)*sqrt((captopril_var/12)+c2);
margin5=(probit(1-0.05)+probit(0.8))*sqrt((moxonidin_var/12)+c2)-probit(1-
0.05)*sqrt((captopril_var/12)+c2);
print margin1 margin2 margin3 margin4 margin5;
w1=u+(probit(1-0.05)+probit(0.8))*sqrt((moxonidin_var/12)+c2)-probit(1-
0.25)*sqrt((captopril_var/12)+c2)
-probit(1-0.05)*sqrt((moxonidin_var/12)+(captopril_var/12));
w2=u+(probit(1-0.05)+probit(0.8))*sqrt((moxonidin_var/12)+c2)-probit(1-
0.2)*sqrt((captopril_var/12)+c2)
-probit(1-0.05)*sqrt((moxonidin_var/12)+(captopril_var/12));
w3=u+(probit(1-0.05)+probit(0.8))*sqrt((moxonidin_var/12)+c2)-probit(1-
0.15)*sqrt((captopril_var/12)+c2)
-probit(1-0.05)*sqrt((moxonidin_var/12)+(captopril_var/12));
w4=u+(probit(1-0.05)+probit(0.8))*sqrt((moxonidin_var/12)+c2)-probit(1-
0.1)*sqrt((captopril_var/12)+c2)
-probit(1-0.05)*sqrt((moxonidin_var/12)+(captopril_var/12));
w5=u+(probit(1-0.05)+probit(0.8))*sqrt((moxonidin_var/12)+c2)-probit(1-
0.05)*sqrt((captopril_var/12)+c2)
-probit(1-0.05)*sqrt((moxonidin_var/12)+(captopril_var/12));
print w1 w2 w3 w4 w5;
w6=u+(probit(1-0.05)+probit(0.8))*sqrt((moxonidin_var/12)+1*1)-probit(1-
0.05)*sqrt((captopril_var/12)+1*1)
-probit(1-0.05)*sqrt((moxonidin_var/12)+(captopril_var/12));
w7=u+(probit(1-0.05)+probit(0.8))*sqrt((moxonidin_var/12)+2*2)-probit(1-
0.05)*sqrt((captopril_var/12)+2*2)
-probit(1-0.05)*sqrt((moxonidin_var/12)+(captopril_var/12));
```

```
w8=u+(probit(1-0.05)+probit(0.8))*sqrt((moxonidin_var/12)+3*3)-probit(1-
0.05)*sqrt((captopril_var/12)+3*3)
-probit(1-0.05)*sqrt((moxonidin_var/12)+(captopril_var/12));
w9=u+(probit(1-0.05)+probit(0.8))*sqrt((moxonidin_var/12)+4*4)-probit(1-
0.05)*sqrt((captopril_var/12)+4*4)
-probit(1-0.05)*sqrt((moxonidin_var/12)+(captopril_var/12));
w10=u+(probit(1-0.05)+probit(0.8))*sqrt((moxonidin_var/12)+5*5)-probit(1-
0.05)*sqrt((captopril_var/12)+5*5)
-probit(1-0.05)*sqrt((moxonidin_var/12)+(captopril_var/12));
print w6 w7 w8 w9 w10;
run;

moxonidin<-c(10.3,11.3,2.0,-6.1,6.2,6.8,3.7,-3.3,-3.6,-3.5,13.7,12.6)
captopril<-c(3.3,17.7,6.7,11.1,-5.8,6.9,5.8,3.0,6.0,3.5,18.7,9.6)
t.test(captopril,moxonidin,conf.level=0.9)
```

Tables

Table A.1. Moments of Moxonidin

Moments			
N	12	Sum Weights	12
Mean	4.175	Sum Observation	50.1
Std Deviation	7.050	Variance	49.704
Skewness	-0.110	Kurtosis	-1.563
Uncorrected SS	755.91	Corrected SS	546.743
Coeff Variation	168.8646	Std Error Mean	2.035

Table A.2. Test for Location of Moxonidin

Tests for Location: Mu0=0				
Test	Statistic		P Value	
Student's t	t	2.051	Pr>|t|	0.065
Sign	M	2	Pr>=|M|	0.388
Singed Rank	S	24	Pr>=|S|	0.064

Table A.3. Quantiles of Moxonidin

Quantiles (Definition 5)	
Quantile	Estimate
100% Max	13.70
99%	13.70
95%	13.70
90%	12.60
75% Q3	10.80
50% Median	4.95
25% Q1	-3.40
10%	-3.60
5%	-6.10
1%	-6.10
0% Min	-6.10

Table A.4. Extreme Observations of Moxonidin

Extreme Observations			
Lowest		Highest	
Value	Ob	Value	Ob
-6.1	4	6.8	6
-3.6	9	10.3	1
-3.5	10	11.3	2
-3.3	8	12.6	12
2.0	3	13.7	11

Table A.5. Moments of Captopril

Moments			
N	12	Sum Weights	12
Mean	7.208	Sum Observations	86.5
Std Deviation	6.625	Variance	43.886
Skewness	0.133	Kurtosis	0.810
Uncorrected SS	1106.27	Corrected SS	482.749
Coeff Variation	91.903	Std Error Mean	1.912

Table A.6. Tests for Loation of Captopril

Tests for Location: Mu0=0				
Test	Statistic		P Value	
Student's t	t	3.769	Pr>|t|	0.003
Sign	M	5	Pr>=|M|	0.006
Signed Rank	S	34.5	Pr>=|S|	0.004

Table A.7. Quantiles of Captopril

Quantiles (Definition 5)	
Quantile	Estimate
100% Max	18.70
99%	18.70
95%	18.70
90%	17.70
75% Q3	10.35
50% Median	6.35
25% Q1	3.40
10%	3.00
5%	-5.80
1%	-5.80
0% Min	-5.80

Table A.8. Extreme Observations of Captopril

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
-5.8	5	6.9	6
3.0	8	9.6	12
3.3	1	11.1	4
3.5	10	17.7	2
5.8	7	18.7	11