Determination of Sample Size and Similarity Criteria Using Regional Error Rate in a Multi-regional Clinical Trial

Sang Jun Youn

The Graduate School
Yonsei University
Department of Biostatistics and Computing

Determination of Sample Size and Similarity Criteria Using Regional Error Rate in a Multi-regional Clinical Trial

A Dissertation
Submitted to the Department of Biostatistics and Computing
and the Graduate School of Yonsei University
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Sang Jun Youn

December 2016

This certifies that the Dissertation of
Sang Jun Youn is approved.

Thesis Supervisor : Prof. Seung Ho Kang

Prof. Chung Mo Nam

Prof. Byung Ho Nam

Prof. Ki Jun Song

Prof. Inkyung Jung

The Graduate School
Yonsei University

December 2016

Contents

List of Tables

List of Figures

Abstracts

There is growing interest in multi-regional clinical trials (MRCTs), which are those conducted in many regions using the common protocol. MRCTs simplify the approval and registration processes for treatment in all regions, and provide an opportunity to reduce costs and time consumption by not repeating similar clinical trials. Accomplishing the goal of MRCTs depends on consistency in the effect size of individual regions following verification of overall treatment effects. However, there are currently no criteria to assess the similarity of treatment effects across regions, or standards for calculating the required number of clinical trial subjects in the region of interest to demonstrate such similarity. In 2007, Japanese MHLW provided guideline on similarity criteria and the required number of clinical trial subjects in Japan for MRCTs. However, this guideline does not offer a statistical perspective.

Based on the MHLW guideline, Ko et al. (2010) proposed a method based on the concept of the assurance probability to calculate the sample size in the region of interest. But this method does not focus on the second purpose of MRCTs, which concerns the similarity of treatment effects across regions. This thesis introduces a method standardized by effect size, which was originally suggested by Kang et al. (2016), as a statistical hypothesis testing procedure to address the second purpose of MRCTs. This thesis also discusses approaches using

the regional type II error rate to calculate critical values for a hypothesis testing on the similarity, as well as the required number of clinical trial subjects in the region of interest through the suggested method.

Using calculated regional type II error rates according to similarity criteria, it was shown that it is easier to control the regional type II error rate if the difference in effect sizes between the region of interest and other regions excluding the region of interest is great or the critical value for the similarity increases. If a pre-determined regional type II error rate is satisfied and parameters such as effect size and the required number of patients in the region of interest are the same respectively, the critical value for the similarity criterion $D_1 \geq \rho D$ is considered more conservative compared to the critical value for the similarity criterion $D_1 \geq \rho D_{1c}$. The proportion of the patients in the region of interest did not monotonically decrease with the regional type II error rate. Such an undesirable property needs to be improved by developing new similarity criteria. Furthermore, the method used to control the regional type I error rate is an expanded form of the method proposed by Ko et al. (2010), in the sense that it is applicable in cases where the effect sizes across regions are heterogeneous.

Recently, there has been an increasing trend of globalization in developing new drugs using MRCTs in Korea. As the frequency of conducting MRCTs increases, it has become important to determine

critical value to evaluate the similarity in treatment effects between ethnic groups and to calculate the sample size in the region of interest, such as Korea. In such clinical development environments, the regional type II error method proposed in this thesis will be useful for MRCTs, and should be further studied for continued improvement.

Key words : Multi-regional clinical trial, Similarity criteria, Critical value, Sample size determination, Assurance probability, Secondary hypothesis, Regional error rate, Regional type I error, Regional type II error

Chapter 1. Introduction

Korea has a short history of new drug development compared to economically developed countries such as the United States (US), European countries, and Japan. Therefore, the infrastructure for clinical trials was inadequate compared to such countries until the 1990s. However, globalization of new drug clinical trials since 2000 increased the number of clinical trials in Korea to 202 cases in 2009. Korea is ranked as the 12th country and Seoul the 3rd city in clinical trial participation in the world (Bae, 2010).

According th the '2015 Clinical trial Protocol Approval Status and Annual Inspection Result Presentation' published in 2016 by the Ministry of Food and Drug Safety, the approved clinical trials in Korea is continuously increasing from 503 cases in 2011 to 652 cases in 2014. In addition, the number of clinical trials including multi-regional clinical trials (MRCTs) in 2015 was 675 cases, which is a 3.5% annual increase compared to 652 cases in 2014. A comparison of Korean and multi-national pharmaceutical companies showed a greater number of approvals for multi-national companies by 55% to 45% in 2015. Confirmatory clinical trials, phase III clinical trials, accounted for 58% of approved clinical trials in multi-national pharmaceutical companies (MFDS, 2016). Due to the increase in drug clinical trials in Korea and globalization of development of drugs, it has become essential to establish developing strategies to explain variances in the intrinsic

(e.g.; genetic, physiological) and extrinsic (e.g.: medical practice, cultural and environmental) characteristic of participating regions on the efficacy and safety of developing products.

In 1998, the International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH) announced a guideline entitled “Ethnic Factors in the Acceptability of Foreign Clinical Data". The ICH E5 guideline proposed a bridging study to produce additional information for utilization of foreign clinical data when this data fails to provide sufficient bridging evidence. Korea acted upon the ICH E5 guideline until the mid-2000s and utilized a bridging strategy that required a bridging study for new foreign drug approval, in order to identify intrinsic and extrinsic difference between ethnicities exposed to the drug. The ICH guideline defines a bridging study as follow: "Bridging study is a supplementary study conducted to allow extrapolation of foreign clinical data on the population of the new region in order to provide pharmacodynamic or clinical data on efficacy, safety, dosage and dose regimen of the drug". However, a bridging strategy is an additional strategy conducted by the country of interest, which involves a bridging study with the aim to utilize all clinical data from the country of origin. The strategy is conducted on products after the required phase I, II and III clinical trials and the drug approval are completed in the country of origin. Although the bridging study, conducted for a bridging strategy as suggested by ICH E5 guideline, provides regulatory strategies to minimize overlap in clinical trials, the need for valuable resources for duplication of large

clinical trials in all regions is inevitable. Furthermore, the bridging study conducted for bridging strategy has an issue similar to the phenomenon known as "Drug lag". This describes the fact that a new country cannot release a drug until a few years after its release in countries including the country of origin, thus reducing the product lifetime and delaying the provision of new treatments to patients in that new country.

Recently, many global pharmaceutical companies have been involved in drug development through MRCTs to overcome the problems discussed above. In a study by Ando et al. (2010) in which cases of MRCTs were introduced, the trial includes various study design with different countries and ethnicities. MRCTs conducted with bridging purposes in the context of a global development program should not only simplify the approval and registration of new drugs in all regions but also provide an opportunity to reduce time consumption and costs from conducting repetitive clinical trials. Such merit has led to continuous interest in MRCTs. The 11th ICH E5 Q&A describes an MRCT as a study that uses a common protocol in more than one region for bridging that allows near simultaneous worldwide registration and thus can be conducted in the context of a global development program.

The objective of MRCTs should therefore be to show that the drug is effective in the region and to compare the results of the study between the regions with the intent of establishing that the drug is not sensitive to ethnic factors. A closer look at the statements in the

ICH E5 guideline states that in an MRCT planning, sufficient numbers of patients, with adequate power to show treatment effect should be registered in each region. The guideline further emphasize that the ability of MRCTs to achieve their purpose depends on "the extent of similarity in drug effects between individual participating regions". However, there are no detailed suggestions on definitions for treatment effects such as "not sensitive", "similar" or "consistent", or sample size determination of participating individual regions. These depend on scientific aspects and regulatory requirements which may vary from region to region.

It could be argued that the design and analysis of MRCTs are similar to those of multi-center clinical trials. The reason may be that in both cases data are collected from multiple units, such as centers or regions, and intended to be analyzed as a whole, where heterogeneity of the treatment effect across units may exist. However, the crucial difference is that in multi-center clinical trials one regulatory agency reviews the material and decides on market release in the country, whereas in MRCTs, there is no such thing as "Global approval". In other words, the significance of the overall results does not guarantee market approval in each country. Therefore, in extreme cases, the same MRCT material can be interpreted differently by regulatory bodies of each region to arrive at different conclusions. In general, in order to investigate the existence of heterogeneity according to center or region, a method to confirm the existence of treatment-by-center or treatment-by-region interaction is used.

According to Chen et al. (2010), in recently published MRCTs, treatment-by-region interaction tests are commonly used to assess heterogeneous treatment effect across regions and a non-significant interaction test would lead to the conclusion that the treatment effect is consistent across regions.

If individual center results in cases that use interaction tests show extreme or opposite results between centers, the results should be discussed. The Q&A for the ICH E9 guideline states that there should be at least 10 patients in each center. Shao and Chow (1993) argued that this number should not be less that the number of centers. These proposals focus on investigation treatment-by-center interactions and the resulting sample sizes may not be adequate to show the efficacy of a new drug in an individual region, as required in an MRCT. In addition, to evaluate the interaction effects, a very large sample size is needed, and the numbers involved can make it unrealistic to conduct a trial.

There are currently no statistical criteria provided by the ICH E5 guideline to assess the treatment similarity or consistency between overall clinical data and data from the region of interest. Moreover, until 2007 no regulatory agency in any country provided guidelines on the required number of patients for allocation in the region of interest to assess the similarity in treatment effects across regions. The first standard to assess such similarities was suggested in September 2007 by the Japanese Ministry of Health, Labor, and Welfare (MHLW) through the publication of a guideline called "The Basic Principles on

Global Clinical Trials". This guideline, similar to ICH E5, describes MRCTs as planned research conducted in medical institutions in numerous countries with a common protocol for the purpose of the development and approval of new drugs. This guideline provides detailed questions on consideration of the required number of Japanese patients, a description of basic concepts and an introduction for conducting MRCTs in a Q&A format.

The guideline also provides two methods to decide on the number of Japanese subjects required to obtain identical results from the participating Japanese patients compared to overall patient data in MRCTs. Let D be the observed difference between the effects of placebo and the drug in patient groups in all regions, whereas D_1 is the observed difference in the Japanese patient group. The first method in the guideline, 'Method 1' for shot, requires that the number of Japanese patients is sufficiently large to ensure $D_1/D > \rho$ with a probability of at least 80%, where ρ is a pre-specified threshold and $\rho \geq 0.5$ is generally recommended. The second method in the guideline, 'Method 2' for short, requires that the sample size is sufficiently large to demonstrate a consistent trend for all individual regions. For example, assume that three regions participated in a global clinical trial and let D_1, D_2 and D_3 denote the observed treatment differences. Then the number of clinical trial subjects is determined such that each individual difference D_1, D_2 and D_3 is larger than 0 with a probability of at least 80%.

Kawai et al. (2008) discussed Method 2 and proposed an approach to rationalize partitioning the total sample size among the regions so that a high probability of observing a consistent trend under the assumed treatment effect across regions can be derived if the treatment effect is positive and uniform across regions in a confirmatory MRCT. Their approach takes an overall perspective in that its main purpose is to estimate the overall treatment effect. On the other hand, Ko et al. (2010) discussed Method 1 taking a regional viewpoint and focused on estimating the treatment effect in a specific region. Specifically, they proposed a method for calculating sample size in the specific region in order to ensure that the assurance probabilities for similarity criteria under the alternative hypothesis for the primary purpose of MRCTs were maintained at a desired level, say, 80%. However, as mentioned above, MRCTs have two purposes. First, there is a need to prove significance of the new drug compared to placebo in the overall clinical trial subject group. Second, there is a need to prove treatment effects are the similarity in patients in the region of interest and the overall patient group. The two purposes result in different hypotheses. The first purpose is related to the overall treatment effects of the clinical trial and the primary hypothesis of an MRCT is established based on the first purpose. The second purpose is to assess consistency of the treatment effect across regions, resulting in the secondary hypothesis. Ko et al. (2010) evaluated the assurance probabilities on similarity criteria under the primary alternative hypothesis. The purpose of the assurance probabilities is to assess the

consistency of treatment effects across regions, following the evaluation of overall treatment effect results. Kang et al. (2016) emphasizes the fact that the assurance probabilities should be assessed via the secondary alternative hypothesis, not the first.

This thesis will focus on the two purposes of MRCTs to discuss the critical values for assessing the similarity and to calculating the sample size in the region of interest using Method 1. Using the method by Kang et al. (2016), a standardized equation on effect size will be suggested and a comparison with the assurance probability by Ko et al. (2010) will be conducted. If the difference in treatment effects between the region of interest and other regions excluding the region of interest is predicted to be significant in the design stage, the similarity criteria of Method 1 cannot be used and an independent clinical trial needs to be conducted on the region of interest. However, such difference in effects cannot be distinctively defined and thus in the design stage of this study, both cases with significant and insignificant differences in treatment effects between patients in the region of interest and other regions excluding the region of interest will be included. This will be done using the similarity criteria discussed above for calculating the required number of clinical trial subjects. For the assumed treatment effects, results from the US Food and Drug Administration (FDA) Statistical Review on approved drugs from 2011 to 2014 will be used.

The contents of this thesis are as follows. Chapter 2 will discuss the similarity criteria proposed by Japan's MHLW, which can be evaluated

to consider similar treatment effects between the region of interest and all regions in MRCTs, as well as the similarity criteria by Ko et al. (2010). Chapter 3 will examine the assurance probabilities proposed by Ko et al. (2010). In chapter 4, we will examine the assurance probabilities using the regional type II error rate as suggested by Kang et al. (2016) and propose a standardized method. A literature review on effect size will be suggested to standardize and utilize the method suggested by Kang et al. (2016). Chapter 5 will discuss the calculated results of regional type II error rates and regional type I error rates according to the effect size. We will also discuss the method for choosing critical values ρ and p_1 according to the effect size and results. Chapters 6 and 7 will discuss examples in real clinical trial conditions, followed by a discussion and conclusion.

Chapter 2. Similarity criteria

The first purpose of MRCTs is to determine the treatment effects in all participating regions, and the second purpose is to assess applicability of the overall clinical trial results to each region. In order to assess the second purpose, similarity criteria to confirm applicability of the trial results in the region of interest is important. This section will discuss such similarity criteria. For the sake of simplicity, this thesis only deals with a phase III confirmatory clinical trial that utilizes parallel group design to compare a new treatment group and placebo group.

The notation used in this thesis to explain the similarity criteria will be described. Continuous primary endpoints that represent efficacy results are denoted as X and Y, each in regard to patients given new treatment and placebo, respectively. The greater the difference in the primary endpoints between the new treatment and placebo groups will be defined to have greater treatment effects.

$$X \sim N(\mu_T , \sigma^2) \qquad Y \sim N(\mu_P , \sigma^2)$$

where the population variance σ^2 is assumed to be known, although σ^2 is actually unknown and must be estimated from actual clinical trial data.

Let μ_T and μ_P be the population means of the new treatment and placebo, respectively, and let the overall treatment difference be $\triangle = \mu_T - \mu_P$.

The hypothesis of the first purpose of the MRCT for testing the overall treatment effect is shown below.

$$H_0 : \triangle \leq 0 \quad vs. \quad H_A : \triangle > 0 \tag{1}$$

The primary hypothesis states that the new treatment is effective at a global level. Although the primary hypothesis is the one sided hypothesis to test the overall treatment effect, the methods proposed by Ko et al. (2010) and Kang et al. (2016) can be straightforwardly extended to the two-sided hypothesis.

Let $2N$ denote the total number of patients planned for the trial, divided equally between the new treatment group and the placebo control group (ie, equally allocation of subjects to the treatment group). If the total sample size for the each group N is for a one-sided test with the desired significant level α and power $1-\beta$ for detecting an expected overall difference $\triangle = \delta$, then the equation for the sample size calculation is as follows.

$$N = \frac{2(z_{1-\alpha} + z_{1-\beta})^2 \sigma^2}{\delta^2} \tag{2}$$

where $z_{1-\alpha}$ is the $(1-\alpha)$th percentile of the standard normal

distribution (Chow et al., 2003).

Let K be the number of regions participating in the MRCT and p_i refers to the proportion of patients in i[th] participating region from the total number of subjects $2N$, resulting in $i=1,2,\cdots,K$, $\sum_{i=1}^{K} p_i = 1$. Furthermore, let $n_i = p_i N$ be the number of patients assigned to each treatment group from the i[th] participating region. Random variables X_{ij} and Y_{ir} refer to the primary endpoints representing treatment effects in the j[th] or r[th] patient given the new treatment or the placebo from the i[th] participating region.

Without a loss of generality, assume that the region of interest is the first region and thus $i=1$ region. Let D_1 be the mean difference in observed effects between new treatment and placebo groups for the first region. Let D denote the mean difference in observed effects for all regions. Let D_{1c} be the mean difference in observed effects for regions other than the first region. Hence,

$$D_1 = \overline{X_{1.}} - \overline{Y_{1.}}, \quad D = \sum_{i=1}^{K}\sum_{j=1}^{n_i} \frac{(X_{ij} - Y_{ij})}{N}, \quad D_{1c} = \sum_{i=2}^{K}\sum_{j=1}^{n_i} \frac{(X_{ij} - Y_{ij})}{N - n_1} \qquad (3)$$

where

$$\overline{X_{1.}} = \frac{1}{n_1}\sum_{j=1}^{n_1} X_{1j}, \quad \overline{Y_{1.}} = \frac{1}{n_1}\sum_{j=1}^{n_1} Y_{1j}$$

In addition, Z_1 is the test statistic for the first region, Z is the test statistic for all regions and Z_{1c} is the test statistic for regions other

than the first region. That is,

$$Z_1 = \frac{\overline{X_{1.}} - \overline{Y_{1.}}}{\sigma\sqrt{\frac{1}{n_1}}}, \quad Z = \frac{\sum_{i=1}^{K}\sum_{j=1}^{n_i}(X_{ij} - Y_{ij})}{\sigma\sqrt{2N}}, \quad Z_{1c} = \frac{\sum_{i=2}^{K}\sum_{j=1}^{n_i}(X_{ij} - Y_{ij})}{\sigma\sqrt{2(N-n_1)}} \quad (4)$$

We will focus on the two similarity criteria suggested by Ko et al. (2010) in order to assess consistency of treatment effects under the condition that the overall treatment effect is significant at the significance level α.

(i) $D_1 \geq \rho D_{1c}$, for some $0 < \rho < 1$

(ii) $D_1 \geq \rho D$, for some $0 < \rho < 1$

Here, the first similarity criterion is that the extent of the treatment effect of the new treatment in the first region has to be similar to the treatment effect in regions other than the first region. The second criterion is that the extent of the treatment effect of the new treatment in the first region has to be similar to the treatment effect in all participating regions. In particular, Japanese MHLW suggested using the second criterion, so that the treatment effect between the first region (Japan) and all populations can be consistent when the ratio of treatment effect estimate for the first region to that for the overall population is greater than ρ, that is, $D_1 \geq \rho D$ with $\rho \geq 0.5$ (MHLW, 2007).

Chapter 3. Assurance probability

3.1. Assurance probability of similarity by Ko et al.

Ko et al. (2010) proposed a method to determine the proportion of patients out of $2N$ in the first region p_1 in the overall significance level α, using the assurance probability on similarity criteria (i) and (ii), when the expected difference in overall treatment effect is $\triangle = \delta$. When the overall treatment effect is $\triangle = \delta$, under the assumption that the treatment effect of the new treatment and placebo is uniform across regions, the assurance probability for the first similarity criterion (i) denoted as AP_1 can be expressed as below.

$$AP_1 = P_\delta(D_1 \geq \rho D_{1c} | Z > z_{1-\alpha}) \tag{5}$$

$$= \frac{\int_{a_1}^{\infty} (\Phi(\frac{u - c_2}{c_1}) - \Phi(\frac{c_5 - c_3 u}{c_4}))\phi(u)\,du}{1-\beta}$$

where P_δ is the probability measure with respect to $\triangle = \delta$ and

$$c_1 = \rho\sqrt{\frac{p_1}{1-p_1}},\ c2 = (\rho - 1)\sqrt{p_1}\,(z_{1-\alpha} + z_{1-\beta}),\quad c_3 = \sqrt{p_1},\ c_4 = \sqrt{1-p_1},$$

$$c_5 = -z_{1-\beta} \text{ and } a_1 = c_2 + \frac{c_1(c_5 - c_2 c_3)}{c_1 c_3 + c_4}$$

and α and β are the overall significance level and type II error rate, respectively.

Similar to AP_1, the assurance probability of the second similarity criterion (ii) denoted as AP_2 can be represented as below.

$$AP_2 = P_\delta(D_1 \geq \rho D \mid Z > z_{1-\alpha}) \tag{6}$$
$$= \frac{\int_{a_2}^{\infty} (\Phi(\frac{u - c_7}{c_6}) - \Phi(\frac{c_5 - c_3 u}{c_4}))\phi(u)\,du}{1-\beta}$$

where

$$c_6 = \frac{\rho\sqrt{p_1(1-p_1)}}{1-\rho p_1},\ c_7 = \frac{(1-\rho)\sqrt{p_1}}{1-\rho p_1}(z_{1-\alpha} + z_{1-\beta}),\ a_2 = c_7 + \frac{c_6(c_5 - c_7 c_3)}{c_6 c_3 + c_4}$$

In conclusion, the method suggested by Ko et al. (2010) determines p_1, the proportion of patients out of $2N$ in the first region, to ensure that the assurance probabilities for similarity criteria (i) and (ii) under the alternative hypothesis for the first purpose of MRCTs (From here on referred to as the primary alternative hypothesis) $H_A : \triangle = \delta$ are maintained at a desired level, say, 80%. Here, a higher assurance probability is better.

3.2. Limitations of Ko et al.'s method

In MRCTs, along with the first purpose of validating the treatment effect of developing drug in all participating regions, there is the second purpose to demonstrate that the effects of the developing drug

is not sensitive to the region. For the second purpose, consistency assessment can be conducted to demonstrate that the result is not sensitive to each region after hypothesis testing for the overall trial results.

The assurance probabilities method suggested by Ko et al. (2010) is inappropriate for evaluating the assurance probabilities under the primary alternative hypothesis which is the hypothesis for the first purpose of MRCTs. According to the method by Kang et al. (2016), as suggested in this thesis, the assurance probabilities should not be assessed based on the primary alternative hypothesis, but on the secondary alternative hypothesis for the second purpose of MRCTs with focus on assessing the similarity of treatment effects across regions. The reasons can be found in the concept of the assurance probabilities suggested by Ko et al. (2010). The purpose of the assurance probabilities is to assess consistency of treatment effects between the region of interest and all participating regions under the condition that the primary hypothesis is accepted. Specifically, the purpose is to confirm the assumption that all regions show an similar treatment effect. Therefore, such a purpose can be interpreted to be a process to test the following secondary hypothesis.

- H_{s0} refers to the secondary null hypothesis that claims the similarity in treatment effects between the region of interest and all participating regions.

- H_{sA} refers to the secondary alternative hypothesis that claims there is no similarity in treatment effects between the region of interest and all participating regions.

There are many ways to express the secondary hypothesis as a statistical hypothesis, but the generally used method is as written below.

H_{s0} : There is no difference in new treatment effects between regions.

$$vs$$

H_{sA} : There is a difference in new treatment effects between regions.

In reference to the aforementioned secondary hypothesis for the second purpose of MRCTs, the assurance probabilities suggested by Ko et al. (2010) can be considered a probability that satisfies the similarity criteria under H_{s0}. In the method suggested by Kang et al. (2016), as introduced in this thesis, the probability of not satisfying the similarity criteria under the secondary null hypothesis is referred to as a regional type I error rate. The assurance probabilities suggested by Ko et al. (2010), under the assumption that the treatment effects in the region of interest and other regions excluding the region of interest are identical, achieve exactly the same value as subtracting the regional type I error rate from 1. This is obtained through the method by Kang et al. (2016), which will be discussed in the next

section.

The regional type I error rate represents the probability of error to arrive at the incorrect conclusion of significant regional difference, despite the lack of difference in treatment effects between regions. On the other hand, the regional type II error represents the probability of error to arrive at the incorrect conclusion of no significance in regional difference, although there is a significant difference in treatment effects between regions. Therefore, when the hypothesis is tested about the second purpose of MRCTs , the regional type II error is more serious than the regional type I error.

To calculate the number of clinical trial subjects required for the region of interest using a method to control regional type I error, as in the method using the assurance probabilities by Ko et al. (2010), the secondary null hypothesis would be assumed to be true. In this case, since it is assumed that there is no difference in new treatment effects between regions, a contradiction arises, which is that there is no need to assign sufficient number of patients for the region of interest. Therefore, it is more important to control regional type II error than regional type I error to calculate the number of required clinical trial subjects for the region of interest under the secondary hypothesis and thus the method proposed by Ko et al. (2010) would not be appropriate.

Chapter 4. Regional error rate

4.1. Proposed method using the assurance probability

In order to introduce the method suggested by Kang et al. (2016), the assurance probability AP_2 based on the second similarity criterion (ii) $D_1 \geq \rho D$ proposed in the Japanese MHLW will be used. The main concepts of the suggested method are as follows.

(i) $D_1/D \geq \rho$ is assessed considering that it is a statistical hypothesis testing procedure to test the secondary hypothesis stated in section 3.2, where D_1/D and ρ are regarded as a test statistic and a critical value, respectively. Under the assumption that the primary alternative hypothesis is accepted, D is greater than 0 and thus $D_1 \geq \rho D$ and $D_1/D \geq \rho$ are the same.

(ii) Considering that the type II error is more serious than the type I error under the secondary hypothesis, the type II error rate with the secondary hypothesis is used to determine the associated parameters and the sample size in the region of interest.

In order to assess whether $D_1 \geq \rho D$ is satisfied through the statistical hypothesis testing procedure, there is a need to formulate the secondary hypothesis in section 3.2 into a statistical hypothesis.

Although there are many ways to express the secondary hypothesis in the forms of statistical hypothesis, in this thesis, the most commonly used hypotheses equations (7) and (8) will be applied. Let $\triangle_1$ and $\triangle$ be the treatment effects in the region of interest and all regions participating in the clinical trial, respectively,

$$\triangle_1 = \mu_{1,T} - \mu_{1,P} \quad \text{and} \quad \triangle = \mu_T - \mu_P$$

with $\mu_{1,T}$ and $\mu_{1,P}$ as the population means of the primary endpoint for the new treatment and placebo groups in the first region. If D_1/D is a natural estimator of $\triangle_1/\triangle$, the secondary hypothesis can be expressed as follows,

$$H_{s0} : \frac{\triangle_1}{\triangle} \geq 1 \quad \text{vs} \quad H_{sA} : \frac{\triangle_1}{\triangle} < 1 \tag{7}$$

Here, subscript “s" was used to indicate that the hypothesis refers to the secondary hypothesis. The hypothesis in (7) is called the secondary hypothesis, because the equation (1) refers to the primary hypothesis for the first purpose of MRCTs. Since the rejection region $D_1/D < \rho$ according to the similarity criterion (ii) is one-sided, the hypothesis in (7) should also be a one-sided. The secondary alternative hypothesis in (7) shows that the new treatment effect in the first region is less than that of all regions, which is a major concern from the perspective of regulatory agencies. In other words, it can be concluded that treatment effects are similar between regions only

when treatment effects in the region of interest is the same or greater compared to that of all regions.

The secondary hypothesis can applied to the similarity criterion (i) $D_1/D_{1c} \geq \rho$ to be expressed as shown below.

$$H_{s0} : \frac{\triangle_1}{\triangle_{1c}} \geq 1 \quad \text{vs} \quad H_{sA} : \frac{\triangle_1}{\triangle_{1c}} < 1 \tag{8}$$

where $\triangle_{1c}$ is the treatment effect in all regions excluding the first region. The fact that $\triangle$ can be expressed in terms of $\triangle_1$, $\triangle_{1c}$ and λ_1 is important in explaining the relationship between equations (7) and (8) in the secondary hypothesis. Here, λ_1 represents the ratio of total number of patients in the first region to the total number of patients in all participating regions $\triangle$ can be expressed as shown below.

$$\triangle = \lambda_1 \triangle_1 + (1 - \lambda_1) \triangle_{1c} \tag{9}$$

The relationship of equation (9) is explained in Appendix A. Assuming the value of λ_1 is approximately known and the primary alternative hypothesis is accepted, the secondary null hypothesis on the similarity criterion (i) $H_{s0} : \triangle_1/\triangle \geq 1$ becomes $H_{s0} : \triangle_1/\triangle \geq 1$. Using equation (9) results in $H_{s0} : \triangle_1 \geq \lambda_1\triangle_1 + (1-\lambda_1)\triangle_{1c}$ and rearranging the right clause gives $H_{s0} : (1-\lambda_1)\triangle_1 \geq (1-\lambda_1)\triangle_{1c}$, deriving the final equation $H_{s0} : \triangle_1 \geq \triangle_{1c}$, which is identical to the equation of the secondary null

hypothesis (8) of the second similarity criterion (ii).

The type I error related to the secondary hypothesis can be interpreted as an error resulting in the incorrect conclusion that there is a difference in treatment effects, despite the similar treatment effects of the new treatment in all participating regions and the first region of the clinical trial. On the other hand, the type II error related to the secondary hypothesis is an error resulting in the incorrect conclusion that there is no difference in treatment effects, despite the difference in treatment effects of new treatment in all participating regions and the first region of the clinical trial. From the perspective of regulatory authorities, an incorrect conclusion that claims the similarity when there is a difference, compared to the incorrect conclusion that claims a difference in treatment effects despite the similarity, is more serious. This is due to the fact that the type I error does not result in drug approval and therefore, although it may be conservative to developers, there is no serious risk for regulatory agencies. In contrast, the type II error would result in drug approval, leading to significant risk for the regulatory agencies. Therefore, decisions should be made with a greater focus on the type II error rather than on the type I error. From here on, the type I error and the type II error related to the secondary hypothesis will be referred to as the regional type I error and the regional type II error.

The second similarity criterion $D_1 \geq \rho D$ results in the equations below for the regional type I error rate (10) and the regional type II

error rate (11).

$$P(accept\ \ H_{sA}|H_{s0}) = P(D_1 < \rho D|H_{s0}) \quad (10)$$

$$P(accept\ \ H_{s0}|H_{sA}) = P(D_1 \geq \rho D|H_{sA}) \quad (11)$$

The regional type II error rate can be understood as a similar concept to the assurance probability, as it is actually assessed based on the secondary alternative hypothesis. As mentioned in chapter 3, the higher assurance probability suggested by Ko et al. (2010) is better but for the regional type II error rate, a lower rate is preferred.

In general, the critical value of a statistical hypothesis testing procedure should be determined under the null hypothesis in order to control the type I error rate under the significance level, as the type I error is more serous than the type II error. However, as mentioned, in the statistical hypothesis testing in equations (7) and (8) on the secondary hypothesis of MRCTs, regulatory agencies consider the regional type II error to be more serious than the regional type I error. Therefore, the regional type II error rate should be used to determine the sample size for clinical trials in the first region as well as associated parameters including critical value ρ and the proportion of patients out of $2N$ in the region of interest p_1.

The regional type I error rate and the regional type II error rate for the first similarity criterion (i) $D_1 \geq \rho D_{1c}$ $(0 < \rho < 1)$ using the concept in equations (10) and (11) are shown below. From here, the regional

type II error rate and the regional type I rate for the first similarity criterion (i) are $\beta_{1,s}$ and $\alpha_{1,s}$, respectively.

$$\beta_{1,s} = P_\delta(D_1 \geq \rho D_{1c} \mid Z > z_{1-\alpha},\ H_{sA} : \triangle_1/\triangle_{1c} < 1) \tag{12}$$
$$= \frac{\int_{b_1}^{\infty} [\Phi(\frac{1}{d_1}u - \frac{d_2}{d_1}) - \Phi(-\frac{d_3}{d_4}u + \frac{d_5}{d_4})]\phi(u)du}{\int_{-\infty}^{\infty} [1 - \Phi(-\frac{d_3}{d_4}u + \frac{d_5}{d_4})]\phi(u)du}$$

$$\alpha_{1,s} = P_\delta(D_1 < \rho D_{1c} \mid Z > z_{1-\alpha},\ H_{s0} : \triangle_1/\triangle_{1c} \geq 1) \tag{13}$$
$$= \frac{\int_{b_3}^{\infty} [\Phi(d_1u + d_2) - \Phi(-\frac{d_4}{d_3}u + \frac{d_5}{d_3})]\phi(u)du}{\int_{-\infty}^{\infty} [1 - \Phi(-\frac{d_3}{d_4}u + \frac{d_5}{d_4})]\phi(u)du}$$

where

$$d_1 = \rho\sqrt{\frac{p_1}{1-p_1}} \quad , \quad d_2 = \sqrt{\frac{Np_1}{2}}\,(\rho\frac{\triangle_{1c}}{\sigma} - \frac{\triangle_1}{\sigma}) \quad , \quad d_3 = \sqrt{p_1} \quad ,$$

$$d_4 = \sqrt{1-p_1} \quad , \quad d_5 = \sqrt{\frac{N}{2}}\,(p_1\frac{\triangle_1}{\sigma} + (1-p_1)\frac{\triangle_{1c}}{\sigma}) \quad ,$$

$$b_1 = d_2 + \frac{d_1(d_5 - d_2d_3)}{d_1d_3 + d_4}, \quad b_3 = \frac{d_5 - d_2d_3}{d_1d_3 + d_4}$$

The regional type II error rate and the regional type I error rate for the second similarity criterion (ii) $D_1 \geq \rho D$ $(0 < \rho < 1)$ are shown below. As with the similarity criterion (i), the regional type II error rate and the regional type I rate for the second similarity criterion (ii)

are $\beta_{2,s}$ and $\alpha_{2,s}$, respectively.

$$\beta_{2,s} = P_\delta(D_1 \geq \rho D \mid Z > z_{1-\alpha},\ H_{sA} : \triangle_1/\triangle < 1) \tag{14}$$
$$= \frac{\int_{b_2}^{\infty} [\Phi(\frac{1}{d_6}u - \frac{d_7}{d_6}) - \Phi(-\frac{d_8}{d_9}u + \frac{d_{10}}{d_9})]\phi(u)du}{\int_{-\infty}^{\infty} [1 - \Phi(-\frac{d_8}{d_9}u + \frac{d_{10}}{d_9})]\phi(u)du}$$

$$\alpha_{2,s} = P_\delta(D_1 < \rho D \mid Z > z_{1-\alpha},\ H_{s0} : \triangle_1/\triangle \geq 1) \tag{15}$$
$$= \frac{\int_{b_4}^{\infty} [\Phi(d_6u + d_7) - \Phi(-\frac{d_9}{d_8}u + \frac{d_{10}}{d_8})]\phi(u)du}{\int_{-\infty}^{\infty} [1 - \Phi(-\frac{d_8}{d_9}u + \frac{d_{10}}{d_9})]\phi(u)du}$$

where

$$d_6 = \rho\frac{\sqrt{p_1(1-p_1)}}{1-\rho p_1} \quad , \quad d_7 = d_6\sqrt{\frac{N(1-p_1)}{2}}\ \frac{\triangle_{1c}}{\sigma} - \sqrt{\frac{Np_1}{2}}\ \frac{\triangle_1}{\sigma} \quad , \quad d_8 = \sqrt{p_1} \quad ,$$

$$d_9 = \sqrt{1-p_1} \quad , \quad d_{10} = z_{1-\alpha} - \sqrt{\frac{N}{2}}\,(p_1\frac{\triangle_1}{\sigma} + (1-p_1)\frac{\triangle_{1c}}{\sigma}) \ ,$$

$$b_2 = d_7 + \frac{d_6(d_{10} - d_7d_8)}{d_6d_8 + d_9} \quad , \quad b_4 = \frac{d_{10} - d_7d_8}{d_6d_8 + d_9}$$

The mathematical derivations of the regional type II error standardized by the effect size of the first region ($\triangle_1/\sigma$) and other regions ($\triangle_{1c}/\sigma$) are provided in Appendix B. In the case of the regional type I error rate, the mathematical derivations are similar to those of the regional type II error rate and therefore is not shown here. Equations (12) and

(14) show that the regional type II error rate is dependent on ρ, p_1, α, N, $\triangle_1/\sigma$ and $\triangle_{1c}/\sigma$. The total sample size N per group calculated during the planning of MRCTs is determined by the overall significance level α, a given power and the overall effect size. Therefore, in calculating the regional type II error rate, N, α, β and $\triangle/\sigma$ should be considered as fixed constants and only ρ, p_1, $\triangle_1/\sigma$ and $\triangle_{1c}/\sigma$ should be recognized as parameters to be considered in the calculation.

In order to assess the similarity in the treatment effects in the first region according to similarity criteria (i) or (ii), the selection of critical value is extremely important. If regional type II error rates $\beta_{1,s}$ or $\beta_{2,s}$ were to be used to set the critical value ρ, then the values for p_1, N, $\triangle_1/\sigma$ and $\triangle_{1c}/\sigma$ should be determined beforehand. As mentioned, N can be considered as a fixed constant determined during the planning of MRCTs and thus there is a need to provide evidence to determine p_1, $\triangle_1/\sigma$ and $\triangle_{1c}/\sigma$. In order to use regional type II error rates to calculate critical value ρ, there are two situations to set p_1. The first situation is that the value can be determined before the initiation of an MRCT based on the conditions of clinical trial recruitment in the region of interest or the cost of the trial rather than on statistical grounds. The second situation is that the value can be determined for the sake of competitive registration of patients between regions to increase the speed of registration in the clinical trial. The second situation, unlike the first situation, provides the proportion of patients

from the first region as soon as the planned number of subjects is registered and follow-ups of patient are complete in the MRCT. Since both cases can be used to set p_1, if effect sizes ($\triangle_1/\sigma$ and $\triangle_{1c}/\sigma$) are assumed based on clinical or regulatory evidence and the regional type II error rate is specified in advance for example 10% or 20% then the critical value ρ that assesses the similarity while satisfying such a regional type II error rate can be calculated. In general, when designing an MRCT, the overall effect size $\triangle/\sigma$ is set based on results from the literature, previous studies or regulations. Similarly, $\triangle_1/\sigma$ and $\triangle_{1c}/\sigma$ can be set using the literature or previous studies, as well as on an empirical or regulatory basis. In this thesis, calculated results of the ρ values that satisfy the pre-specified regional type II error rate and its examples will be discussed in section 5.2 and section 6.1, respectively.

If regulatory agencies determine ρ, $\triangle_1/\sigma$ and $\triangle_{1c}/\sigma$ on an empirical or regulatory basis, regional type II error rates will be functions only dependent on p_1. Therefore, p_1 values that satisfy the pre-specified regional type II error rate, for example at 20%, can be determined. In other words, under the condition given by ρ, $\triangle_1/\sigma$ and $\triangle_{1c}/\sigma$, if the proportion of the number of patients in the first region to the number of patients in all participating region in the MRCT is p_1, the regulatory authority in the first region will use the determined regional type II error rate to discover the difference between $\triangle_1/\sigma$ and $\triangle_{1c}/\sigma$.

4.2. Distribution of the effect size through literature reviews

As mentioned in section 4.1, the regional type II error rate depends on ρ, p_1, $\triangle_1/\sigma$ and $\triangle_{1c}/\sigma$. The critical value ρ for assessing the similarity and the proportion of patients to be assigned in the first region that satisfy the pre-specified regional type II error rate change with the effect sizes of the first region and other regions excluding the region of interest. This section investigates the results of actual clinical trials through literature review in order to determine the assumed range for effect sizes $\triangle_1/\sigma$ and $\triangle_{1c}/\sigma$. This was done to allow calculations for the proportion of patients out of $2N$ in the first region p_1 which satisfies the given regional type II error rate and critical value ρ to assess the similarity.

The literature review was conducted using the results of the FDA statistical review on new molecular entity and therapeutic biological products for market release approval by the US FDA from 2011 to 2014. To apply the investigated clinical trial results to this thesis, phase III clinical trials with placebo group and a continuous variable as the primary endpoint were selected. The following procedure was used to investigate FDA approved drugs from 2011 to 2014.

(i) First step of investigation

An annual list of drugs was generated from the website below for the period between 2011 and 2014.

Research site:
http://www.fda.gov/Drugs/DevelopmentApprovalProcess/DrugInnovation/default.htm

(ii) Second step of investigation

For each of the listed drugs, the website stated below was used to search for active ingredients and the results were organized using the FDA statistical Review.

Search site:
http://www.accessdata.fda.gov/scripts/cder/drugsatfda/

The effect sizes of new drugs that were investigated from 2011 to 2014 are listed in Table A1 of Appendix C. Table 1 summarizes the effect sizes from Table A1 to suggest basic statistics.

Table 1. Summary of the effect size of approved drugs from 2011 to 2014

Mean	SD	Median	25% percentile	75% percentile	MIN	MAX
0.68	0.39	0.58	0.41	0.96	0.03	1.7

Fukunaga et al. (2014) investigated effect sizes and research designs of Phase II and III clinical trials on approved drugs for depression, schizophrenia, asthma, high blood pressure and diabetes in Japan

from 1970 to 2011. The range of effect sizes investigated by Fukunaga et al. (2014) was from -0.64 to 1.94 and the average was 0.19. Since not only did this result include placebo controlled comparison clinical trials, active comparator controlled treatment clinical trials and failed clinical trial results but also a relatively limited number of diseases was investigated, the results were broader than that of this thesis. However, the box-whisker plot of effect sizes of each disease separated by an comparator controlled and placebo controlled groups confirms the appropriateness of the effect size range suggested in this thesis.

In this thesis, the assumed range of Δ_1/σ and Δ_{1c}/σ for calculating ρ and p_1 that satisfy the pre-specified regional type II error rate is set from 0.4 to 0.9 which corresponds roughly between the 20% and 75% percentiles in Table 1.

Chapter 5. Results

5.1. $\beta_{1,s}$ and $\beta_{2,s}$ according to the effect size

In this section, regional type II error rates $\beta_{1,s}$ and $\beta_{2,s}$ according to the effect sizes of the region of interest and other regions excluding the region of interest, the proportion of patients assigned to the region of interest and the critical value to assess the similarity were calculated and the changes were confirmed. In particular, in cases where ρ and p_1 are fixed (for example, in cases where regulatory agencies use empirical or regulatory basis), the changes in regional type II error rates with changes in the effect sizes in the region of interest and other regions were studied. Calculations for the regional type II error rate were conducted based on the conditions described below.

1) $\alpha = 0.025$
2) $100 \leq N \leq 1{,}000$, 100 unit change
3) Effect size : $0.1 \leq \Delta_1/\sigma \leq 1.4$, 0.1 unit change
   $0.2 \leq \Delta_{1c}/\sigma \leq 1.5$ unit change
4) $0.5 \leq \rho \leq 0.9$, 0.1 unit change
5) $0.1 \leq p_1 \leq 0.9$, 0.1 unit change

Calculated regional type II error rates according to the conditions above are not all provided in a table but the representative regional type II error rates $\beta_{1,s}$ and $\beta_{2,s}$ for N=500, ρ=0.5, 0.7, 0.9 and effect size in region of interest $\triangle_1/\sigma$=0.1 are shown in Table 2. Regional type II error rates are written as 0.00 in the table if they were smaller than 0 after rounding at the 3rd decimal place.

Taking the first row of Table 2 as an example, the values are regional type II error rates $\beta_{1,s}$ and $\beta_{2,s}$ according to p_1 when the significance level is 0.025, ρ is 0.5, $\triangle_1/\sigma$ is 0.1 and $\triangle_{1c}/\sigma$ is 0.2. If p_1 is 0.1, $\beta_{1,s}$ and $\beta_{2,s}$ corresponding to similarity criteria (i) and (ii) are 0.52 and 0.53, respectively.

Figure 1 shows the relationship between $\beta_{1,s}$ or $\beta_{2,s}$ and the proportion of patient out of $2N$ in the first region p_1 for ρ values from some of the results in Table 1. This includes combinations of $\triangle_1/\sigma$ and $\triangle_{1c}/\sigma$ to be (0.1, 0.2), (0.1, 0.3), and (0.1, 0.4).

It can be confirmed through Table 2 and figure 1 that if ρ, $\triangle_1/\sigma$, $\triangle_{1c}/\sigma$ and p_1 are the same, respectively, $\beta_{1,s}$ associated with the similarity criterion (i) is smaller than $\beta_{2,s}$ associated with the similarity criterion (ii). When the proportion of patients in the region of interest p_1 is small, there is little difference between $\beta_{1,s}$ and $\beta_{1,s}$, but as the value of p_1 approaches 1, the difference between $\beta_{1,s}$ and $\beta_{1,s}$ increases. Furthermore, when ρ and p_1 are the same, respectively, as the difference between $\triangle_1/\sigma$ and $\triangle_{1c}/\sigma$ increases, the difference

between $\beta_{1,s}$ and $\beta_{2,s}$ decreases. And when p_1, $\triangle_1/\sigma$ and $\triangle_{1c}/\sigma$ are the same, respectively, as ρ increases $\beta_{1,s}$ and $\beta_{2,s}$ decrease.

When $\triangle_1/\sigma = 0.1$ and $\triangle_{1c}/\sigma = 0.2$, $\beta_{1,s}$ and $\beta_{2,s}$ cannot be lower than 20%, regardless of ρ or similarity criteria. However as the difference between $\triangle_1/\sigma$ and $\triangle_{1c}/\sigma$ and ρ value increases, $\beta_{1,s}$ and $\beta_{2,s}$ show a tendency to be lower than 20% but increases with p_1 increase. Concerning the similarity criterion (i), as the difference between $\triangle_1/\sigma$ and $\triangle_{1c}/\sigma$ increases and ρ and p_1 increases, $\beta_{1,s}$ shows decreases and then increases after a certain point. However, the range is small. On the other hand, for the similarity criterion (ii), if the difference between $\triangle_1/\sigma$ and $\triangle_{1c}/\sigma$ are not small and the ρ value is small, as p_1 increases $\beta_{2,s}$ also increases greatly. For example, if ρ is 0.5 and $\triangle_1/\sigma$ and $\triangle_{1c}/\sigma$ are 0.1 and 0.2, respectively, as p_1 approaches 1, $\beta_{1,s}$ increases from 0.5 to over 0.6 and then decreases, but $\beta_{2,s}$ increases with p_1 increase. In the same condition, if ρ is 0.9, $\beta_{1,s}$ decreases until p_1 is 0.5 and increases as p_1 approaches 1. However, the difference between the decreasing magnitude before $p_1 = 0.5$ and the increasing magnitude after $p_1 = 0.5$ are not significantly large and similar. On the other hand, $\beta_{2,s}$ showed a tendency to decrease until p_1 is around 0.5 and then considerably increased as p_1 approached 1.

Table 2. The regional type II error rate when $N=500$

ρ	$\frac{\triangle_1}{\sigma}$	$\frac{\triangle_{1c}}{\sigma}$	$p_1 : \beta_{1,s}$									$p_1 : \beta_{2,s}$								
			0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.5	0.1	0.2	0.52	0.53	0.54	0.56	0.57	0.59	0.60	0.61	0.60	0.53	0.56	0.61	0.67	0.74	0.82	0.90	0.97	1.00
		0.3	0.40	0.37	0.35	0.33	0.33	0.34	0.35	0.38	0.42	0.42	0.41	0.43	0.47	0.54	0.64	0.77	0.91	1.00
		0.4	0.31	0.25	0.21	0.18	0.16	0.16	0.16	0.19	0.26	0.33	0.30	0.29	0.32	0.37	0.46	0.61	0.82	0.99
		0.5	0.23	0.15	0.11	0.08	0.07	0.06	0.06	0.08	0.14	0.25	0.20	0.19	0.20	0.24	0.31	0.45	0.69	0.97
		0.6	0.16	0.09	0.05	0.03	0.02	0.02	0.02	0.03	0.06	0.18	0.13	0.11	0.12	0.14	0.21	0.32	0.56	0.93
		0.7	0.11	0.04	0.02	0.01	0.01	0.00	0.00	0.01	0.02	0.13	0.07	0.06	0.06	0.08	0.12	0.22	0.44	0.88
		0.8	0.07	0.02	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.08	0.04	0.03	0.03	0.04	0.07	0.15	0.34	0.81
		0.9	0.04	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.02	0.01	0.01	0.02	0.03	0.09	0.26	0.73
		1.0	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.01	0.00	0.00	0.01	0.02	0.05	0.19	0.64
		1.1	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.01	0.03	0.13	0.56
		1.2	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.09	0.49
0.7	0.1	0.2	0.43	0.41	0.40	0.39	0.39	0.40	0.41	0.42	0.44	0.44	0.44	0.45	0.48	0.53	0.60	0.70	0.83	0.96
		0.3	0.30	0.23	0.20	0.17	0.17	0.17	0.18	0.21	0.27	0.31	0.26	0.25	0.26	0.29	0.36	0.47	0.66	0.91
		0.4	0.19	0.11	0.08	0.06	0.05	0.05	0.06	0.08	0.14	0.20	0.14	0.12	0.11	0.13	0.17	0.26	0.46	0.83
		0.5	0.11	0.05	0.02	0.02	0.01	0.01	0.01	0.02	0.06	0.12	0.06	0.05	0.04	0.05	0.07	0.13	0.28	0.72
		0.6	0.06	0.02	0.01	0.00	0.00	0.00	0.00	0.00	0.02	0.07	0.03	0.01	0.01	0.01	0.02	0.05	0.16	0.58
		0.7	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.03	0.01	0.00	0.00	0.00	0.01	0.02	0.08	0.44
		0.8	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.01	0.03	0.32
		0.9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.22
0.9	0.1	0.2	0.35	0.31	0.28	0.27	0.26	0.27	0.28	0.30	0.35	0.36	0.32	0.30	0.30	0.31	0.34	0.38	0.47	0.63
		0.3	0.21	0.14	0.10	0.09	0.08	0.08	0.10	0.13	0.20	0.21	0.15	0.12	0.11	0.11	0.12	0.16	0.25	0.45
		0.4	0.11	0.05	0.03	0.02	0.02	0.02	0.02	0.04	0.09	0.11	0.05	0.03	0.03	0.02	0.03	0.05	0.10	0.28
		0.5	0.05	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.04	0.05	0.01	0.01	0.00	0.00	0.01	0.01	0.03	0.15
		0.6	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.07
		0.7	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03

$\rho = 0.5$	$\rho = 0.7$	$\rho = 0.9$

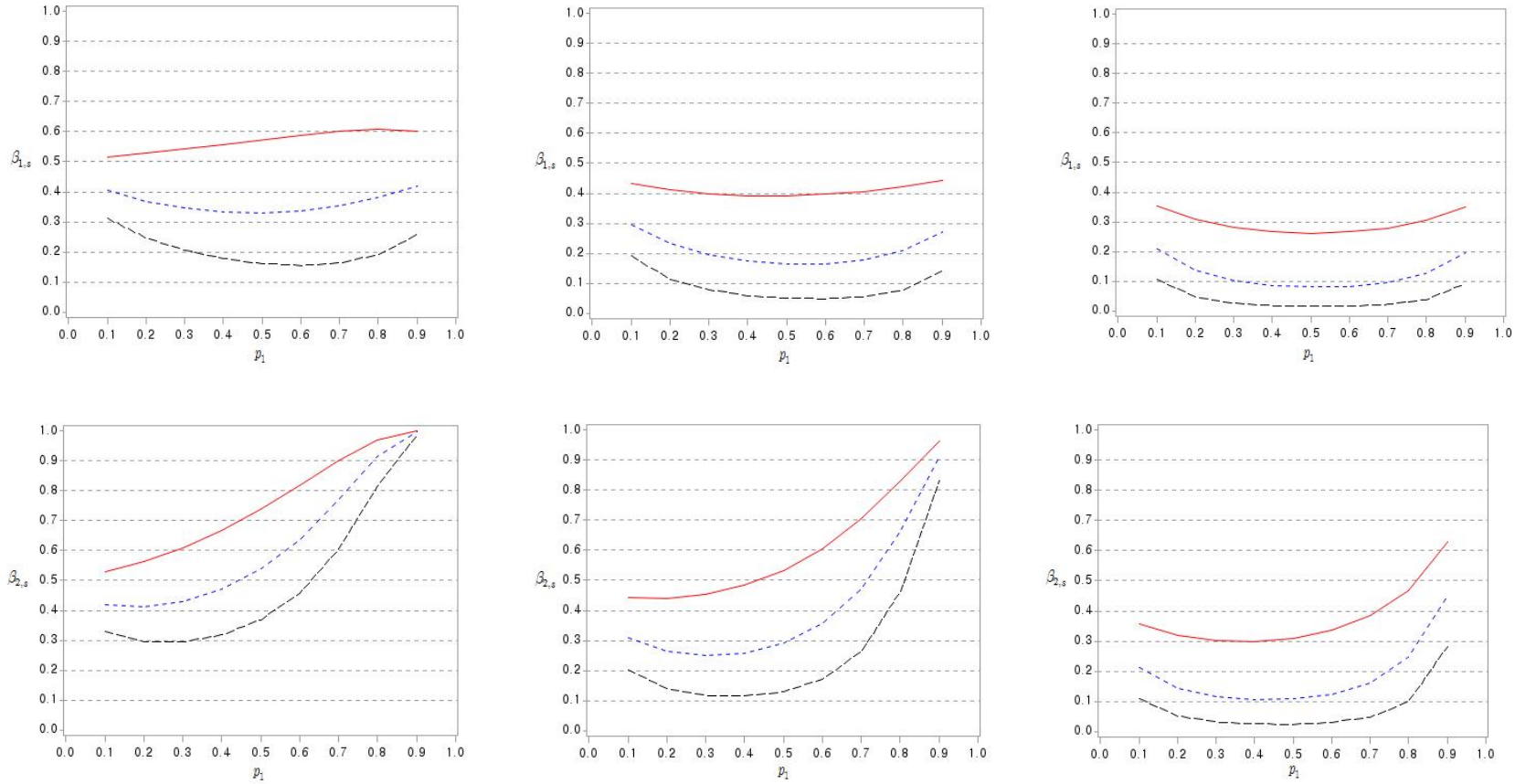


Figure 1. Graph of the regional type II error rate vs p_1 when $N = 500$

5.2. Critical value ρ selection according to the effect size

Selection of critical value ρ to assess the similarity is an important standard that determines the similarity between treatment effects in the region of interest and all regions. According to Ko et. al. (2010), selection of ρ needs to consider the ethnic factors between the region of interest and other regions, and it can be decided by the regulatory authorities in the region. The Japanese MHLW guideline recommends that the ρ value should be over 0.5 in the similarity criterion (ii) $D_1 \geq \rho D$, but does not discuss the standard for the selection from a statistical perspective.

In this section, the changes in critical value ρ according to the regional type II error rate will be investigated in order to discuss a method to select a ρ value that satisfies the pre-specified regional type II error rate.

5.2.1. Conditions for $\beta_{1,s}$ and $\beta_{2.s}$ calculation to select ρ

The critical value ρ for the similarity criteria should be determined during the planning of the clinical trial design. In order to select a ρ value during clinical trial planning, the effect sizes of the region of interest and other regions excluding the region of interest should be assumed. In this section, in order to select a ρ value in a condition similar to that of a real clinical trial environment, effect sizes based

on literature review will be used to calculate regional type II error rates $\beta_{1,s}$ and $\beta_{2,s}$.

Effect sizes based on literature review of results of US FDA review on approved drugs from 2011 to 2014, as mentioned in section 4.2, were set to change in units of 0.05, within the range of 0.5-0.9. The total sample size per group N was assumed to be 100, 300, 500, 700, 1,000, and 1,500, and the critical value ρ for similarity criteria was set in units of 0.001, within the range 0.5-0.99.

The proportion of patients out of $2N$ in the region of interest p_1 was determined using the presented results from a Pharmaceutical and Medical Devices Agency (PMDA) biostatistics summer workshop in 2012 (Ando, 2012). From the literature review of Ando (2012), MRCT results of Japanese PMDA approved drugs from Feb 2006 to June 2012 were used to investigate the assigned ratio in the Japan and Asia region. The range of the ratio was set to 0.05-0.5, which included the 25^{th} percentile to 75^{th} percentile from the investigated results. And Regional type II error rate calculation was conducted within this range by changing by 0.05.

[Summary of $\beta_{1,s}$ and $\beta_{2,s}$calculation conditions to select ρ value]

(1) $\alpha = 0.025$

(2) Effect size : $0.4 \leq \Delta_1/\sigma \leq 0.85$, $0.45 \leq \Delta_{1c}/\sigma \leq 0.9$, 0.05 unit change

(3) N=100, 300, 500, 700, 1,000, 1,500

(4) $0.5 \leq \rho \leq 0.99$, 0.001 unit change

(5) $0.05 \leq p_1 \leq 0.5$, 0.05 unit change

5.2.2. Selection of a ρ value that satisfies the pre-specified $\beta_{1,s}$ and $\beta_{2.s}$

To determine a ρ value that satisfies the pre-specified regional type II error rate, ρ values were calculated for when regional type II error rates are 10%, 15%, 20%, 25%, 30% and 35% for each similarity criterion according to the changes in N, p_1, $\triangle_1/\sigma$ and $\triangle_{1c}/\sigma$.

The calculated results from the conditions above are not all listed in this thesis due to space limitations. However, the results for when N is 500 and the ρ values that satisfy the pre-determined regional type II error rates (10%, 20%, and 30%) according to the combination of effect sizes ($\triangle_1/\sigma$, $\triangle_{1c}/\sigma$) are listed by p_1 values in Table A2 of Appendix D. In Table A2, the empty cells refer to ρ values larger than 1 after rounding at the 3rd decimal place. For these cells, ρ can be considered to be 1 because the range of ρ values is between 0 and 1. In addition, the combinations of $\triangle_1/\sigma$ and $\triangle_{1c}/\sigma$ that are not listed in Table A2 among the range of effect size in section 5.2.1, are the cases where the ρ value that satisfies the regional type II error rate is not between 0 and 1. In other words, if the ρ value that satisfies the regional type II error rate is greater than 1, it is not listed in the table and those empty cells can be assumed to be 1.

Table 3 is an excerpt of Table A2 only when $\triangle_1/\sigma$ is 0.4. Table 3 shows that when $\beta_{1,s}$ or $\beta_{2,s}$ is the same, and N, $\triangle_1/\sigma$, $\triangle_{1c}/\sigma$ and p_1 are the same, respectively, the critical value ρ for the similarity criterion (ii) is greater than the critical value ρ for the similarity criterion (i). When all other parameter conditions are the same, respectively, the similarity criterion (ii) can be seen as more conservative in assessing the similarity of effect size in the region of interest than the similarity criterion (i). In other words, if all parameters are the same, respectively, even if there is no actual similarity, there is a higher probability of the similarity criterion (i) to assess the case to be similar than the similarity criterion (ii). Furthermore, if N and p_1 are the same, respectively, the ρ value that satisfies the pre-determined $\beta_{1,s}$ or $\beta_{2,s}$ decreases as the difference between the effect sizes of the first region $\triangle_1/\sigma$ and other regions $\triangle_{1c}/\sigma$ increases. This phenomenon indicates that when the effect size in the region of interest is considerably smaller than that of other regions, the critical value ρ of the similarity is further reduced in assessing the similarity. For additional explanation, suppose that the effect size in the region of interest is predicted to be much smaller than that of other regions. In spite of this difference of the effect size, if the MRCT is conducted because there are no clinically ethnic differences in the treatment effect of the new treatment, critical value ρ needs to be reduced in order to increase the possibility of confirming the similarity.

The above-mentioned properties can also be seen in Figure 2. The figure shows the ρ values that satisfy $\beta_{1,s}$ and $\beta_{2,s}$ at 20% in relation to the changes of p_1 when N is 500 and the combination of the effect sizes in the region of interest and other regions ($\triangle_1/\sigma, \triangle_{1c}/\sigma$) are (0.4, 0.6), (0.4, 0.7), (0.4, 0.8) and (0.4, 0.9). As shown in Figure 2, when N, p_1, $\triangle_1/\sigma$ and $\beta_{1,s}$ (or $\beta_{2,s}$) are the same, respectively, as $\triangle_{1c}/\sigma$ increases the ρ value decreases. Accordingly, the ρ value for the similarity criterion (i) is smaller than that of the similarity criterion (ii). Furthermore, the ρ value that satisfies the pre-determined $\beta_{1,s}$ and $\beta_{2,s}$ decreases as p_1 increases and increases after a certain p_1, regardless of similarity criteria (i) or (ii). The regional type II error rate also decreases as ρ increases. This can be easy to verify that the regional type II error rates are decreasing functions of ρ, since functions (12) and (13) for the regional type II error rate include ρ only in d_1 and d_6, respectively.

Table 3. Selection of ρ that satisfies the pre-determined regional type II error rate when $N=500$

β_{1s} or $\beta_{2,s}$	$\frac{\triangle_1}{\sigma}$	$\frac{\triangle_{1c}}{\sigma}$	p_1 : Similarity criteria (i)										p_1 : Similarity criteria (ii)									
			0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
0.10	0.40	0.60				1.00	0.97	0.95	0.94	0.93	0.92	0.92				1.00	0.98	0.97	0.96	0.96	0.96	0.96
		0.65			0.96	0.92	0.89	0.87	0.86	0.85	0.85	0.84			0.96	0.93	0.92	0.91	0.90	0.91	0.91	0.91
		0.70		0.95	0.88	0.85	0.82	0.81	0.79	0.79	0.78	0.77		0.95	0.90	0.87	0.86	0.86	0.86	0.86	0.87	0.87
		0.75		0.88	0.82	0.79	0.77	0.75	0.74	0.73	0.72	0.72		0.89	0.85	0.82	0.81	0.81	0.81	0.82	0.83	0.84
		0.80	0.95	0.83	0.77	0.74	0.72	0.70	0.69	0.68	0.67	0.67	0.96	0.84	0.80	0.78	0.77	0.77	0.77	0.78	0.79	0.80
		0.85	0.90	0.78	0.72	0.69	0.67	0.66	0.65	0.64	0.63	0.63	0.90	0.79	0.75	0.74	0.73	0.73	0.74	0.75	0.76	0.77
		0.90	0.85	0.73	0.68	0.65	0.63	0.62	0.61	0.60	0.59	0.59	0.85	0.75	0.72	0.70	0.70	0.70	0.71	0.71	0.73	0.74
0.20	0.40	0.55			0.99	0.96	0.94	0.93	0.92	0.92	0.91	0.91			0.99	0.97	0.96	0.95	0.95	0.95	0.95	0.95
		0.60		0.96	0.91	0.88	0.86	0.85	0.84	0.83	0.83	0.83		0.96	0.92	0.90	0.89	0.89	0.89	0.89	0.90	0.91
		0.65	0.98	0.88	0.84	0.81	0.79	0.78	0.77	0.77	0.76	0.76	0.98	0.89	0.86	0.84	0.84	0.84	0.84	0.85	0.85	0.86
		0.70	0.91	0.82	0.77	0.75	0.73	0.72	0.71	0.71	0.70	0.70	0.92	0.83	0.80	0.79	0.79	0.79	0.79	0.80	0.81	0.82
		0.75	0.85	0.76	0.72	0.70	0.68	0.67	0.66	0.66	0.65	0.65	0.86	0.78	0.75	0.74	0.74	0.75	0.75	0.76	0.78	0.79
		0.80	0.80	0.71	0.68	0.65	0.64	0.63	0.62	0.62	0.61	0.61	0.81	0.73	0.71	0.70	0.70	0.71	0.72	0.73	0.74	0.76
		0.85	0.75	0.67	0.64	0.61	0.60	0.59	0.58	0.58	0.57	0.57	0.76	0.69	0.67	0.67	0.67	0.67	0.68	0.70	0.71	0.73
		0.90	0.71	0.63	0.60	0.58	0.57	0.56	0.55	0.55	0.54	0.54	0.72	0.66	0.64	0.63	0.64	0.64	0.65	0.67	0.68	0.70
0.30	0.40	0.50			0.98	0.96	0.95	0.94	0.93	0.93	0.93	0.93			0.98	0.97	0.96	0.96	0.96	0.96	0.96	0.96
		0.55	1.00	0.92	0.89	0.87	0.86	0.85	0.84	0.84	0.84	0.84	1.00	0.93	0.90	0.89	0.89	0.89	0.89	0.90	0.90	0.91
		0.60	0.91	0.84	0.81	0.80	0.79	0.78	0.77	0.77	0.77	0.76	0.92	0.86	0.84	0.83	0.83	0.83	0.84	0.85	0.86	0.87
		0.65	0.84	0.78	0.75	0.73	0.72	0.72	0.71	0.71	0.70	0.70	0.85	0.80	0.78	0.78	0.78	0.78	0.79	0.80	0.81	0.83
		0.70	0.78	0.72	0.70	0.68	0.67	0.66	0.66	0.66	0.65	0.65	0.79	0.74	0.73	0.73	0.73	0.74	0.75	0.76	0.77	0.79
		0.75	0.73	0.67	0.65	0.64	0.63	0.62	0.61	0.61	0.61	0.61	0.74	0.70	0.69	0.69	0.69	0.70	0.71	0.72	0.74	0.75
		0.80	0.68	0.63	0.61	0.60	0.59	0.58	0.57	0.57	0.57	0.57	0.69	0.66	0.65	0.65	0.65	0.66	0.68	0.69	0.71	0.72
		0.85	0.64	0.59	0.57	0.56	0.55	0.55	0.54	0.54	0.53	0.53	0.65	0.62	0.61	0.61	0.62	0.63	0.64	0.66	0.68	0.69
		0.90	0.61	0.56	0.54	0.53	0.52	0.51	0.51	0.51	0.50	0.50	0.62	0.59	0.58	0.58	0.59	0.60	0.62	0.63	0.65	0.67

$\beta_{1,s}$	$\beta_{2,s}$

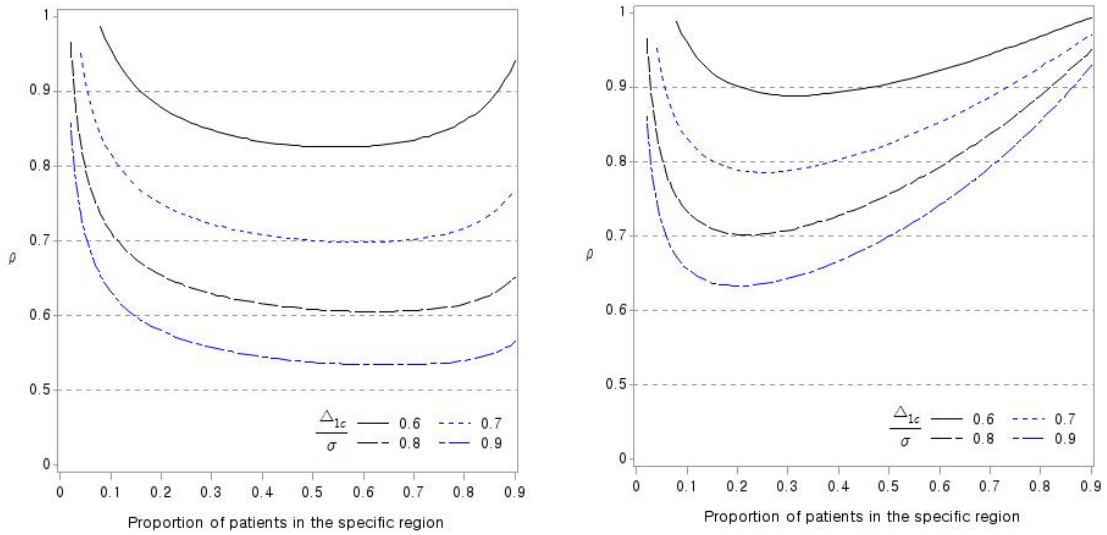


Figure 2. Graph of ρ vs p_1 when the regional type II error rate is 20%

5.3. Selection of p_1 according to the effect size

The sample size in the region of interest is investigated by selecting p_1 that satisfies $\beta_{1,s}$ for the similarity criterion (i) or $\beta_{2,s}$ for the similarity criterion (ii) according to the critical value ρ to assess the similarity, the changes of the effect size in the region of interest Δ_1/σ and the changes of effect size in other regions Δ_{1c}/σ.

5.3.1. Conditions for $\beta_{1,s}$ and $\beta_{2.s}$ calculation to select p_1

The proportion of patients out of $2N$ in the region of interest p_1 was calculated according to the size of $\beta_{1,s}$ and $\beta_{2,s}$ for the similarity criteria by changing the ranges of effect size ($\triangle_1/\sigma$, $\triangle_{1c}/\sigma$), N and ρ. Here, the range of effect sizes was limited to 0.4-0.9 and 0.1-0.5, respectively with unit change of 0.05 to present p_1. The reason for dividing the effect size range into two was as follows: the effect size range of 0.4-0.9 was chosen using the FDA review results of approved drugs from 2011 to 2014 from section 4.2 and the range was between the 25th percentile and the 75th percentile of the investigation results. In general, if a significance level α, power $1-\beta$ and an expected treatment effect $\triangle=\delta$ are given, then the equation to calculate the overall sample size for each group in the clinical trial is as shown in the widely known equation (2). In order to calculate the number of required clinical trial subjects, it is customary to assume power $1-\beta$ to be between 80%-90%. The expected treatment effect is estimated based on previous clinical trials or literature review to compute the sample size. Using equation (2) for the required number of clinical trial subjects, if the significance level α, N and power $1-\beta$ are pre-determined, then the expected effect size $\triangle/\sigma=\delta/\sigma$ can be calculated as shown below.

$$\frac{\delta}{\sigma} = (z_{1-\alpha} + z_{1-\beta})\sqrt{\frac{2}{N}} \quad (17)$$

As mentioned, since power in clinical trials is conventionally set between 80%-90%, when the number of subjects for each group in the trial N is set to 100, 300, 500, 700, 1,000, and 1,500 with the significance level $\alpha = 0.025$ and the given power range, the expected effect size commonly used in the sample size calculation is as shown in Table 4. The table shows that when N is 100 and power is between 80% and 90%, the effect size is between 0.4 and 0.46; when N is 500, the effect size is between 0.18 and 0.21; and when N is 1,500, the effect size is between 0.1 and 0.12. This shows that as N increases, the effect size decreases. If Table 4 was to be used to design clinical trial with power within 80%-90% and N below 1,500, the configurable effect size is predicted to be between 0.1 and 0.5. For this reason, to calculate the proportion of required subjects for clinical trial in the region of interest in real clinical trial conditions, the range of 0.1-0.5 was additionally included to be used in the regional type II error calculation for p_1 selection.

As a calculation condition for selecting p_1, the overall sample size per group N was set to 100, 300, 500, 700, 1,000 and 1,500 and critical value ρ for the similarity assessment was set between 0.5 and 0.9 with 0.1 unit change. Furthermore, p_1 was changed within the range of 0-0.9, with 0.0002 unit change to calculate $\beta_{1,s}$ and $\beta_{2,s}$ for

selecting a p_1 that satisfies the pre-specified regional type II error rate. Here, the significance level α was 0.025.

Table 4. Expected effect size according to $1-\beta$ and N

$1-\beta$	N	α	$z_{1-\alpha}$	$z_{1-\beta}$	$\triangle/\sigma$
0.9	100	0.025	1.96	1.28	0.46
0.9	300	0.025	1.96	1.28	0.26
0.9	500	0.025	1.96	1.28	0.21
0.9	700	0.025	1.96	1.28	0.17
0.9	1,000	0.025	1.96	1.28	0.14
0.9	1,500	0.025	1.96	1.28	0.12
0.85	100	0.025	1.96	1.04	0.42
0.85	300	0.025	1.96	1.04	0.24
0.85	500	0.025	1.96	1.04	0.19
0.85	700	0.025	1.96	1.04	0.16
0.85	1,000	0.025	1.96	1.04	0.13
0.85	1,500	0.025	1.96	1.04	0.11
0.8	100	0.025	1.96	0.84	0.40
0.8	300	0.025	1.96	0.84	0.23
0.8	500	0.025	1.96	0.84	0.18
0.8	700	0.025	1.96	0.84	0.15
0.8	1,000	0.025	1.96	0.84	0.13
0.8	1,500	0.025	1.96	0.84	0.10

[Summary of $\beta_{1,s}$ and $\beta_{2,s}$ calculation conditions for p_1 value selection]

(1) $\alpha = 0.025$

(2) Effect size

· Literature review: $0.4 \leq \Delta_1/\sigma \leq 0.85,\ 0.45 \leq \Delta_{1c}/\sigma \leq 0.9$

0.05 unit change

· Actual clinical trial condition : $0.1 \leq \Delta_1/\sigma \leq 0.45,\ 0.15 \leq \Delta_{1c}/\sigma \leq 0.5$

0.05 unit change

(3) N=100, 300, 500, 700, 1,000, 1,500

(4) $0.5 \leq \rho \leq 0.9$, 0.1 unit change

(5) $0 \leq p_1 \leq 0.9$, 0.0002 unit change

5.3.2. p_1 calculation for the sample size in the region of interest

p_1 values that satisfy $\beta_{1,s}$ and $\beta_{2,s}$ according to the effect sizes and the changes in the ρ value were calculated and the results are listed in Table A3 of Appendix E, when ρ values are 0.8 and 0.9, regional type II rates are 10%, 15%, 20%, 25%, 30% and 35% and effect sizes of the region of interest are 0.2 and 0.4. Table 5 is the table when ρ is 0.9 and $\beta_{1,s}$ and $\beta_{2,s}$ satisfy 0.15 and 0.2 in Table A3 of Appendix E. The empty cells in the table refer to cases without p_1 that satisfies the pre-specified regional type II error rate for the given condition of N, ρ, Δ_1/σ and Δ_{1c}/σ and thus the regional type II error rate is not

below 0.15 or 0.2 within the p_1 range of 0-1. For example, when $\rho = 0.9$, $\triangle_1/\sigma = 0.2$, $\triangle_{1c} = 0.3$ and $N = 100$, $\beta_{1,s}$ or $\beta_{2,s}$ is never lower than 0.2 within the range of p_1. In addition, p_1 values, which is calculated in the ranges of $\triangle_1/\sigma$ and $\triangle_{1c}/\sigma$ not shown in the table, are not included in the table because $\beta_{1,s}$ or $\beta_{2,s}$ is never lower than 0.1 or 0.2 with the p_1 range of 0-1.

As an example of the interpretation of the values in the table, when the overall sample size per group is 1,000, $\triangle_1/\sigma = 0.4$, $\triangle_{1c}/\sigma = 0.6$, $\rho = 0.9$ and the similarity criterion (i) is used to assess the similarity, p_1 value which satisfies $\beta_{1,s} = 0.2$ is 0.2598. Therefore, the sample size for the region of interest is 260 per group, which means that the total number of patients for the region of interest in the MRCT is 520. When the similarity criterion (ii) is used to assess the similarity, p_1 value which satisfies $\beta_{2,s} = 0.2$ is 0.0792 and therefore the sample size per group for the region of interest is 80 making the total number of patient 160.

When N, ρ, $\triangle_1/\sigma$ and $\triangle_{1c}/\sigma$ are given, p_1 does not monotonically increase or decrease with regional type II error rates $\beta_{1,s}$, $\beta_{2,s}$. As a result of this relationship, there can be two p_1 values between 0 and 1 for a pre-determined regional type II error rate, as can be seen in the p_1 graph in relation to $\beta_{1,s}$ or $\beta_{2,s}$. Figure 3 is a scatter plot of $\beta_{1,s}$ and p_1, and $\beta_{2,s}$ and p_1 for when $N = 1{,}000$, $\rho = 0.9$, $\triangle_1/\sigma = 0.4$ and $\triangle_{1c}/\sigma = 0.6$. The plot shows that as p_1 changes, there is no monotone

increase or decrease in $\beta_{1,s}$ or $\beta_{2,s}$. And it can be observed that there are no or more than one p_1 value satisfying the pre-determined $\beta_{1,s}$ and $\beta_{2,s}$. Furthermore, when other parameters (effect size, ρ) are fixed, the regional type II error rate does not go below a certain value as N decreases. This phenomenon occurs when there are cases with smaller difference between $\triangle_1/\sigma$ and $\triangle_{1c}/\sigma$ and smaller ρ value assuming other parameters are fixed, and is also shown in Figure A1 and A2 in Appendix F.

Table 5. p_1 to calculate the sample size in the region of interest when $\rho = 0.9$

$\beta_{1,s}$ or $\beta_{2,s}$	$\triangle_1/\sigma$	$\triangle_{1c}/\sigma$	$N : D_1 \geq \rho D_{1c}$						$N : D_1 \geq \rho D$					
			100	300	500	700	1000	1500	100	300	500	700	1000	1500
0.15	0.2	0.35				0.3000	0.1846	0.1150					0.2228	0.1252
		0.40		0.4486	0.1922	0.1290	0.0868	0.0564			0.2242	0.1396	0.0910	0.0580
		0.45		0.1960	0.1078	0.0748	0.0512	0.0338		0.2240	0.1138	0.0774	0.0524	0.0342
		0.50		0.1224	0.0700	0.0492	0.0340	0.0224		0.1296	0.0722	0.0502	0.0344	0.0226
	0.4	0.55					0.3126	0.1792						
		0.60			0.2750	0.1764	0.1164	0.0748				0.2392	0.1330	0.0804
		0.65		0.2572	0.1358	0.0934	0.0636	0.0416			0.1548	0.1006	0.0668	0.0428
		0.70		0.1482	0.0838	0.0586	0.0404	0.0266		0.1682	0.0886	0.0608	0.0414	0.0270
		0.75	0.4704	0.0990	0.0572	0.0404	0.0280	0.0186		0.1054	0.0592	0.0412	0.0284	0.0188
		0.80	0.2582	0.0714	0.0418	0.0296	0.0206	0.0136		0.0740	0.0426	0.0300	0.0208	0.0138
		0.85	0.1828	0.0540	0.0318	0.0226	0.0158	0.0106	0.2076	0.0554	0.0324	0.0228	0.0158	0.0106
		0.90	0.1388	0.0424	0.0252	0.0178	0.0124	0.0084	0.1496	0.0432	0.0254	0.0180	0.0126	0.0084
0.2	0.2	0.30						0.2292						
		0.35			0.2656	0.1712	0.1134	0.0728				0.2014	0.1232	0.0764
		0.40		0.2170	0.1176	0.0812	0.0556	0.0366		0.2666	0.1262	0.0850	0.0572	0.0372
		0.45		0.1196	0.0686	0.0482	0.0334	0.0220		0.1272	0.0708	0.0492	0.0338	0.0222
		0.50	0.3050	0.0774	0.0452	0.0320	0.0222	0.0148		0.0800	0.0460	0.0324	0.0224	0.0148
	0.4	0.55				0.2840	0.1766	0.1104						0.1322
		0.60		0.3182	0.1600	0.1088	0.0738	0.0482			0.2026	0.1226	0.0792	0.0502
		0.65		0.1514	0.0854	0.0596	0.0412	0.0272		0.1772	0.0914	0.0624	0.0424	0.0276
		0.70	0.3932	0.0928	0.0538	0.0380	0.0264	0.0174		0.0988	0.0556	0.0388	0.0268	0.0176
		0.75	0.2210	0.0632	0.0372	0.0264	0.0184	0.0122	0.2840	0.0654	0.0378	0.0266	0.0184	0.0122
		0.80	0.1520	0.0460	0.0272	0.0194	0.0136	0.0090	0.1682	0.0470	0.0276	0.0196	0.0136	0.0090
		0.85	0.1124	0.0350	0.0208	0.0148	0.0104	0.0070	0.1196	0.0356	0.0210	0.0150	0.0104	0.0070
		0.90	0.0870	0.0276	0.0164	0.0118	0.0082	0.0056	0.0908	0.0280	0.0166	0.0118	0.0082	0.0056

Note: A lower value is presented when there are two p_1 values satisfying $\beta_{1,s}$ or $\beta_{2,s}$.

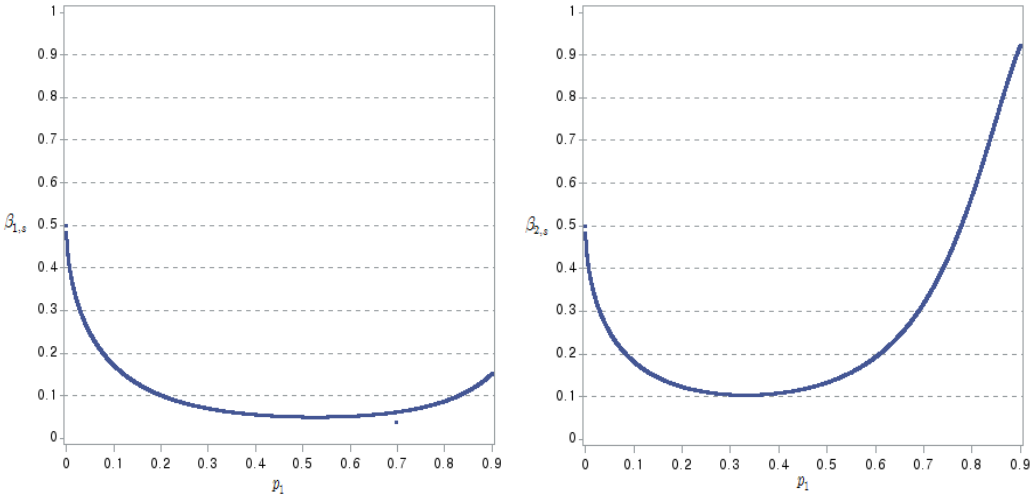


Figure 3. Scatter plot of $\beta_{1,s}$ (or $\beta_{2,s}$) vs p_1 when $N=1{,}000$, $\rho=0.9$, $\triangle_1/\sigma=0.4$ and $\triangle_{1c}/\sigma=0.6$

Another property of p_1 is the existence of outliers. The outlier is observed regardless of $\beta_{1,s}$ or $\beta_{2,s}$, especially in greater frequency when N and the effect sizes of $\triangle_1/\sigma$ and $\triangle_{1c}/\sigma$ are greater. Since this thesis focuses on the application of ρ and p_1, considering the regional type II error rate in actual clinical trial conditions, the cause of such outliers were not further investigated. Further research is needed to study the causes of the outliers. In Figure 4, outliers are highlighted in red dotted circles.

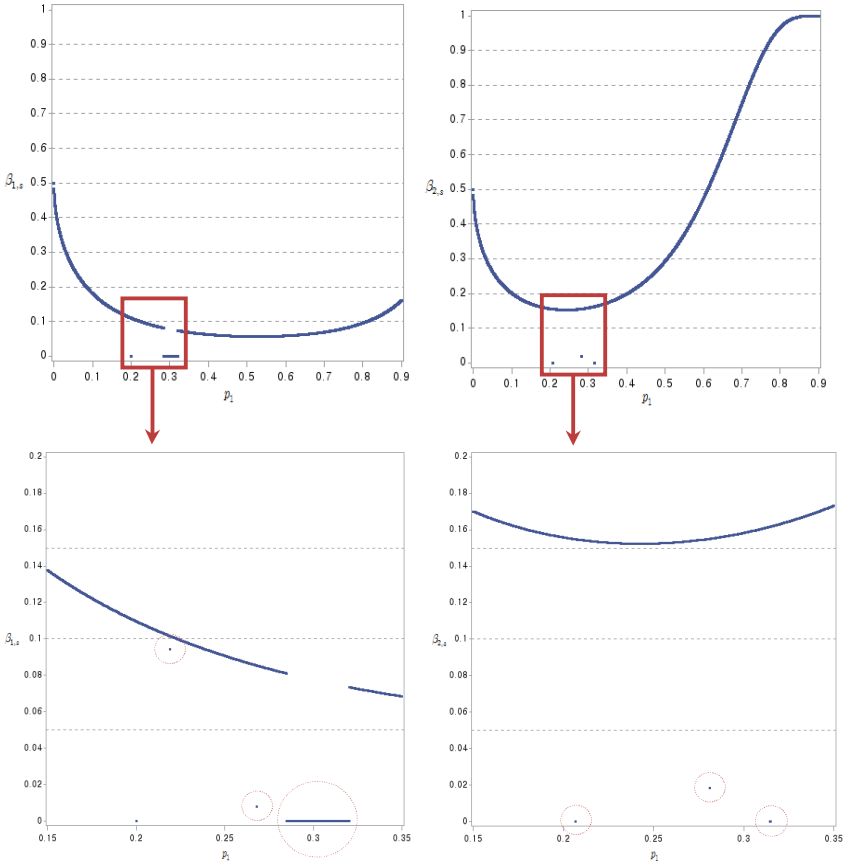


Figure 4. p_1 outliers when $N=1{,}500$, $\rho=0.9$, $\triangle_1/\sigma=0.7$, $\triangle_{1c}/\sigma=0.9$

5.4. Regional type I error rate according to the effect size

In this section, regional type I error rates $\alpha_{1,s}$ and $\alpha_{2,s}$ are calculated in relation to the changes in the effect sizes of the region of interest $\triangle_1/\sigma$ and other regions $\triangle_{1c}/\sigma$ according to the similarity criterion (i) and (ii) as well as ρ and p_1 values. Also, the changing pattern of $\alpha_{1,s}$ and $\alpha_{2,s}$ from the changes in parameter is investigated and the relationship between the regional type I error rate and the assurance probability proposed by Ko et al. (2010) is confirmed.

5.4.1. Conditions for calculating $\alpha_{1,s}$ and $\alpha_{2,s}$

Regional type I error rates $\alpha_{1,s}$ and $\alpha_{2,s}$ were obtained by changing the ranges of p_1, ρ, N, $\triangle_1/\sigma$ and $\triangle_{1c}/\sigma$ according to similarity criteria. The ranges used to calculate and are as follows. The range of the effect size was 0.1-1.5 with 0.1 unit change and the overall sample size for each group N was 100-1,000 with 100 unit change. The range of the critical value ρ for the similarity was 0.5-0.9 with 0.1 unit change and the proportion of number of patient for the clinical trial in the region of interest p_1 was 0.1-0.9 with 0.1 unit change.

[Summary of $\alpha_{1,s}$ and $\alpha_{2,s}$ calculation conditions]

(1) $\alpha = 0.025$

(2) Effect size: $0.1 \leq \triangle_1/\sigma,\ \triangle_{1c}/\sigma \leq 1.5\ (\triangle_1/\sigma \geq \triangle_{1c}/\sigma)$, 0.1 unit change

(3) $100 \leq N \leq 1{,}000$, 100 unit change

(4) $0.5 \leq \rho \leq 0.9$, 0.1 unit change

(5) $0.1 \leq p_1 \leq 0.9$, 0.1 unit change

5.4.2. $\alpha_{1,s}$ and $\alpha_{2,s}$ according to the effect size

The regional type I error rate calculated from changes in the effect sizes of the region of interest and other regions, N, ρ and p_1 are not all listed in this thesis due to limited space. However, Table 6 shows an example of $\alpha_{1,s}$ and $\alpha_{2,s}$ values for when N is 500, ρ is 0.8, 0.9, $\triangle_1/\sigma$ is 0.2-0.6, $\triangle_{1c}/\sigma$ is 0.2-0.4, and p_1 is 0.1-0.9. The first row shows the regional type I error rate according to p_1 when the significance level is 0.025, ρ is 0.8 and effect sizes $\triangle_1/\sigma$ and $\triangle_{1c}/\sigma$ are 0.2. If p_1 is 0.1, $\alpha_{1,s}$ and $\alpha_{2,s}$ are 0.42 and 0.41, respectively. Table 6 shows that in every case, $\alpha_{2,s}$ is smaller than $\alpha_{1,s}$. Especially if p_1 is small, $\alpha_{1,s}$ and $\alpha_{2,s}$ are similar but as p_1 increases $\alpha_{2,s}$ becomes much smaller than $\alpha_{1,s}$. In other words, if all other parameters are the same, respectively, as the proportion of patients for clinical trial in the region of interest p_1 increases, the regional type I error rate according to the similarity criterion (ii) decreases. Therefore, the

possibility of error resulting in the incorrect conclusion that there is a difference in treatment effects despite the similar treatment effects, is reduced when assessing the similarity for the effect size of the region of interest. In contrast, if the regional type II error rate in section 5.1 is used and all parameter conditions are the same, respectively, the similarity criterion (i) is less likely than the similarity criterion (ii) to make the wrong conclusion that the treatment effect in the region of interest is similar to that of other regions when it is not.

Table 6. The regional type I error rate when $N=500$

ρ	$\frac{\triangle_{1c}}{\sigma}$	$\frac{\triangle_1}{\sigma}$	$p_1 : \alpha_{1,s}$									$p_1 : \alpha_{2,s}$								
			0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.8	0.2	0.2	0.42	0.39	0.37	0.36	0.35	0.35	0.36	0.37	0.40	0.41	0.37	0.33	0.29	0.25	0.21	0.16	0.10	0.03
		0.3	0.25	0.18	0.14	0.12	0.11	0.11	0.12	0.15	0.21	0.24	0.16	0.11	0.08	0.06	0.04	0.03	0.01	0.00
		0.4	0.12	0.06	0.03	0.02	0.02	0.02	0.02	0.04	0.08	0.12	0.05	0.02	0.01	0.01	0.00	0.00	0.00	0.00
		0.5	0.05	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.02	0.05	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		0.6	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	0.3	0.3	0.39	0.35	0.32	0.31	0.30	0.30	0.31	0.33	0.36	0.38	0.32	0.27	0.22	0.18	0.13	0.08	0.04	0.01
		0.4	0.22	0.15	0.11	0.09	0.08	0.08	0.09	0.12	0.18	0.21	0.13	0.08	0.05	0.03	0.02	0.01	0.00	0.00
		0.5	0.10	0.04	0.02	0.01	0.01	0.01	0.01	0.03	0.07	0.10	0.04	0.01	0.01	0.00	0.00	0.00	0.00	0.00
		0.6	0.04	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.04	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	0.4	0.4	0.35	0.30	0.27	0.25	0.24	0.24	0.25	0.27	0.32	0.34	0.26	0.21	0.15	0.11	0.07	0.03	0.01	0.00
		0.5	0.19	0.12	0.08	0.07	0.06	0.06	0.07	0.09	0.15	0.18	0.10	0.05	0.03	0.02	0.01	0.00	0.00	0.00
		0.6	0.09	0.03	0.02	0.01	0.01	0.01	0.01	0.02	0.05	0.08	0.03	0.01	0.00	0.00	0.00	0.00	0.00	0.00
0.9	0.2	0.2	0.46	0.45	0.44	0.43	0.43	0.43	0.43	0.44	0.46	0.46	0.43	0.41	0.39	0.37	0.34	0.30	0.25	0.16
		0.3	0.28	0.22	0.18	0.17	0.16	0.16	0.17	0.20	0.27	0.28	0.21	0.17	0.14	0.12	0.10	0.09	0.07	0.04
		0.4	0.15	0.08	0.05	0.04	0.03	0.04	0.04	0.07	0.13	0.14	0.07	0.04	0.03	0.02	0.02	0.01	0.01	0.00
		0.5	0.06	0.02	0.01	0.01	0.00	0.00	0.01	0.01	0.05	0.06	0.02	0.01	0.00	0.00	0.00	0.00	0.00	0.00
		0.6	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	0.3	0.3	0.44	0.42	0.41	0.40	0.40	0.40	0.41	0.42	0.44	0.44	0.41	0.38	0.35	0.32	0.28	0.24	0.18	0.09
		0.4	0.27	0.20	0.17	0.15	0.14	0.14	0.16	0.19	0.25	0.26	0.19	0.14	0.11	0.09	0.07	0.05	0.04	0.01
		0.5	0.14	0.07	0.04	0.03	0.03	0.03	0.04	0.06	0.12	0.13	0.06	0.03	0.02	0.01	0.01	0.01	0.00	0.00
		0.6	0.06	0.02	0.01	0.00	0.00	0.00	0.01	0.01	0.04	0.06	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00
	0.4	0.4	0.42	0.40	0.38	0.37	0.37	0.37	0.38	0.39	0.42	0.42	0.38	0.34	0.30	0.26	0.22	0.17	0.11	0.03
		0.5	0.25	0.18	0.15	0.13	0.12	0.12	0.14	0.17	0.23	0.25	0.17	0.12	0.09	0.07	0.05	0.03	0.02	0.00
		0.6	0.13	0.06	0.04	0.03	0.02	0.02	0.03	0.05	0.11	0.12	0.05	0.03	0.02	0.01	0.01	0.00	0.00	0.00

Figures 5 and 6 show the relationship between the proportion of patients out of $2N$ in the region of interest p_1 and the regional type I error rate.

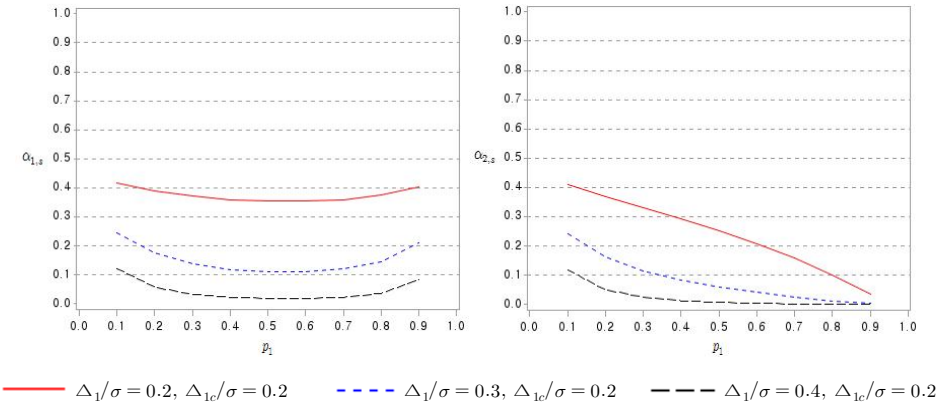


Figure 5. Graph of the regional type I error rate vs p_1 when $\rho = 0.8$ and $N = 500$

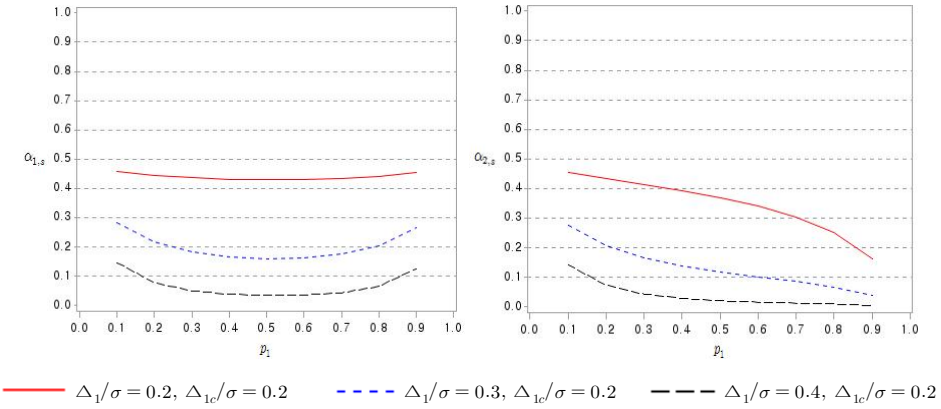


Figure 6. Graph of the regional type I error rate vs p_1 when $\rho = 0.9$ and $N = 500$

As shown in Table 6 and Figures 5 and 6, if ρ and p_1 are the same, respectively, $\alpha_{1,s}$ and $\alpha_{2,s}$ decrease as the difference between $\triangle_1/\sigma$ and $\triangle_{1c}/\sigma$ increases. On the other hand, if p_1, $\triangle_1/\sigma$ and $\triangle_{1c}/\sigma$ are the same, respectively, as ρ increases, $\alpha_{1,s}$ and $\alpha_{2,s}$ increase. In the similarity criterion (i), there is no case where $\alpha_{1,s}$ is lower than 20%, regardless of ρ, in all cases where $\triangle_1/\sigma$ and $\triangle_{1c}/\sigma$ are the same. However, in the similarity criterion (ii), as a p_1 increases, $\alpha_{2,s}$ is lower than 20% after a certain p_1. For example, if $\rho=0.8$, $\triangle_1/\sigma=\triangle_{1c}/\sigma=0.2$ and p_1 is greater than 0.62, then $\alpha_{2,s}$ is lower than 20%. Additionally, if N, ρ and p_1 are the same, respectively, the regional type I error rate decreases as $\triangle_1/\sigma=\triangle_{1c}/\sigma$ increases in both similarity criteria. Table 6 shows that when ρ is 0.8 and $\triangle_1/\sigma=\triangle_{1c}/\sigma$ is 0.2, 0.3 and 0.4, $\alpha_{1,s}$ and $\alpha_{2,s}$ values are 0.42, 0.39, and 0.35 and 0.41, 0.38, and 0.34, respectively, demonstrating a decrease.

The regional type I error rate related to the similarity criterion (ii) continuously decreases as p_1 increases. As stated by Ko et al. (2010), this makes intuitive sense, since the observed overall treatment effect D will be increasingly dominated by the observed result from the region of interest D_1, as p_1 increases. On the contrast, the regional type I error rate related to the similarity criterion (i) decreases as p_1 increases and then decreases after a certain value of p_1. As mentioned as a property of the assurance probability by Ko et al. (2010), this phenomenon arises from the fact that the observed result

from regions other than the region of interest D_{1c} is gradually dominated by D_1 at first, and is then overwhelmingly dominated by D_1 later as p_1 increases.

The method for controlling the regional type I error rate when $\triangle_1/\sigma = \triangle_{1c}/\sigma$ is the same as the assurance probability method by Ko et al. (2010). Ko et al. (2010) used the equation (16) which is the formula for calculating the sample size and applied the conversion of $\delta/(\sigma\sqrt{2/N})$ into $Z_{1-\alpha}+Z_{1-\beta}$, when deriving the assurance probability. Therefore, if α and N are the same, respectively, the effect size δ/σ for the pre-determined β can be calculated and using this effect size, $\alpha_{1,s}$ and $\alpha_{2,s}$ can be calculated to obtain the same results as $1-AP_1$ and $1-AP_2$ by Ko et al. (2010). This can be verified by using the formula for deriving the assurance probability and the formula for deriving the regional type I error rate, as shown in Appendix G.

Table 7 shows the calculated $\alpha_{1,s}$ and $\alpha_{2,s}$ using the same parameters as in Table 1 of the study by Ko et al. (2010). These values were compared to AP_1 and AP_2. In order to calculate the regional type I error rate, N was assumed to be 100 and 500, and the effect size that type II error is 20% under $N=100,\ 500$ was calculated using the equation (16). In conclusion, if the assurance probability proposed by Ko et al. (2010) is subtracted from 1, it can be confirmed that it is exactly the same as the regional type I error rate. Therefore, the method by Ko et al. (2010) is a method to control the regional type I error rate, which is limited to cases where the effect size of the

region of interest is the same as that of other regions. On the other hand, the method using the regional type I error rate can be used for cases where the effect size of the region of interest is not only the same as that of other regions but also different from that of other regions excluding the region of interest. Therefore, this method is an expansion on the method proposed by Ko et al. (2010).

Table 7. AP_1, AP_2 and $\alpha_{1,s}$, $\alpha_{2,s}$ when $\alpha=0.025$, $\beta=0.2$, $\rho=0.2$

p_1	AP_1	$1-AP_1$	AP_2	$1-AP_2$	$N=100$		$N=500$	
					$\alpha_{1,s}$	$\alpha_{2,s}$	$\alpha_{1,s}$	$\alpha_{2,s}$
0.05	0.7146	0.2854	0.7165	0.2835	0.2854	0.2835	0.2854	0.2835
0.10	0.7899	0.2101	0.7948	0.2052	0.2101	0.2052	0.2101	0.2052
0.15	0.8396	0.1604	0.8476	0.1524	0.1604	0.1524	0.1604	0.1524
0.20	0.8757	0.1243	0.8863	0.1137	0.1243	0.1137	0.1243	0.1137
0.25	0.9029	0.0971	0.9157	0.0843	0.0971	0.0843	0.0971	0.0843
0.30	0.9239	0.0761	0.9383	0.0617	0.0761	0.0617	0.0761	0.0617
0.35	0.9402	0.0598	0.9557	0.0443	0.0598	0.0443	0.0598	0.0443
0.40	0.9531	0.0469	0.9691	0.0309	0.0469	0.0309	0.0469	0.0309
0.45	0.9633	0.0367	0.9791	0.0209	0.0367	0.0209	0.0367	0.0209
0.50	0.9713	0.0287	0.9866	0.0134	0.0287	0.0134	0.0287	0.0134
0.55	0.9775	0.0225	0.9919	0.0081	0.0225	0.0081	0.0225	0.0081
0.60	0.9824	0.0176	0.9955	0.0045	0.0176	0.0045	0.0176	0.0045
0.65	0.9862	0.0138	0.9978	0.0022	0.0138	0.0022	0.0138	0.0022
0.70	0.9890	0.0110	0.9991	0.0009	0.0110	0.0009	0.0110	0.0009
0.75	0.9909	0.0091	0.9997	0.0003	0.0091	0.0003	0.0091	0.0003
0.80	0.9921	0.0079	0.9999	0.0001	0.0079	0.0001	0.0079	0.0001
0.85	0.9923	0.0077	1.0000	0.0000	0.0077	0.0000	0.0077	0.0000
0.90	0.9904	0.0096	1.0000	0.0000	0.0096	0.0000	0.0096	0.0000

Note: $\frac{\triangle_1}{\sigma}=\frac{\triangle_{1c}}{\sigma}=0.3962$ when $N=100$, $\frac{\triangle_1}{\sigma}=\frac{\triangle_{1c}}{\sigma}=0.17719$ when $N=500$

Chapter 6. Examples

6.1. Determination of ρ that satisfies the pre-determined regional type II error

A randomized, multi-center, double-blind, parallel group, placebo-controlled clinical trial is to be planned to confirm the effects of the SGLT-2 drug on type II diabetes patients with difficulties controlling their blood glucose levels. Korea, Japan, and China will participate in the trial and the efficacy will be assessed according to the changes after 24 weeks of administration based on HbA1c levels in the blood samples.

Based on the results observed from previous exploratory study, the total sample size is set to be 1,000, resulting in 500 for each group. The sample size to be allocated to Korea among 1,000 patients is planned to be 300 patients, which is 30% of the total patients. The primary purpose of this trial is to assess whether the efficacy result for Koreans is similar to that of other countries, as well as to assess the overall efficacy. The similarity between the Korean result and the result of other countries is accessed using the similarity criterion (i). To determine ρ, the effect size of HbA1c change in Korea (Δ_{1c}/σ) and the effect size of HbA1c change in Japan and China (Δ_{1c}/σ) are assumed to be 0.4 and 0.7, respectively, and the regional type II error

rate to be 20%. The obtained value of ρ using these parameters is 0.72 and this value will be used as a standard for assessing the similarity between result in Korea and that of other countries.

6.2. Determination of the sample size in the region of interest using the regional type II error

A multi-regional phase III clinical trial to assess the efficacy of inhaled corticosteroid (ICS) and long acting beta-agonist (LABA) on patients over 40 years of age with severe chronic obstructive pulmonary disease is to be planned. To assess the efficacy of an ICS + LABA combination drug, a randomized, double-blind, parallel group, placebo-controlled clinical trial will be conducted using a dry power inhaler for drug delivery. The primary endpoint is the changes in FEV_1 measured 60 minutes later after the administration of investigational product, and Five countries including Korea, Japan, China, Taiwan and Malaysia will participate in the trial. The effect size ($\triangle/\sigma$) of the difference between the treatment and placebo groups is 0.20501 in the phase II clinical trial conducted in two countries before this phase III trial. Using this result, the total sample size for the phase III clinical trial is calculated to be 1,000 (500 patients per group) at a significance level of 0.025 (one-sided) and 90% power.

The purpose of this trial is to examine whether the overall treatment effect from the MRCT can be applied to Korea, in addition

to demonstrate the overall treatment effect. Therefore, it is necessary to verify whether the overall treatment effect can be applied to Korea, under the condition that overall treatment effect is statistically significant. In this regard, the proportion of the patients recruited in Korea needs to be determined at the design phase of the trial to ensure the similarity between Korea and all regions.

For assessment, similarity criterion (ii) $D_1 > \rho D$ with $\rho = 0.9$ is used. And the effect size in Korea ($\triangle_1/\sigma$) is assumed to be 0.1, and the effect size in all countries excluding Korea ($\triangle_1/\sigma$) to be 0.25. Under these conditions, the proportion of patients recruited in Korea p_1 is 0.2516 so that the regional type II error rate will be 20%. And thus the required sample size in Korea per group is at least 500 × 0.2516 ≃ 126, corresponding to a total of 252 patients for the trial.

6.3. Determination of the sample size in the region of interest using the regional type I error

The method to determine the required sample size in the region of interest using the regional type I error rate is similar to the method proposed by Ko et al. (2010) using the assurance probability. The method by Ko et al. (2010) is a method to determine the sample size in the region of interest using p_1, where the assurance probability is over a certain value, when the treatment effect is uniform across regions. On the other hand, the method using the regional type I

error rate can calculate the sample size in the region of interest using p_1, where the regional type I error rate is lower than a certain value, even when the effect size in the region of interest is the same or greater than that of other regions.

Below, a similar example to those suggested in the published paper by Ko et al. (2010) is described. It is based on a calculation example using the regional type I error rate to determine the required sample size in the region of interest.

A randomized, double-blind, multi-regional clinical trial will be conducted in patients with hypercholesterolemia, atherosclerotic or coronary artery disease for comparing a new drug for lowering low-density-lipoprotein cholesterol (LDL-C) and a placebo control. In this trial, patients of age 18 years or older with documented LDL-C level between 2.5 mmol/L and ≤4.20 mmol/L are planned to be recruited from 3 regions including Taiwan, the United States and Europe. The primary efficacy variable is the percent change from baseline in LDL-C. The total sample size is calculated based on the results observed from previous exploratory study. The effect size of the primary endpoint (percent change in LDL-C) in the previous study was 0.228. Using this result, the total sample size is calculated to be 1,000 (500 patients per group) at a significance level of 0.025 (one-sided) and 95% power.

In addition to demonstrate an overall treatment effect from all regions, this trial is also interested in examining whether the overall

results from the MRCT can be applied to Taiwan under the condition that the overall treatment effect is statistically significant in the overall region. In this regard, the proportion of the patients recruited in Taiwan needs to be determined at the design phase of the trial to ensure the similarity between Taiwan and all regions.

If similarity criterion (ii) is used and $\rho = 0.8$, $\triangle_1/\sigma = 0.3$ and $\triangle_{1c}/\sigma = 0.2$ are chosen, then the proportion of the patients recruited in Taiwan needs to be at least 0.34 so that the regional type I error rate will be at most 10%. In this case, the required total sample size from Taiwan is around 340 (170 patients per group). On the other hand, if $\rho = 0.9$, then the required proportion of patients in Taiwan will increase to 0.6. That is, the total sample size of patients recruited from Taiwan needs to be at least 600 (300 patients per group).

Chapter 7. Conclusion and discussion

The benefits of MRCTs include the use of the same protocol in various regions to conduct clinical trials, such that the new drug can be approved simultaneously in multiple regions, thus saving time and costs. If the overall treatment effect of MRCTs is significant, it is essential to confirm that there is no ethnic difference in treatment effects. In general, the interaction between treatment and regions was tested to identify the existence of ethnic differences. However, to evaluate the interaction effects, a very large sample size is needed, and thus the numbers involved can make it unrealistic to conduct a trial (Uesaka, 2009). Currently, Japan is the only country with a regulatory body that provides a guideline on assessing consistency between regions in MRCTs which is mentioned in the 11^{th} Q&A to ICH E5. But there has been no research on the similarity criterion provided by MHLW in Japan as a decision process in terms of statistical hypothesis testing.

Ko et al. (2010) proposed the assurance probabilities based on the MHLW guideline to calculate the sample size in the region of interest. However, the assurance probability by Ko et al. (2010) is not appropriate in that it focuses on the alternative hypothesis that the overall treatment is significant, which is related to the first of two objectives in a MRCT.

This thesis introduced a method standardized by effect size, which was originally suggested by Kang et al. (2016), as a statistical hypothesis testing procedure to address the second purpose of MRCTs. This thesis also discussed approaches using the regional type II error rate to calculate critical values for a hypothesis testing on the similarity, as well as the required number of clinical trial subjects in the region of interest through the suggested method.

The results of this thesis demonstrate a difference in the regional type II error rate according to the similarity criteria. And when an effect size, a critical value and other parameters are the same, respectively, the regional type II error rate of the similarity criterion $D_1 \geq \rho D_{1c}$ is smaller than that of the similarity criterion $D_1 \geq \rho D$. In particular, if the number of patients in the region of interest in a MRCT is very large, choosing the similarity criterion $D_1 \geq \rho D_{1c}$ is a way to reduce the regional type II error rate. The regional type II error rate is easier to control when the difference between effect sizes in the region of interest and other regions is great or the critical value for similarity is great.

The selection of critical value ρ is extremely important in assessing the similarity between the data from the region of interest and in all regions. Currently, the Japanese MHLW is the only regulatory authority that provides a condition for critical value, which is over 0.5. This thesis introduced a method to determine the critical value of the similarity assessment using the statistical hypothesis testing procedure

by controlling the regional type II error rate. As a result, if the regional type II error rate is pre-determined and parameters such as the effect sizes and the proportion of patients in the region of interest are the same, respectively, the critical value for the similarity criterion $D_1 \geq \rho D$ is greater than that of $D_1 \geq \rho D_{1c}$. In other words, the second similarity criterion is more conservative in assessing the similarity than the first criterion. Furthermore, regardless of the similarity criteria, if the regional type II error rate is pre-determined and other parameters are the same, respectively, as the difference in effect size between the region of interest and other regions increases, the critical value for the similarity becomes smaller. In the method presented in this thesis, this phenomenon means the following. Suppose that the MRCT should be conducted even if there is a large difference between the region of interest and other regions which is not clinically meaningful. In this case, the critical value in the method to control the regional type II error becomes relatively small and thus increases the likelihood of proving the similarity.

As confirmed in the results of this thesis, in cases using a certain combination of parameters according to similarity criteria, the regional type II error rate is never lower than the pre-determined level, such as 20%. For example, if the effect size in the region of interest is comparable to that of other regions, the regional type II error rate is maintained at a considerably high level. In such cases where treatment effects are assumed to be similar, the regional type II error

rate cannot be maintained at low levels. And consequently, there is high likelihood of the incorrect conclusion that there is no difference in treatment effects, despite the difference in treatment effects in all participating regions and the region of interest. Even if it is evaluated that there is a statistically significant difference of the effect size despite a slight difference, this slight difference can be evaluated as the similarity of the treatment effects in actual clinical situations. Therefore, this error is not a major problem in applying the regional type II error rate.

It may seem reasonable that as the proportion of the patients in the region of interest increases, the regional type II error rate decreases. However, there is no monotone decrease in the regional type II error rate with the proportion of the patients in the region of interest. In light of this property, it may not be appropriate to use the regional type II error rate method with the similarity criteria of the Japanese MHLW and those suggested by Ko et al. (2010). That is why there is a need for new studies to develop similarity criteria that satisfy the monotonically decreasing relationship between the regional type II error rate and the proportion of the patients in the region of interest. In summary, it is important to develop and apply the best similarity criteria, in order to select the proportion of patients in the region of interest using the regional type II error rate or to determine the critical value for the similarity assessment, as suggested in this thesis.

If the effect size is uniform across regions, the method to control the regional type I error rate is the same as that using the assurance

probability method proposed by Ko et al. (2010). Since the method to control the regional type I error rate can also be used in cases where effect sizes across regions are heterogeneous, it is an expanded form of the method by Ko et al. (2010) and it can be utilized in various clinical trial environments. Under the clinical trial condition where parameters such as N, ρ and p_1 are the same, respectively, the regional type I error rate in the similarity criterion (ii) is smaller than that of the similarity criterion (i). In other words, it means that using the similarity criterion (ii) is less likely to make the wrong conclusion that the treatment effect in the region of interest is not similar to that of all regions when it is similar. Furthermore, as the proportion of the patients in the region of interest p_1 increases, the regional type I error rate for similarity criterion (ii) decreases, thus the likelihood of the wrong conclusion that there is no similarity when there is the actual similarity of the treatment effects in the region of interest, is reduced.

This thesis includes a comparison between methods to control the regional type I error and regional type II error. Regional type II error, which leads to the incorrect conclusion that there is no difference when there is a difference in treatment effects between ethnic groups, is more critical. From the perspective of regulatory authorities, an incorrect conclusion that claims the similarity when there is a difference in treatment effects between ethnic groups is more serious, because it results in drug approval. On the other hand, for new drug

developers, an incorrect conclusion that there is a difference when there is no difference in treatment effects between ethnic groups leads to a delay in market approval and additional developmental processes resulting in greater consumption of resources and time. The method suggested in this thesis to use the regional type I error and the regional type II error independently can only satisfy one party. This being so, there is a need for further studies that minimize both regional type I error and regional type II error, in order to satisfy both parties, the regulatory bodies and the pharmaceutical company.

In this thesis, the regional error rate was applied by setting the null and alternative hypotheses for the second purpose of MRCTs as "There is the similarity in treatment effects between regions" and "There is no similarity in treatment effects between regions", respectively. The alternative hypothesis that claims no similarity focuses on the aim to confirm that there is a possibility of difference between regions that may or may not exist when the similarity is predicted. But the hypothesis for the second purpose of MRCTs can be expressed in various forms. A hypothesis with the opposite concept to the secondary hypothesis proposed in this thesis can be used to assess the similarity between regions, and the regional error rate using this hypothesis can be used to determine the critical value for the similarity assessment and the number of patients in the region of interest. Under the conditions of this hypothesis, the regional type I error is more critical than the regional type II error, and thus the regional type I error should be the primary concern in assessing the

required parameters. The limitation of this thesis is not only that the properties of the regional error rate are not investigated under the alternative hypothesis that treatment effects between regions are similar, but also the results with those in this thesis are not compared. Such results of comparisons and assessments need to be confirmed through follow-up studies.

Recently, there has been a trend towards globalization in new drug development using MRCTs in Korea. As the frequency of conducting MRCTs increase, it becomes more important to determine the number of patients assigned to the region of interest, for example, Korea, as well as the critical value to assess the similarity in treatment effects between regions. Previous studies have investigated the required sample size in the region of interest in a MRCT. However, there are no studies that consider the statistical hypothesis testing procedure for the two purposes of the MRCT described in the ICH E5 guideline. Recently, the ICH has been conducting meetings with experts to provide a guideline for general principle on planning and designing MRCTs. The currently published guideline for MRCTs is a draft version for Step 2. According to this draft guideline, there is a need for a appropriate plan for sample size allocation to describe the treatment effects in the multi-regional setting, when there may be some variations in treatment effect due to different distributions of intrinsic and extrinsic factors among regions. This draft guideline also describes that there are several approaches to allocate the overall sample size to regions considering some variation in treatment effect

and one of them is described as follow (ICH E17, 2916). “One approach is to determine the sample size needed in one or more regions based on the ability to show that the region-specific treatment effect preserves some pre-specified proportion of the overall treatment effect”. Although this is a draft version, the method mentioned in ICH E17 focuses on sample size allocation in an environment where the application of similarity criteria and the treatment effects between regions are not homogeneous. Therefore, there is a need for various statistical methods for MRCT that is applicable to not only the two purposes of MRCT mentioned above, but also to a heterogeneous environment of the treatment effects between regions. The method propose in this thesis reflects the statistical hypothesis testing procedure for two purposes of MRCTs as mentioned in ICH E5, and allows calculation of the critical value for the similarity assessment, as well as the number of patients in the region of interest even when there is a difference in treatment effects between regions. Therefore, it can be said that this study considers recent topics of interest concerning MRCTs. Considering the recent increase in frequency of conducting MRCTs, there is a need for further research on various methods using regional type errors and such methods are expected to be useful in the actual environment of MRCTs.

Appendix

Appendix A : Relationship between the overall treatment effect and the regional treatment effects

Let K be the total number of participating regions and m_i be the total number of patients in the ith region ($i = 1, 2, \cdots, K$). The Difference in the treatment effects between the test product and the placebo in all participating regions ($\triangle$) can be expressed as follows.

$$\begin{aligned}
\triangle &= \mu_T - \mu_P \\
&= \frac{1}{\sum_{i=1}^{K} m_i}\left[\sum_{j=1}^{m_1}(X_{1j} - Y_{1j}) + \sum_{i=2}^{K}\sum_{j=1}^{m_i}(X_{ij} - Y_{ij})\right] \\
&= \frac{m_1}{\sum_{i=1}^{K} m_i}\frac{\sum_{j=1}^{m_1}(X_{1j} - Y_{1j})}{m_1} + \frac{\sum_{i=2}^{K} m_i}{\sum_{i=1}^{K} m_i}\frac{\sum_{i=2}^{K}\sum_{j=1}^{m_i}(X_{ij} - Y_{ij})}{\sum_{i=2}^{K} m_i} \\
&= \lambda_1 \triangle_1 + (1-\lambda_1)\triangle_{1c}
\end{aligned}$$

where

$$\lambda_1 = \frac{m_1}{\sum_{i=1}^{K} m_i} \quad , \quad \triangle_1 = \frac{\sum_{j=1}^{m_1}(X_{1j} - Y_{1j})}{m_1} \quad , \quad \triangle_{1c} = \frac{\sum_{i=2}^{K}\sum_{j=1}^{m_i}(X_{ij} - Y_{ij})}{\sum_{i=2}^{K} m_i}$$

Appendix B : The derivation of regional type II error rate

The derivation of $\beta_{1,s}$ is as follows.

$$\begin{aligned}
\beta_{1,s} &= P_\delta(D_1 \geq \rho D_{1c} \mid Z > z_{1-\alpha}, H_{sA} : \triangle_1/\triangle_{1c} < 1) \\
&= P_\delta(Z_1 \geq \rho\sqrt{\frac{p_1}{1-p_1}}\, Z_{1c} \mid \sqrt{p_1}\, Z_1 + \sqrt{1-p_1}\, Z_{1c} > z_{1-\alpha}, H_{sA}) \\
&= P_\delta(Z_1 - \sqrt{\frac{n_1}{2}}\,\frac{\triangle_1}{\sigma} \geq \rho\sqrt{\frac{p_1}{1-p_1}}\,(Z_{1c} - \sqrt{\frac{N-n_1}{2}}\,\frac{\triangle_{1c}}{\sigma}) \\
&\quad + \rho\sqrt{\frac{p_1}{1-p_1}}\,\sqrt{\frac{N-n_1}{2}}\,\frac{\triangle_{1c}}{\sigma} - \sqrt{\frac{n_1}{2}}\,\frac{\triangle_1}{\sigma} \mid \sqrt{p_1}(Z_1 - \sqrt{\frac{n_1}{2}}\,\frac{\triangle_1}{\sigma}) \\
&\quad + \sqrt{1-p_1}(Z_{1c} - \sqrt{\frac{N-n_1}{2}}\,\frac{\triangle_{1c}}{\sigma}) > z_{1-\alpha} - \sqrt{\frac{p_1 n_1}{2}}\,\frac{\triangle_1}{\sigma} \\
&\quad - \sqrt{\frac{(1-p_1)(N-n_1)}{2}}\,\frac{\triangle_{1c}}{\sigma}, H_{sA}) \\
&= P_0(Z_1 \geq \rho\sqrt{\frac{p_1}{1-p_1}}\, Z_{1c} + \sqrt{\frac{p_1 N}{2}}\,(\rho\frac{\triangle_{1c}}{\sigma} - \frac{\triangle_1}{\sigma}) \mid \sqrt{p_1}\, Z_1 + \sqrt{1-p_1}\, Z_{1c} \\
&\quad > z_{1-\alpha} - \sqrt{\frac{N}{2}}\,(p_1\frac{\triangle_1}{\sigma} + (1-p_1)\frac{\triangle_{1c}}{\sigma})) \\
&= P_0(Z_1 \geq d_1 Z_{1c} + d_2 \mid d_3 Z_1 + d_4 Z_{1c} > d_5) \\
&= \frac{P_0(Z_1 \geq d_1 Z_{1c} + d_2,\ d_3 Z_1 + d_4 Z_{1c} > d_5)}{P_0(d_3 Z_1 + d_4 Z_{1c} > z_{1-\alpha} - d_5)}
\end{aligned}$$

$$= \frac{\int_{b_1}^{\infty}[\Phi(\frac{1}{d_1}u - \frac{d_2}{d_1}) - \Phi(-\frac{d_3}{d_4}u + \frac{d_5}{d_4})]\phi(u)du}{\int_{-\infty}^{\infty}[1 - \Phi(-\frac{d_3}{d_4}u + \frac{d_5}{d_4})]\phi(u)du}$$

where

$$d_1 = \rho\sqrt{\frac{p_1}{1-p_1}} \quad , \quad d_2 = \sqrt{\frac{Np_1}{2}}\,(\rho\frac{\triangle_{1c}}{\sigma} - \frac{\triangle_1}{\sigma}) \quad , \quad d_3 = \sqrt{p_1} \quad , \quad d_4 = \sqrt{1-p_1} \quad ,$$

$$d_5 = \sqrt{\frac{N}{2}}\,(p_1\frac{\triangle_1}{\sigma} + (1-p_1)\frac{\triangle_{1c}}{\sigma}) \quad , \quad b_1 = d_2 + \frac{d_1(d_5 - d_2d_3)}{d_1d_3 + d_4}$$

The derivation of b_1 is as below.

$$Z_{1c} = \frac{1}{d_1}Z_1 - \frac{d_2}{d_1} \quad , \quad Z_{1c} = -\frac{d_3}{d_4}Z_1 + \frac{d_5}{d_4}$$

$$\frac{Z_1 - d_2}{d_1} = \frac{-d_3Z_1 + d_5}{d_4} \quad \Rightarrow \quad (d_4 + d_1d_3)Z_1 = d_1d_5 + d_2d_4$$

$$\Rightarrow \quad Z_1 = \frac{d_1d_5 + d_2d_4}{d_4 + d_1d_3} = d_2 + \frac{d_1(d_5 - d_2d_3)}{d_1d_3 + d_4} = b_1$$

Similarly, the derivation of $\beta_{2,s}$ is as follows.

$$
\begin{aligned}
\beta_{2,s} &= P_\delta(D_1 \geq \rho D \mid Z > z_{1-\alpha}, H_{sA} : \triangle_1/\triangle < 1) \\
&= P_\delta(D_1 \geq \rho[p_1 D_1 + (1-p_1) D_{1c}] \mid Z > z_{1-\alpha}, H_{sA}) \\
&= P_\delta(D_1 \geq \frac{\rho(1-p_1)}{1-\rho p_1} D_{1c} \mid Z > z_{1-\alpha}, H_{sA}) \\
&= P_\delta(Z_1 \geq \frac{\rho\sqrt{p_1(1-p_1)}}{1-\rho p_1} Z_{1c} \mid \sqrt{p_1} Z_1 + \sqrt{1-p_1} Z_{1c} > z_{1-\alpha}, H_{sA}) \\
&= P_\delta(Z_1 - \sqrt{\frac{n_1}{2}}\frac{\triangle_1}{\sigma} \geq \frac{\rho\sqrt{p_1(1-p_1)}}{1-\rho p_1}(Z_{1c} - \sqrt{\frac{N-n_1}{2}}\frac{\triangle_{1c}}{\sigma}) \\
&\quad + \frac{\rho\sqrt{p_1(1-p_1)}}{1-\rho p_1}\sqrt{\frac{N-n_1}{2}}\frac{\triangle_{1c}}{\sigma} - \sqrt{\frac{n_1}{2}}\frac{\triangle_1}{\sigma} \mid \sqrt{p_1}(Z_1 - \sqrt{\frac{n_1}{2}}\frac{\triangle_1}{\sigma}) \\
&\quad + \sqrt{1-p_1}(Z_{1c} - \sqrt{\frac{N-n_1}{2}}\frac{\triangle_{1c}}{\sigma}) > z_{1-\alpha} - \sqrt{\frac{p_1 n_1}{2}}\frac{\triangle_1}{\sigma} \\
&\quad - \sqrt{\frac{(1-p_1)(N-n_1)}{2}}\frac{\triangle_{1c}}{\sigma}, H_{sA}) \\
&= P_0(Z_1 \geq \frac{\rho\sqrt{p_1(1-p_1)}}{1-\rho p_1} Z_{1c} + \frac{\rho\sqrt{p_1(1-p_1)}}{1-\rho p_1}\sqrt{\frac{N(1-p_1)}{2}}\frac{\triangle_{1c}}{\sigma} - \sqrt{\frac{Np_1}{c}}\frac{\triangle_1}{\sigma} \\
&\qquad \mid \sqrt{p_1} Z_1 + \sqrt{1-p_1} Z_{1c} > z_{1-\alpha} - \sqrt{\frac{N}{2}}(p_1\frac{\triangle_1}{\sigma} + (1-p_1)\frac{\triangle_{1c}}{\sigma})) \\
&= P_0(Z_1 \geq d_6 Z_{1c} + d_7 \mid d_8 Z_1 + d_9 Z_{1c} > d_{10}) \\
&= \frac{P_0(Z_1 \geq d_6 Z_{1c} + d_7,\ d_8 Z_1 + d_9 Z_{1c} > d_{10})}{P_0(d_8 Z_1 + d_9 Z_{1c} > d_{10})}
\end{aligned}
$$

$$= \frac{\int_{b_2}^{\infty}[\Phi(\frac{1}{d_6}u - \frac{d_7}{d_6}) - \Phi(-\frac{d_8}{d_9}u + \frac{d_{10}}{d_9})]\phi(u)du}{\int_{-\infty}^{\infty}[1 - \Phi(-\frac{d_8}{d_9}u + \frac{d_{10}}{d_9})]\phi(u)du}$$

where

$$d_6 = \rho\frac{\sqrt{p_1(1-p_1)}}{1-\rho p_1} \quad , \quad d_7 = d_6\sqrt{\frac{N(1-p_1)}{2}}\ \frac{\triangle_{1c}}{\sigma} - \sqrt{\frac{Np_1}{2}}\ \frac{\triangle_1}{\sigma} \quad , \quad d_8 = \sqrt{p_1} \quad ,$$

$$d_9 = \sqrt{1-p_1} \quad , \quad d_{10} = \sqrt{\frac{N}{2}}\,(p_1\frac{\triangle_1}{\sigma} + (1-p_1)\frac{\triangle_{1c}}{\sigma}) \quad , \quad b_2 = d_7 + \frac{d_6(d_{10} - d_7d_8)}{d_6d_8 + d_9}$$

The derivation of b_2 is as below.

$$Z_{1c} = \frac{1}{d_6}Z_1 - \frac{d_7}{d_6} \quad , \quad Z_{1c} = -\frac{d_8}{d_9}Z_1 + \frac{d_{10}}{d_9}$$

$$\frac{Z_1 - d_7}{d_6} = \frac{-d_8Z_1 + d_{10}}{d_9} \quad \Rightarrow \quad (d_9 + d_6d_8)Z_1 = d_6d_{10} + d_7d_9$$

$$\Rightarrow \quad Z_1 = \frac{d_6d_{10} + d_7d_9}{d_9 + d_6d_8} = d_7 + \frac{d_6(d_{10} - d_7d_8)}{d_6d_8 + d_9} = b_2$$

Appendix C : Table A1. Result of the effect size through literature reviews

No	FDA approval year	Active Ingredient	Study name	Dose	Treatment - Placebo		|Effect size|
					Mean	SD	
1	2011	ARCPTA	B2336	Ind 150 mcg	0.17	0.36	0.48
				Salmeterol	0.11	0.36	0.31
			B2354	Ind 75 mcg	0.12	0.12	0.98
			B2355	Ind 75 mcg	0.14	0.24	0.58
2		BUPIVACAINE	SKY0402-C-316	300mg	-60.00	106.37	0.56
			SKY0402-C-317	120mg	-21.00	44.53	0.47
3		GABAPENTIN ENACARBIL	XP052	1200mg	-4.48	8.93	0.50
			XP053	600mg	-3.98	7.91	0.50
			XP060	1200mg	-3.11	8.49	0.37
4		CLOBAZAM	STUDY OV-1012	0.25mg	29.10	61.60	0.47
				0.5mg	35.30	67.51	0.52
				1mg	57.00	59.78	0.95
5		CELLEGESIC	REC-C-001	-	-5.00	33.90	0.15
6		VILAZODONE	CLDA-07-DP-02	40mg/day	-2.50	20.59	0.12
			GNSC-04-DP-02	40mg/day	-3.20	19.06	0.17
7	2012	LORCASERIN	STUDY 009	10mg bid	-3.70	6.21	0.60
			STUDY 010	10mg bid	-3.10	6.32	0.49
				10mg qd	-3.10	3.69	0.84
			STUDY 011	10mg bid	-3.00	7.88	0.38
				10mg qd	-1.90	7.17	0.27
8		MP29-02	MP4002	-	-2.70	4.16	0.65
			MP4004	-	-2.42	4.28	0.56
			MP4006	-	-2.16	4.35	0.50
9		LINAGLIPTIN	STUDY 46	5mg	-0.60	1.06	0.57
				2.5mg/500mg Twice	-1.30	1.07	1.22
		METFORMIN		500mg	-0.70	1.08	0.65
				1000mg Twice	-1.10	1.07	1.03
		LINAGLIPTIN &METFORMIN		2.5mg+1000mg Twice	-1.70	1.08	1.58
10		PREGABALIN	STUDY 125	-	-1.46	1.92	0.76
			STUDY 1107	-	-0.59	1.46	0.41
11		BECLOMETHASONE DIPROPIONATE	STUDY 301	320mcg	-1.00	2.01	0.50
			STUDY 302	320mcg	-0.90	2.14	0.42
			STUDY 303	320mcg	-1.00	2.25	0.44
12		PHENETERMINE	OB-301	7.5mg	-3.90	6.03	0.65
				46mg	-3.50	6.03	0.58
		TOPIRAMATE		15mg	-4.30	6.10	0.70
				92mg	-4.80	6.09	0.79
		PHENETERMINE & TOPIRAMATE		7.5mg+46mg	-7.10	6.00	1.18
				15mg+92mg	-7.70	6.04	1.27

No	FDA approval year	Active Ingredient	Study name	Dose	Treatment - Placebo		|Effect size|
					Mean	SD	
		PHENETERMINE & TOPIRAMATE	OB-302	15mg+92mg	-9.90	8.82	1.12
				3.75mg+23mg	-3.70	8.43	0.44
			OB-303	15mg+92mg	-8.60	6.20	1.39
				7.5mg+46mg	-6.50	6.32	1.03
13		METHYLPHENIDATE HCL	NWP06-ADD-100	-	-12.10	7.14	1.69
14		ACLIDINIUM BROMIDE	M33	200ug	0.09	0.20	0.42
				400ug	0.12	0.20	0.62
			M38a	200ug	0.05	0.20	0.25
				400ug	0.07	0.21	0.35
15		PANCRELIPASE	VIO16EPI07-01	-	36.50	21.44	1.70
16	2013	RIOCIGUAT	STUDY 12934	1.0-2.5mg	35.20	72.91	0.48
				1.0-1.5mg	36.70	83.50	0.44
			STUDY 11348	1.0-2.5mg	44.40	81.01	0.55
17		PAROXETINE MESYLATE	STUDY-003	7.5mg	-1.35	4.34	0.31
				7.5mg	-0.05	0.24	0.19
			STUDY 004	7.5mg	-1.42	4.17	0.34
				7.5mg	-0.03	0.23	0.14
18		CANAGLIFLOZIN	DIA3005	100mg	-0.91	0.83	1.10
				300mg	-1.17	0.83	1.41
		CANAGLIFLOZIN METFORMIN	DIA3006	100mg	-0.62	0.78	0.80
				300mg	-0.77	0.78	0.99
19		MIPOMERSEN	ISIS 301012-CS5	200mg	-21.40	18.98	1.13
			MIPO3500108	200mg	-48.40	33.18	1.46
			ISIS 301012-CS7	200mg	-33.20	24.39	1.36
			ISIS 301012-CS12	200mg	-32.40	26.01	1.25
20		OSPEMIFENE	15-50310	60mg	-0.50	1.16	0.43
					8.15	13.41	0.61
					-40.28	25.77	1.56
					-0.97	0.95	1.02
		OSPEMIFENE	15-50821	60mg	-0.26	1.13	0.23
					10.66	11.84	0.90
					-40.01	27.32	1.46
					-0.87	0.86	1.01
21		POLIDOCANOL	VAP-VV015	-	-3.31	3.60	0.92
			VAP-VV016	-	-3.53	3.54	1.00
22	2014	BELSOMRA	STUDY 28	Low	8.20	55.92	0.15
				High	20.10	57.22	0.35
			STUDY 29	Low	23.90	61.46	0.39
				High	22.60	64.16	0.35
23		CONTRAVE	NB-301	16mg	-4.60	6.65	0.69
				32mg	-4.80	6.65	0.72
			NB-302	32mg	-4.20	8.66	0.49
			NB-303	32mg	-4.60	5.99	0.77

No	FDA approval year	Active Ingredient	Study name	Dose	Treatment - Placebo		|Effect size|
					Mean	SD	
			NB-304	32mg	-3.20	4.94	0.65
24		ESBRIET	STUDY 016	1 Capsule TID	2.90	6.70	0.43
			STUDY 004	1 Capsule TID	4.40	17.52	0.25
			STUDY 006	1 Capsule TID	0.60	19.35	0.03
25		FARXIGA	D1690C00019	10mg	-0.40	1.01	0.40
26		JARDIANCE	1245.23	10mg	-0.59	0.73	0.81
				25mg	-0.62	0.80	0.77
				10mg	-0.62	0.75	0.83
				25mg	-0.59	0.74	0.79
			1245.19	10mg	-0.43	0.97	0.45
				25mg	-0.56	0.97	0.58
			1245.20	10mg	-0.72	0.75	0.96
				25mg	-0.83	0.83	1.00
27		TANZEUM	GLP112755	30mg	-0.76	0.87	0.88
28		TARGINIQ ER	OUN3701	OXN	-0.50	1.73	0.29
29		TIVORBEX	IND3_08_04b	40mg TID	442.00	887.14	0.50
				40mg BID	260.00	884.68	0.29
				20mg TID	313.00	884.68	0.35
			IND3-10-06	40mg TID	318.00	1027.71	0.31
				40mg BID	342.00	1024.99	0.33
				20mg TID	62.00	1027.02	0.06
30		TRULICITY	GBCF_GBDF	0.75mg	-1.04	0.98	1.06
				1.5mg	-1.23	0.92	1.33
		SITAGLPTIN		-	-0.63	0.97	0.65
		TRULICITY	GBCF_GBDA	0.75mg	-0.84	0.97	0.87
				1.5mg	-1.05	0.97	1.08
		EXENATIDE			-0.53	0.97	0.55
31		XARTEMIS XR	STUDY 0182	OC/APAP 7.5mg/325mg	48.00	86.16	0.56

Appendix D : Table A2. ρ selection when N=500

β_{1s} or $\beta_{2,s}$	$\frac{\triangle_1}{\sigma}$	$\frac{\triangle_{1c}}{\sigma}$	p_1 : Similarity criteria I										p_1 : Similarity criteria II									
			0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
0.10	0.40	0.60				1.00	0.97	0.95	0.94	0.93	0.92	0.92				1.00	0.98	0.97	0.96	0.96	0.96	0.96
		0.65			0.96	0.92	0.89	0.87	0.86	0.85	0.85	0.84			0.96	0.93	0.92	0.91	0.90	0.91	0.91	0.91
		0.70		0.95	0.88	0.85	0.82	0.81	0.79	0.79	0.78	0.77		0.95	0.90	0.87	0.86	0.86	0.86	0.86	0.87	0.87
		0.75		0.88	0.82	0.79	0.77	0.75	0.74	0.73	0.72	0.72		0.89	0.85	0.82	0.81	0.81	0.81	0.82	0.83	0.84
		0.80	0.95	0.83	0.77	0.74	0.72	0.70	0.69	0.68	0.67	0.67	0.96	0.84	0.80	0.78	0.77	0.77	0.77	0.78	0.79	0.80
		0.85	0.90	0.78	0.72	0.69	0.67	0.66	0.65	0.64	0.63	0.63	0.90	0.79	0.75	0.74	0.73	0.73	0.74	0.75	0.76	0.77
		0.90	0.85	0.73	0.68	0.65	0.63	0.62	0.61	0.60	0.59	0.59	0.85	0.75	0.72	0.70	0.70	0.70	0.71	0.71	0.73	0.74
	0.45	0.65				1.00	0.97	0.96	0.94	0.94	0.93	0.93				1.00	0.98	0.97	0.96	0.96	0.96	0.96
		0.70			0.96	0.92	0.90	0.88	0.87	0.86	0.86	0.85			0.97	0.94	0.92	0.92	0.91	0.91	0.92	0.92
		0.75		0.95	0.89	0.86	0.84	0.82	0.81	0.80	0.80	0.79		0.96	0.91	0.88	0.87	0.87	0.87	0.87	0.88	0.88
		0.80		0.89	0.83	0.80	0.78	0.77	0.75	0.75	0.74	0.74		0.90	0.86	0.84	0.83	0.82	0.83	0.83	0.84	0.85
		0.85	0.96	0.84	0.78	0.75	0.73	0.72	0.71	0.70	0.69	0.69	0.96	0.85	0.81	0.79	0.79	0.78	0.79	0.80	0.81	0.82
		0.90	0.90	0.79	0.74	0.71	0.69	0.68	0.67	0.66	0.65	0.65	0.91	0.81	0.77	0.75	0.75	0.75	0.75	0.76	0.77	0.79
	0.50	0.70				1.00	0.98	0.96	0.95	0.94	0.94	0.93				1.00	0.98	0.97	0.97	0.96	0.96	0.97
		0.75			0.96	0.93	0.91	0.89	0.88	0.87	0.87	0.87			0.97	0.94	0.93	0.92	0.92	0.92	0.92	0.93
		0.80		0.95	0.90	0.87	0.85	0.83	0.82	0.81	0.81	0.81		0.96	0.91	0.89	0.88	0.88	0.88	0.88	0.89	0.89
		0.85		0.90	0.84	0.81	0.79	0.78	0.77	0.76	0.76	0.75		0.91	0.86	0.85	0.84	0.84	0.84	0.84	0.85	0.86
		0.90	0.96	0.85	0.80	0.77	0.75	0.73	0.73	0.72	0.71	0.71	0.96	0.86	0.82	0.81	0.80	0.80	0.80	0.81	0.82	0.83
	0.55	0.75				1.00	0.98	0.96	0.95	0.95	0.94	0.94				1.00	0.98	0.97	0.97	0.97	0.97	0.97
		0.80			0.96	0.93	0.91	0.90	0.89	0.88	0.88	0.87			0.97	0.95	0.93	0.93	0.92	0.93	0.93	0.93
		0.85		0.96	0.91	0.88	0.86	0.84	0.83	0.83	0.82	0.82		0.96	0.92	0.90	0.89	0.88	0.88	0.89	0.89	0.90
		0.90		0.90	0.85	0.83	0.81	0.79	0.78	0.78	0.77	0.77		0.91	0.87	0.86	0.85	0.85	0.85	0.85	0.86	0.87
	0.60	0.80				1.00	0.98	0.96	0.96	0.95	0.95	0.94				1.00	0.98	0.98	0.97	0.97	0.97	0.97
		0.85			0.97	0.94	0.92	0.90	0.90	0.89	0.88	0.88			0.97	0.95	0.94	0.93	0.93	0.93	0.93	0.94
		0.90		0.96	0.91	0.88	0.86	0.85	0.84	0.84	0.83	0.83		0.96	0.92	0.90	0.90	0.89	0.89	0.90	0.90	0.91
	0.65	0.85				1.00	0.98	0.97	0.96	0.95	0.95	0.95				1.00	0.99	0.98	0.97	0.97	0.97	0.97
		0.90			0.97	0.94	0.92	0.91	0.90	0.90	0.89	0.89			0.97	0.95	0.94	0.94	0.93	0.94	0.94	0.94
	0.70	*0.90*				1.00	0.98	0.97	0.96	0.96	0.95	0.95				1.00	0.99	0.98	0.97	0.97	0.97	0.98
0.15	0.40	0.55					1.00	0.98	0.97	0.96	0.96	0.96					1.00	0.99	0.98	0.98	0.98	0.98
		0.60			0.97	0.93	0.91	0.89	0.88	0.88	0.87	0.87			0.97	0.94	0.93	0.92	0.92	0.92	0.92	0.93
		0.65		0.94	0.89	0.86	0.84	0.82	0.81	0.80	0.80	0.79		0.95	0.90	0.88	0.87	0.87	0.87	0.87	0.88	0.89
		0.70	0.99	0.87	0.82	0.79	0.77	0.76	0.75	0.74	0.74	0.73	0.99	0.89	0.85	0.83	0.82	0.82	0.82	0.83	0.84	0.85
		0.75	0.93	0.81	0.77	0.74	0.72	0.71	0.70	0.69	0.68	0.68	0.93	0.83	0.79	0.78	0.77	0.77	0.78	0.79	0.80	0.81
		0.80	0.87	0.76	0.72	0.69	0.67	0.66	0.65	0.64	0.64	0.64	0.87	0.78	0.75	0.74	0.73	0.74	0.74	0.75	0.76	0.78

β_{1s} or $\beta_{2,s}$	$\frac{\triangle_1}{\sigma}$	$\frac{\triangle_{1c}}{\sigma}$	p_1 : Similarity criteria I										p_1 : Similarity criteria II									
			0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
		0.85	0.82	0.72	0.67	0.65	0.63	0.62	0.61	0.60	0.60	0.60	0.82	0.74	0.71	0.70	0.70	0.70	0.71	0.72	0.73	0.75
		0.90	0.77	0.68	0.64	0.61	0.60	0.58	0.58	0.57	0.56	0.56	0.78	0.70	0.67	0.66	0.66	0.67	0.68	0.69	0.70	0.72
	0.45	0.60					1.00	0.98	0.97	0.97	0.96	0.96					1.00	0.99	0.98	0.98	0.98	0.98
		0.65			0.97	0.94	0.92	0.90	0.89	0.89	0.88	0.88			0.97	0.95	0.94	0.93	0.93	0.93	0.93	0.94
		0.70		0.95	0.90	0.87	0.85	0.83	0.83	0.82	0.81	0.81		0.95	0.91	0.89	0.88	0.88	0.88	0.88	0.89	0.90
		0.75	0.99	0.88	0.84	0.81	0.79	0.78	0.77	0.76	0.76	0.75	0.99	0.89	0.86	0.84	0.83	0.83	0.84	0.84	0.85	0.86
		0.80	0.93	0.83	0.78	0.76	0.74	0.73	0.72	0.71	0.71	0.70	0.93	0.84	0.81	0.79	0.79	0.79	0.80	0.80	0.81	0.83
		0.85	0.87	0.78	0.73	0.71	0.69	0.68	0.67	0.67	0.66	0.66	0.88	0.79	0.77	0.75	0.75	0.75	0.76	0.77	0.78	0.79
		0.90	0.83	0.73	0.69	0.67	0.65	0.64	0.63	0.63	0.62	0.62	0.83	0.75	0.73	0.72	0.72	0.72	0.73	0.74	0.75	0.77
	0.50	0.65					1.00	0.98	0.98	0.97	0.97	0.96					1.00	0.99	0.98	0.98	0.98	0.98
		0.70			0.97	0.94	0.92	0.91	0.90	0.89	0.89	0.89			0.98	0.95	0.94	0.94	0.93	0.93	0.94	0.94
		0.75		0.95	0.90	0.88	0.86	0.85	0.84	0.83	0.83	0.82		0.96	0.92	0.90	0.89	0.89	0.89	0.89	0.90	0.90
		0.80	0.99	0.89	0.85	0.82	0.80	0.79	0.78	0.78	0.77	0.77	0.99	0.90	0.87	0.85	0.84	0.84	0.85	0.85	0.86	0.87
		0.85	0.93	0.84	0.79	0.77	0.75	0.74	0.73	0.73	0.72	0.72	0.94	0.85	0.82	0.81	0.80	0.80	0.81	0.82	0.83	0.84
		0.90	0.88	0.79	0.75	0.73	0.71	0.70	0.69	0.69	0.68	0.68	0.89	0.81	0.78	0.77	0.77	0.77	0.78	0.78	0.80	0.81
	0.55	0.70					1.00	0.99	0.98	0.97	0.97	0.97					1.00	0.99	0.99	0.98	0.98	0.98
		0.75			0.97	0.95	0.93	0.92	0.91	0.90	0.90	0.90			0.98	0.96	0.95	0.94	0.94	0.94	0.94	0.95
		0.80		0.95	0.91	0.88	0.87	0.86	0.85	0.84	0.84	0.84		0.96	0.92	0.91	0.90	0.90	0.90	0.90	0.90	0.91
		0.85	0.99	0.90	0.86	0.83	0.79	0.80	0.80	0.79	0.79	0.78	0.99	0.91	0.87	0.86	0.85	0.85	0.86	0.86	0.87	0.88
		0.90	0.94	0.85	0.81	0.78	0.77	0.76	0.75	0.74	0.74	0.74	0.94	0.86	0.83	0.82	0.82	0.82	0.82	0.83	0.84	0.85
	0.60	0.75					1.00	0.99	0.98	0.97	0.97	0.97					1.00	0.99	0.99	0.98	0.98	0.99
		0.80			0.97	0.95	0.93	0.92	0.91	0.91	0.91	0.90			0.98	0.96	0.95	0.94	0.94	0.94	0.95	0.95
		0.85		0.96	0.92	0.89	0.88	0.87	0.86	0.85	0.85	0.85		0.96	0.93	0.91	0.90	0.90	0.90	0.91	0.91	0.92
		0.90	0.99	0.90	0.86	0.84	0.83	0.82	0.81	0.80	0.80	0.80	1.00	0.91	0.88	0.87	0.86	0.86	0.87	0.87	0.88	0.89
	0.65	0.80					1.00	0.99	0.98	0.98	0.97	0.97					1.00	0.99	0.99	0.99	0.99	0.99
		0.85			0.98	0.95	0.94	0.93	0.92	0.91	0.91	0.91			0.98	0.96	0.95	0.95	0.95	0.95	0.95	0.95
		0.90		0.96	0.92	0.90	0.88	0.87	0.87	0.86	0.86	0.86		0.96	0.93	0.92	0.91	0.91	0.91	0.91	0.92	0.92
	0.70	0.85					1.00	0.99	0.98	0.98	0.97	0.97					1.00	0.99	0.99	0.99	0.99	0.99
		0.90			0.98	0.96	0.94	0.93	0.92	0.92	0.92	0.92			0.98	0.96	0.96	0.95	0.95	0.95	0.95	0.96
	0.75	*0.90*					1.00	0.99	0.98	0.98	0.98	0.98					1.00	0.99	0.99	0.99	0.99	0.99
0.20	0.40	0.55			0.99	0.96	0.94	0.93	0.92	0.92	0.91	0.91			0.99	0.97	0.96	0.95	0.95	0.95	0.95	0.95
		0.60		0.96	0.91	0.88	0.86	0.85	0.84	0.83	0.83	0.83		0.96	0.92	0.90	0.89	0.89	0.89	0.89	0.90	0.91
		0.65	0.98	0.88	0.84	0.81	0.79	0.78	0.77	0.77	0.76	0.76	0.98	0.89	0.86	0.84	0.84	0.84	0.84	0.85	0.85	0.86
		0.70	0.91	0.82	0.77	0.75	0.73	0.72	0.71	0.71	0.70	0.70	0.92	0.83	0.80	0.79	0.79	0.79	0.79	0.80	0.81	0.82
		0.75	0.85	0.76	0.72	0.70	0.68	0.67	0.66	0.66	0.65	0.65	0.86	0.78	0.75	0.74	0.74	0.75	0.75	0.76	0.78	0.79
		0.80	0.80	0.71	0.68	0.65	0.64	0.63	0.62	0.62	0.61	0.61	0.81	0.73	0.71	0.70	0.70	0.71	0.72	0.73	0.74	0.76

β_{1s} or $\beta_{2,s}$	$\frac{\triangle_1}{\sigma}$	$\frac{\triangle_{1c}}{\sigma}$	p_1 : Similarity criteria I										p_1 : Similarity criteria II									
			0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
		0.85	0.75	0.67	0.64	0.61	0.60	0.59	0.58	0.58	0.57	0.57	0.76	0.69	0.67	0.67	0.67	0.67	0.68	0.70	0.71	0.73
		0.90	0.71	0.63	0.60	0.58	0.57	0.56	0.55	0.55	0.54	0.54	0.72	0.66	0.64	0.63	0.64	0.64	0.65	0.67	0.68	0.70
	0.45	0.60			0.99	0.97	0.95	0.94	0.93	0.92	0.92	0.92			1.00	0.97	0.96	0.96	0.95	0.95	0.95	0.96
		0.65		0.96	0.92	0.89	0.87	0.86	0.85	0.85	0.84	0.84		0.96	0.93	0.91	0.90	0.90	0.90	0.90	0.91	0.91
		0.70	0.99	0.89	0.85	0.82	0.81	0.80	0.79	0.78	0.78	0.78	0.99	0.90	0.87	0.85	0.85	0.85	0.85	0.86	0.87	0.87
		0.75	0.92	0.83	0.79	0.77	0.75	0.74	0.73	0.73	0.72	0.72	0.92	0.84	0.82	0.80	0.80	0.80	0.81	0.82	0.83	0.84
		0.80	0.86	0.78	0.74	0.72	0.70	0.69	0.69	0.68	0.68	0.67	0.87	0.79	0.77	0.76	0.76	0.76	0.77	0.78	0.79	0.81
		0.85	0.81	0.73	0.70	0.68	0.66	0.65	0.65	0.64	0.64	0.63	0.82	0.75	0.73	0.72	0.72	0.73	0.74	0.75	0.76	0.78
		0.90	0.76	0.69	0.66	0.64	0.62	0.61	0.61	0.60	0.60	0.60	0.77	0.71	0.69	0.69	0.69	0.70	0.70	0.72	0.73	0.75
	0.50	0.65			0.99	0.97	0.95	0.94	0.93	0.93	0.93	0.92			1.00	0.98	0.96	0.96	0.96	0.96	0.96	0.96
		0.70		0.96	0.92	0.90	0.88	0.87	0.86	0.86	0.86	0.85		0.97	0.93	0.92	0.91	0.91	0.91	0.91	0.92	0.92
		0.75	0.99	0.90	0.86	0.84	0.82	0.81	0.80	0.80	0.80	0.79	0.99	0.91	0.88	0.86	0.86	0.86	0.86	0.87	0.88	0.88
		0.80	0.92	0.84	0.80	0.78	0.77	0.76	0.75	0.75	0.74	0.74	0.93	0.85	0.83	0.82	0.82	0.82	0.82	0.83	0.84	0.85
		0.85	0.87	0.79	0.76	0.74	0.72	0.71	0.71	0.70	0.70	0.69	0.87	0.81	0.78	0.78	0.78	0.78	0.79	0.80	0.81	0.82
		0.90	0.82	0.75	0.71	0.69	0.68	0.67	0.67	0.66	0.66	0.65	0.83	0.76	0.74	0.74	0.74	0.75	0.75	0.76	0.78	0.79
	0.55	0.70			0.99	0.97	0.96	0.95	0.94	0.93	0.93	0.93			1.00	0.98	0.97	0.96	0.96	0.96	0.96	0.96
		0.75		0.97	0.93	0.90	0.89	0.88	0.87	0.87	0.87	0.86		0.97	0.94	0.92	0.92	0.91	0.91	0.92	0.92	0.93
		0.80	0.99	0.90	0.87	0.85	0.83	0.82	0.82	0.81	0.81	0.81	0.99	0.91	0.89	0.87	0.87	0.87	0.87	0.88	0.89	0.89
		0.85	0.93	0.85	0.82	0.80	0.78	0.77	0.77	0.76	0.76	0.76	0.93	0.86	0.84	0.83	0.83	0.83	0.84	0.84	0.85	0.86
		0.90	0.88	0.80	0.77	0.75	0.74	0.73	0.72	0.72	0.71	0.71	0.88	0.82	0.80	0.79	0.79	0.79	0.80	0.81	0.82	0.83
	0.60	0.75			1.00	0.97	0.96	0.95	0.94	0.94	0.94	0.94			1.00	0.98	0.97	0.96	0.96	0.96	0.96	0.97
		0.80		0.97	0.93	0.91	0.90	0.89	0.88	0.88	0.87	0.87		0.97	0.94	0.93	0.92	0.92	0.92	0.92	0.93	0.93
		0.85	0.99	0.91	0.88	0.86	0.84	0.83	0.83	0.82	0.82	0.82	0.99	0.92	0.89	0.88	0.88	0.88	0.88	0.89	0.89	0.90
		0.90	0.93	0.86	0.83	0.81	0.80	0.79	0.78	0.78	0.77	0.77	0.94	0.87	0.85	0.84	0.84	0.84	0.85	0.85	0.86	0.87
	0.65	0.80			1.00	0.98	0.96	0.95	0.95	0.94	0.94	0.94			1.00	0.98	0.97	0.97	0.97	0.97	0.97	0.97
		0.85		0.97	0.94	0.92	0.90	0.90	0.89	0.89	0.88	0.88		0.97	0.94	0.93	0.93	0.92	0.93	0.93	0.93	0.94
		0.90	0.99	0.91	0.88	0.86	0.85	0.84	0.84	0.83	0.83	0.83	0.99	0.92	0.90	0.89	0.89	0.89	0.89	0.89	0.90	0.91
	0.70	0.85			1.00	0.98	0.96	0.96	0.95	0.95	0.94	0.94			1.00	0.98	0.97	0.97	0.97	0.97	0.97	0.97
		0.90		0.97	0.94	0.92	0.91	0.90	0.90	0.89	0.89	0.89		0.97	0.95	0.94	0.93	0.93	0.93	0.93	0.94	0.94
	0.75	*0.90*			1.00	0.98	0.97	0.96	0.95	0.95	0.95	0.95			1.00	0.98	0.98	0.97	0.97	0.97	0.97	0.97
0.25	0.40	0.50					0.99	0.98	0.97	0.97	0.97	0.96					1.00	0.99	0.98	0.98	0.98	0.98
		0.55		0.98	0.94	0.91	0.90	0.89	0.88	0.88	0.87	0.87		0.98	0.95	0.93	0.92	0.92	0.92	0.92	0.93	0.93
		0.60	0.99	0.90	0.86	0.84	0.82	0.81	0.80	0.80	0.80	0.79	0.99	0.91	0.88	0.86	0.86	0.86	0.86	0.87	0.88	0.88
		0.65	0.91	0.83	0.79	0.77	0.76	0.75	0.74	0.73	0.73	0.73	0.91	0.84	0.82	0.81	0.81	0.81	0.81	0.82	0.83	0.84
		0.70	0.84	0.77	0.73	0.71	0.70	0.69	0.69	0.68	0.68	0.67	0.85	0.79	0.76	0.76	0.76	0.76	0.77	0.78	0.79	0.81
		0.75	0.79	0.71	0.68	0.67	0.65	0.64	0.64	0.63	0.63	0.63	0.80	0.74	0.72	0.71	0.72	0.72	0.73	0.74	0.76	0.77

β_{1s} or β_{2s}	$\frac{\triangle_1}{\sigma}$	$\frac{\triangle_{1c}}{\sigma}$	p_1 : Similarity criteria I										p_1 : Similarity criteria II									
			0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
		0.80	0.74	0.67	0.64	0.62	0.61	0.60	0.60	0.59	0.59	0.59	0.75	0.69	0.68	0.67	0.68	0.68	0.70	0.71	0.72	0.74
		0.85	0.69	0.63	0.60	0.59	0.57	0.57	0.56	0.56	0.55	0.55	0.70	0.65	0.64	0.64	0.64	0.65	0.66	0.68	0.69	0.71
		0.90	0.65	0.59	0.57	0.55	0.54	0.53	0.53	0.53	0.52	0.52	0.67	0.62	0.61	0.61	0.61	0.62	0.63	0.65	0.66	0.68
	0.45	0.55					0.99	0.98	0.98	0.97	0.97	0.97					1.00	0.99	0.99	0.98	0.98	0.98
		0.60		0.98	0.94	0.92	0.91	0.90	0.89	0.89	0.88	0.88		0.98	0.95	0.94	0.93	0.93	0.93	0.93	0.93	0.94
		0.65	0.99	0.91	0.87	0.85	0.84	0.83	0.82	0.82	0.81	0.81	0.99	0.91	0.89	0.88	0.87	0.87	0.88	0.88	0.89	0.90
		0.70	0.92	0.84	0.81	0.79	0.77	0.77	0.76	0.75	0.75	0.75	0.92	0.85	0.83	0.82	0.82	0.82	0.83	0.84	0.85	0.86
		0.75	0.85	0.78	0.75	0.73	0.72	0.71	0.71	0.70	0.70	0.70	0.86	0.80	0.78	0.77	0.78	0.78	0.79	0.80	0.81	0.82
		0.80	0.80	0.73	0.70	0.69	0.68	0.67	0.66	0.66	0.65	0.65	0.81	0.75	0.74	0.73	0.74	0.74	0.75	0.76	0.77	0.79
		0.85	0.75	0.69	0.66	0.65	0.63	0.63	0.62	0.62	0.61	0.61	0.76	0.71	0.70	0.69	0.70	0.71	0.72	0.73	0.74	0.76
		0.90	0.71	0.65	0.62	0.61	0.60	0.59	0.59	0.58	0.58	0.58	0.72	0.67	0.66	0.66	0.67	0.67	0.69	0.70	0.71	0.73
	0.50	0.60					0.99	0.99	0.98	0.97	0.97	0.97					1.00	0.99	0.99	0.98	0.98	0.99
		0.65		0.98	0.95	0.93	0.92	0.91	0.90	0.90	0.89	0.89		0.99	0.96	0.94	0.94	0.93	0.93	0.94	0.94	0.94
		0.70	0.99	0.91	0.88	0.86	0.85	0.84	0.83	0.83	0.83	0.82	0.99	0.92	0.90	0.89	0.88	0.88	0.89	0.89	0.90	0.90
		0.75	0.92	0.85	0.82	0.80	0.79	0.78	0.78	0.77	0.77	0.77	0.93	0.86	0.84	0.83	0.83	0.84	0.84	0.85	0.86	0.87
		0.80	0.86	0.80	0.77	0.75	0.74	0.73	0.73	0.72	0.72	0.72	0.87	0.81	0.80	0.79	0.79	0.80	0.80	0.81	0.82	0.84
		0.85	0.81	0.75	0.72	0.71	0.69	0.69	0.68	0.68	0.67	0.67	0.82	0.77	0.75	0.75	0.75	0.76	0.77	0.78	0.79	0.80
		0.90	0.77	0.71	0.68	0.67	0.66	0.65	0.64	0.64	0.64	0.63	0.78	0.73	0.72	0.71	0.72	0.72	0.74	0.75	0.76	0.78
	0.55	0.65					0.99	0.99	0.98	0.98	0.97	0.97					1.00	0.99	0.99	0.99	0.99	0.99
		0.70		0.98	0.95	0.93	0.92	0.91	0.91	0.90	0.90	0.90		0.99	0.96	0.95	0.94	0.94	0.94	0.94	0.94	0.95
		0.75	0.99	0.92	0.89	0.87	0.86	0.85	0.84	0.84	0.84	0.84	0.99	0.93	0.90	0.89	0.89	0.89	0.89	0.90	0.90	0.91
		0.80	0.93	0.86	0.83	0.81	0.80	0.80	0.79	0.79	0.78	0.78	0.93	0.87	0.85	0.85	0.85	0.85	0.85	0.86	0.87	0.88
		0.85	0.87	0.81	0.78	0.77	0.75	0.75	0.74	0.74	0.74	0.73	0.88	0.82	0.81	0.80	0.80	0.81	0.82	0.83	0.84	0.85
		0.90	0.82	0.76	0.74	0.72	0.71	0.71	0.70	0.70	0.69	0.69	0.83	0.78	0.77	0.76	0.77	0.77	0.78	0.79	0.80	0.82
	0.60	0.70					1.00	0.99	0.98	0.98	0.98	0.98					1.00	0.99	0.99	0.99	0.99	0.99
		0.75		0.99	0.96	0.94	0.93	0.92	0.91	0.91	0.91	0.91		0.99	0.96	0.95	0.94	0.94	0.94	0.94	0.95	0.95
		0.80	0.99	0.92	0.89	0.88	0.87	0.86	0.85	0.85	0.85	0.85	0.99	0.93	0.91	0.90	0.90	0.90	0.90	0.91	0.91	0.92
		0.85	0.93	0.87	0.84	0.83	0.82	0.81	0.80	0.80	0.80	0.80	0.93	0.88	0.86	0.86	0.85	0.86	0.86	0.87	0.88	0.89
		0.90	0.88	0.82	0.79	0.78	0.77	0.76	0.76	0.75	0.75	0.75	0.88	0.83	0.82	0.81	0.82	0.82	0.83	0.84	0.85	0.86
	0.65	0.75					1.00	0.99	0.98	0.98	0.98	0.98					1.00	0.99	0.99	0.99	0.99	0.99
		0.80		0.99	0.96	0.94	0.93	0.92	0.92	0.92	0.91	0.91		0.99	0.96	0.95	0.95	0.95	0.95	0.95	0.95	0.96
		0.85	0.99	0.93	0.90	0.89	0.88	0.87	0.86	0.86	0.86	0.86	0.99	0.93	0.91	0.91	0.90	0.90	0.91	0.91	0.92	0.92
		0.90	0.93	0.88	0.85	0.84	0.83	0.82	0.81	0.81	0.81	0.81	0.94	0.89	0.87	0.86	0.86	0.87	0.87	0.88	0.89	0.89
	0.70	0.80					1.00	0.99	0.98	0.98	0.98	0.98					1.00	0.99	0.99	0.99	0.99	0.99
		0.85		0.99	0.96	0.95	0.94	0.93	0.92	0.92	0.92	0.92		0.99	0.97	0.96	0.95	0.95	0.95	0.95	0.95	0.96
		0.90	0.99	0.93	0.91	0.89	0.88	0.88	0.87	0.87	0.87	0.87	0.99	0.94	0.92	0.91	0.91	0.91	0.91	0.92	0.92	0.93

β_{1s} or $\beta_{2,s}$	$\frac{\triangle_1}{\sigma}$	$\frac{\triangle_{1c}}{\sigma}$	p_1 : Similarity criteria I										p_1 : Similarity criteria II									
			0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
	0.75	0.85					1.00	0.99	0.99	0.98	0.98	0.98					1.00	0.99	0.99	0.99	0.99	0.99
		0.90		0.99	0.96	0.95	0.94	0.93	0.93	0.93	0.92	0.92		0.99	0.97	0.96	0.95	0.95	0.95	0.95	0.96	0.96
	0.80	0.90					1.00	0.99	0.99	0.98	0.98	0.98					1.00	0.99	0.99	0.99	0.99	0.99
0.30	0.40	0.50			0.98	0.96	0.95	0.94	0.93	0.93	0.93	0.93			0.98	0.97	0.96	0.96	0.96	0.96	0.96	0.96
		0.55	1.00	0.92	0.89	0.87	0.86	0.85	0.84	0.84	0.84	0.84	1.00	0.93	0.90	0.89	0.89	0.89	0.89	0.90	0.90	0.91
		0.60	0.91	0.84	0.81	0.80	0.79	0.78	0.77	0.77	0.77	0.76	0.92	0.86	0.84	0.83	0.83	0.83	0.84	0.85	0.86	0.87
		0.65	0.84	0.78	0.75	0.73	0.72	0.72	0.71	0.71	0.70	0.70	0.85	0.80	0.78	0.78	0.78	0.78	0.79	0.80	0.81	0.83
		0.70	0.78	0.72	0.70	0.68	0.67	0.66	0.66	0.66	0.65	0.65	0.79	0.74	0.73	0.73	0.73	0.74	0.75	0.76	0.77	0.79
		0.75	0.73	0.67	0.65	0.64	0.63	0.62	0.61	0.61	0.61	0.61	0.74	0.70	0.69	0.69	0.69	0.70	0.71	0.72	0.74	0.75
		0.80	0.68	0.63	0.61	0.60	0.59	0.58	0.57	0.57	0.57	0.57	0.69	0.66	0.65	0.65	0.65	0.66	0.68	0.69	0.71	0.72
		0.85	0.64	0.59	0.57	0.56	0.55	0.55	0.54	0.54	0.53	0.53	0.65	0.62	0.61	0.61	0.62	0.63	0.64	0.66	0.68	0.69
		0.90	0.61	0.56	0.54	0.53	0.52	0.51	0.51	0.51	0.50	0.50	0.62	0.59	0.58	0.58	0.59	0.60	0.62	0.63	0.65	0.67
	0.45	0.55			0.98	0.96	0.95	0.95	0.94	0.94	0.93	0.93			0.99	0.97	0.96	0.96	0.96	0.96	0.96	0.97
		0.60	1.00	0.93	0.90	0.88	0.87	0.86	0.86	0.85	0.85	0.85	1.00	0.94	0.91	0.90	0.90	0.90	0.90	0.91	0.91	0.92
		0.65	0.92	0.86	0.83	0.81	0.80	0.80	0.79	0.79	0.78	0.78	0.92	0.87	0.85	0.84	0.84	0.85	0.85	0.86	0.87	0.88
		0.70	0.85	0.79	0.77	0.75	0.74	0.74	0.73	0.73	0.73	0.72	0.86	0.81	0.80	0.79	0.80	0.80	0.81	0.82	0.83	0.84
		0.75	0.80	0.74	0.72	0.70	0.69	0.69	0.68	0.68	0.68	0.67	0.80	0.76	0.75	0.75	0.75	0.76	0.77	0.78	0.79	0.81
		0.80	0.75	0.69	0.67	0.66	0.65	0.64	0.64	0.64	0.63	0.63	0.76	0.72	0.71	0.71	0.71	0.72	0.73	0.74	0.76	0.77
		0.85	0.70	0.65	0.63	0.62	0.61	0.60	0.60	0.60	0.59	0.59	0.71	0.68	0.67	0.67	0.68	0.69	0.70	0.71	0.73	0.74
		0.90	0.66	0.62	0.60	0.58	0.58	0.57	0.57	0.56	0.56	0.56	0.67	0.64	0.63	0.64	0.64	0.66	0.67	0.68	0.70	0.72
	0.50	0.60			0.98	0.97	0.96	0.95	0.94	0.94	0.94	0.94			0.99	0.97	0.97	0.96	0.96	0.96	0.97	0.97
		0.65	1.00	0.93	0.91	0.89	0.88	0.87	0.87	0.87	0.86	0.86	1.00	0.94	0.92	0.91	0.91	0.91	0.91	0.92	0.92	0.93
		0.70	0.93	0.87	0.84	0.83	0.82	0.81	0.81	0.80	0.80	0.80	0.93	0.88	0.86	0.86	0.86	0.86	0.86	0.87	0.88	0.89
		0.75	0.86	0.81	0.78	0.77	0.76	0.76	0.75	0.75	0.74	0.74	0.87	0.82	0.81	0.81	0.81	0.82	0.82	0.83	0.84	0.85
		0.80	0.81	0.76	0.73	0.72	0.71	0.71	0.70	0.70	0.70	0.70	0.82	0.78	0.77	0.76	0.77	0.78	0.78	0.80	0.81	0.82
		0.85	0.76	0.71	0.69	0.68	0.67	0.66	0.66	0.66	0.65	0.65	0.77	0.73	0.72	0.73	0.73	0.74	0.75	0.76	0.78	0.79
		0.90	0.72	0.67	0.65	0.64	0.63	0.63	0.62	0.62	0.62	0.62	0.73	0.70	0.69	0.69	0.70	0.71	0.72	0.73	0.75	0.76
	0.55	0.65			0.99	0.97	0.96	0.95	0.95	0.95	0.94	0.94			0.99	0.98	0.97	0.97	0.97	0.97	0.97	0.97
		0.70	1.00	0.94	0.91	0.90	0.89	0.88	0.88	0.88	0.87	0.87	1.00	0.95	0.93	0.92	0.92	0.92	0.92	0.92	0.93	0.93
		0.75	0.93	0.88	0.85	0.84	0.83	0.82	0.82	0.82	0.81	0.81	0.93	0.89	0.87	0.87	0.87	0.87	0.87	0.88	0.89	0.90
		0.80	0.87	0.82	0.80	0.79	0.78	0.77	0.77	0.76	0.76	0.76	0.88	0.84	0.82	0.82	0.82	0.83	0.84	0.84	0.85	0.86
		0.85	0.82	0.77	0.75	0.74	0.73	0.72	0.72	0.72	0.72	0.71	0.83	0.79	0.78	0.78	0.78	0.79	0.80	0.81	0.82	0.83
		0.90	0.77	0.73	0.71	0.70	0.69	0.68	0.68	0.68	0.67	0.67	0.78	0.75	0.74	0.74	0.75	0.76	0.77	0.78	0.79	0.80
	0.60	0.70			0.99	0.97	0.96	0.96	0.95	0.95	0.95	0.95			0.99	0.98	0.97	0.97	0.97	0.97	0.97	0.97
		0.75	1.00	0.94	0.92	0.91	0.90	0.89	0.89	0.88	0.88	0.88	1.00	0.95	0.93	0.92	0.92	0.92	0.92	0.93	0.93	0.94
		0.80	0.94	0.88	0.86	0.85	0.84	0.83	0.83	0.83	0.83	0.82	0.94	0.89	0.88	0.88	0.88	0.88	0.88	0.89	0.90	0.90

β_{1s} or $\beta_{2,s}$	$\frac{\triangle_1}{\sigma}$	$\frac{\triangle_{1c}}{\sigma}$	p_1 : Similarity criteria I										p_1 : Similarity criteria II									
			0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
		0.85	0.88	0.83	0.81	0.80	0.79	0.78	0.78	0.78	0.78	0.77	0.89	0.85	0.83	0.83	0.83	0.84	0.85	0.85	0.86	0.87
		0.90	0.83	0.78	0.76	0.75	0.75	0.74	0.74	0.73	0.73	0.73	0.84	0.80	0.79	0.79	0.80	0.80	0.81	0.82	0.83	0.84
	0.65	0.75			0.99	0.97	0.97	0.96	0.96	0.95	0.95	0.95			0.99	0.98	0.97	0.97	0.97	0.97	0.97	0.98
		0.80	1.00	0.95	0.93	0.91	0.90	0.90	0.90	0.89	0.89	0.89	1.00	0.95	0.94	0.93	0.93	0.93	0.93	0.93	0.94	0.94
		0.85	0.94	0.89	0.87	0.86	0.85	0.84	0.84	0.84	0.84	0.84	0.94	0.90	0.89	0.88	0.88	0.89	0.89	0.90	0.90	0.91
		0.90	0.89	0.84	0.82	0.81	0.80	0.80	0.79	0.79	0.79	0.79	0.89	0.85	0.84	0.84	0.84	0.85	0.86	0.86	0.87	0.88
	0.70	0.80			0.99	0.98	0.97	0.96	0.96	0.96	0.96	0.95			0.99	0.98	0.98	0.97	0.97	0.97	0.98	0.98
		0.85	1.00	0.95	0.93	0.92	0.91	0.91	0.90	0.90	0.90	0.90	1.00	0.96	0.94	0.93	0.93	0.93	0.93	0.94	0.94	0.95
		0.90	0.94	0.90	0.88	0.87	0.86	0.85	0.85	0.85	0.85	0.85	0.95	0.91	0.89	0.89	0.89	0.89	0.90	0.90	0.91	0.92
	0.75	0.85			0.99	0.98	0.97	0.97	0.96	0.96	0.96	0.96			0.99	0.98	0.98	0.98	0.98	0.98	0.98	0.98
		0.90	1.00	0.95	0.93	0.92	0.92	0.91	0.91	0.90	0.90	0.90	1.00	0.96	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.95
	0.80	*0.90*			0.99	0.98	0.97	0.97	0.96	0.96	0.96	0.96			0.99	0.98	0.98	0.98	0.98	0.98	0.98	0.98
0.35	0.40	0.45							1.00	1.00	0.99	0.99							1.00	1.00	1.00	1.00
		0.50		0.96	0.93	0.92	0.91	0.90	0.90	0.89	0.89	0.89		0.96	0.94	0.93	0.93	0.93	0.93	0.93	0.94	0.94
		0.55	0.92	0.87	0.85	0.83	0.82	0.82	0.81	0.81	0.81	0.81	0.93	0.88	0.87	0.86	0.86	0.86	0.87	0.88	0.88	0.89
		0.60	0.85	0.80	0.77	0.76	0.75	0.75	0.74	0.74	0.74	0.74	0.85	0.81	0.80	0.80	0.80	0.81	0.82	0.83	0.84	0.85
		0.65	0.78	0.73	0.71	0.70	0.69	0.69	0.68	0.68	0.68	0.68	0.79	0.75	0.75	0.75	0.75	0.76	0.77	0.78	0.79	0.81
		0.70	0.72	0.68	0.66	0.65	0.64	0.64	0.64	0.63	0.63	0.63	0.73	0.70	0.70	0.70	0.71	0.72	0.73	0.74	0.76	0.77
		0.75	0.68	0.64	0.62	0.61	0.60	0.60	0.59	0.59	0.59	0.59	0.69	0.66	0.66	0.66	0.67	0.68	0.69	0.71	0.72	0.74
		0.80	0.63	0.60	0.58	0.57	0.56	0.56	0.55	0.55	0.55	0.55	0.65	0.62	0.62	0.62	0.63	0.64	0.66	0.67	0.69	0.71
		0.85	0.60	0.56	0.54	0.54	0.53	0.52	0.52	0.52	0.52	0.52	0.61	0.59	0.58	0.59	0.60	0.61	0.63	0.64	0.66	0.68
		0.90	0.56	0.53	0.51	0.51	0.50						0.58	0.56	0.55	0.56	0.57	0.58	0.60	0.62	0.63	0.65
	0.45	0.50							1.00	1.00	1.00	0.99							1.00	1.00	1.00	1.00
		0.55		0.96	0.94	0.92	0.92	0.91	0.91	0.90	0.90	0.90		0.96	0.95	0.94	0.94	0.94	0.94	0.94	0.94	0.95
		0.60	0.93	0.88	0.86	0.85	0.84	0.83	0.83	0.83	0.82	0.82	0.93	0.89	0.88	0.87	0.87	0.88	0.88	0.89	0.90	0.90
		0.65	0.86	0.81	0.79	0.78	0.77	0.77	0.76	0.76	0.76	0.76	0.86	0.83	0.82	0.82	0.82	0.83	0.83	0.84	0.85	0.86
		0.70	0.80	0.75	0.73	0.72	0.72	0.71	0.71	0.71	0.70	0.70	0.80	0.77	0.77	0.77	0.77	0.78	0.79	0.80	0.81	0.83
		0.75	0.74	0.70	0.69	0.67	0.67	0.66	0.66	0.66	0.66	0.65	0.75	0.72	0.72	0.72	0.73	0.74	0.75	0.76	0.78	0.79
		0.80	0.70	0.66	0.64	0.63	0.63	0.62	0.62	0.62	0.61	0.61	0.71	0.68	0.68	0.68	0.69	0.70	0.71	0.73	0.74	0.76
		0.85	0.66	0.62	0.60	0.59	0.59	0.58	0.58	0.58	0.58	0.58	0.67	0.64	0.64	0.65	0.66	0.67	0.68	0.70	0.71	0.73
		0.90	0.62	0.59	0.57	0.56	0.56	0.55	0.55	0.55	0.54	0.54	0.63	0.61	0.61	0.62	0.63	0.64	0.65	0.67	0.68	0.70
	0.50	0.55							1.00	1.00	1.00	1.00							1.00	1.00	1.00	1.00
		0.60		0.96	0.94	0.93	0.92	0.92	0.91	0.91	0.91	0.91		0.97	0.95	0.94	0.94	0.94	0.94	0.95	0.95	0.95
		0.65	0.94	0.89	0.87	0.86	0.85	0.85	0.84	0.84	0.84	0.84	0.94	0.90	0.89	0.88	0.88	0.89	0.89	0.90	0.90	0.91
		0.70	0.87	0.83	0.81	0.80	0.79	0.78	0.78	0.78	0.78	0.78	0.87	0.84	0.83	0.83	0.83	0.84	0.85	0.85	0.86	0.87
		0.75	0.81	0.77	0.75	0.74	0.74	0.73	0.73	0.73	0.72	0.72	0.82	0.79	0.78	0.78	0.79	0.80	0.80	0.82	0.83	0.84

β_{1s} or $\beta_{2,s}$	$\frac{\triangle_1}{\sigma}$	$\frac{\triangle_{1c}}{\sigma}$	p_1 : Similarity criteria I										p_1 : Similarity criteria II									
			0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
		0.80	0.76	0.72	0.71	0.70	0.69	0.68	0.68	0.68	0.68	0.68	0.77	0.74	0.74	0.74	0.75	0.76	0.77	0.78	0.79	0.81
		0.85	0.71	0.68	0.66	0.65	0.65	0.64	0.64	0.64	0.64	0.64	0.72	0.70	0.70	0.70	0.71	0.72	0.73	0.75	0.76	0.78
		0.90	0.67	0.64	0.63	0.62	0.61	0.61	0.60	0.60	0.60	0.60	0.69	0.66	0.66	0.67	0.68	0.69	0.70	0.72	0.73	0.75
	0.55	0.60							1.00	1.00	1.00	1.00							1.00	1.00	1.00	1.00
		0.65		0.97	0.95	0.94	0.93	0.92	0.92	0.92	0.92	0.92		0.97	0.96	0.95	0.95	0.95	0.95	0.95	0.95	0.96
		0.70	0.94	0.90	0.88	0.87	0.86	0.86	0.85	0.85	0.85	0.85	0.94	0.91	0.90	0.89	0.89	0.90	0.90	0.91	0.91	0.92
		0.75	0.88	0.84	0.82	0.81	0.80	0.80	0.80	0.79	0.79	0.79	0.88	0.85	0.84	0.84	0.85	0.85	0.86	0.87	0.87	0.88
		0.80	0.82	0.78	0.77	0.76	0.75	0.75	0.75	0.74	0.74	0.74	0.83	0.80	0.80	0.80	0.80	0.81	0.82	0.83	0.84	0.85
		0.85	0.77	0.74	0.72	0.71	0.71	0.70	0.70	0.70	0.70	0.70	0.78	0.76	0.75	0.76	0.76	0.77	0.78	0.79	0.81	0.82
		0.90	0.73	0.70	0.68	0.67	0.67	0.66	0.66	0.66	0.66	0.66	0.74	0.72	0.72	0.72	0.73	0.74	0.75	0.76	0.78	0.79
	0.60	0.65							1.00	1.00	1.00	1.00							1.00	1.00	1.00	1.00
		0.70		0.97	0.95	0.94	0.93	0.93	0.93	0.92	0.92	0.92		0.97	0.96	0.95	0.95	0.95	0.95	0.95	0.96	0.96
		0.75	0.94	0.90	0.89	0.88	0.87	0.87	0.86	0.86	0.86	0.86	0.95	0.91	0.90	0.90	0.90	0.90	0.91	0.91	0.92	0.92
		0.80	0.89	0.85	0.83	0.82	0.82	0.81	0.81	0.81	0.81	0.80	0.89	0.86	0.85	0.85	0.86	0.86	0.87	0.87	0.88	0.89
		0.85	0.83	0.80	0.78	0.77	0.77	0.76	0.76	0.76	0.76	0.76	0.84	0.81	0.81	0.81	0.81	0.82	0.83	0.84	0.85	0.86
		0.90	0.79	0.75	0.74	0.73	0.72	0.72	0.72	0.72	0.71	0.71	0.79	0.77	0.77	0.77	0.78	0.79	0.80	0.81	0.82	0.83
	0.65	0.70							1.00	1.00	1.00	1.00							1.00	1.00	1.00	1.00
		0.75		0.97	0.95	0.95	0.94	0.93	0.93	0.93	0.93	0.93		0.97	0.96	0.96	0.95	0.95	0.95	0.96	0.96	0.96
		0.80	0.95	0.91	0.89	0.89	0.88	0.88	0.87	0.87	0.87	0.87	0.95	0.92	0.91	0.91	0.91	0.91	0.91	0.92	0.92	0.93
		0.85	0.89	0.86	0.84	0.83	0.83	0.82	0.82	0.82	0.82	0.82	0.90	0.87	0.86	0.86	0.86	0.87	0.88	0.88	0.89	0.90
		0.90	0.84	0.81	0.79	0.79	0.78	0.78	0.77	0.77	0.77	0.77	0.85	0.82	0.82	0.82	0.83	0.83	0.84	0.85	0.86	0.87
	0.70	0.75							1.00	1.00	1.00	1.00							1.00	1.00	1.00	1.00
		0.80		0.97	0.96	0.95	0.94	0.94	0.94	0.93	0.93	0.93		0.98	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.97
		0.85	0.95	0.92	0.90	0.89	0.89	0.88	0.88	0.88	0.88	0.88	0.95	0.92	0.91	0.91	0.91	0.92	0.92	0.92	0.93	0.93
		0.90	0.90	0.86	0.85	0.84	0.84	0.83	0.83	0.83	0.83	0.83	0.90	0.88	0.87	0.87	0.87	0.88	0.88	0.89	0.90	0.91
	0.75	0.80							1.00	1.00	1.00	1.00							1.00	1.00	1.00	1.00
		0.85		0.98	0.96	0.95	0.95	0.94	0.94	0.94	0.94	0.94		0.98	0.97	0.96	0.96	0.96	0.96	0.96	0.96	0.97
		0.90	0.95	0.92	0.91	0.90	0.89	0.89	0.89	0.89	0.88	0.88	0.96	0.93	0.92	0.92	0.92	0.92	0.92	0.93	0.93	0.94
	0.80	0.85							1.00	1.00	1.00	1.00							1.00	1.00	1.00	1.00
		0.90		0.98	0.96	0.95	0.95	0.95	0.94	0.94	0.94	0.94		0.98	0.97	0.96	0.96	0.96	0.96	0.96	0.97	0.97
	0.85	*0.90*							1.00	1.00	1.00	1.00							1.00	1.00	1.00	1.00

Appendix E : Table A3. p_1 to calculate the sample size in the region of interest

ρ	$\beta_{1,s}$ or $\beta_{2,s}$	$\frac{\triangle_1}{\sigma}$	$\frac{\triangle_{1c}}{\sigma}$	$N: D_1 \geq \rho D_{1c}$						$N: D_1 \geq \rho D$					
				100	300	500	700	1000	1500	100	300	500	700	1000	1500
0.8	0.1	0.2	0.40					0.2700	0.1638						0.2464
			0.45			0.3194	0.2042	0.1352	0.0870					0.1636	0.0960
			0.50		0.3542	0.1792	0.1224	0.0832	0.0544			0.2372	0.1406	0.0904	0.0572
		0.4	0.65					0.2700	0.1638						
			0.70			0.3194	0.2042	0.1352	0.0870					0.2068	0.1034
			0.75		0.3536	0.1792	0.1224	0.0832	0.0544				0.1564	0.0950	0.0588
			0.80		0.2136	0.1186	0.0826	0.0568	0.0374			0.1442	0.0924	0.0610	0.0390
			0.85		0.1486	0.0850	0.0596	0.0412	0.0274		0.1924	0.0946	0.0638	0.0432	0.0280
			0.90	0.5272	0.1106	0.0642	0.0452	0.0314	0.0208		0.1270	0.0688	0.0474	0.0324	0.0212
	0.15	0.2	0.35						0.2642						
			0.40			0.4186	0.2476	0.1608	0.1026					0.2354	0.1196
			0.45		0.3714	0.1844	0.1258	0.0854	0.0558			0.2758	0.1490	0.0942	0.0592
			0.50		0.1992	0.1114	0.0776	0.0534	0.0352		0.2960	0.1258	0.0838	0.0562	0.0364
		0.4	0.60						0.2642						
			0.65			0.4178	0.2476	0.1608	0.1026						0.1410
			0.70		0.3688	0.1844	0.1258	0.0854	0.0558				0.1770	0.1010	0.0614
			0.75		0.1992	0.1114	0.0776	0.0534	0.0352			0.1372	0.0876	0.0576	0.0370
			0.80		0.1308	0.0754	0.0530	0.0368	0.0244		0.1652	0.0834	0.0566	0.0384	0.0250
			0.85	0.3566	0.0936	0.0546	0.0386	0.0268	0.0178		0.1056	0.0580	0.0402	0.0276	0.0182
			0.90	0.2426	0.0704	0.0414	0.0294	0.0204	0.0136		0.0760	0.0432	0.0302	0.0208	0.0138
	0.2	0.2	0.35				0.4886	0.2598	0.1586						
			0.40			0.2234	0.1500	0.1012	0.0658				0.2048	0.1178	0.0716
			0.45		0.2066	0.1150	0.0800	0.0552	0.0364			0.1332	0.0876	0.0584	0.0376
			0.50		0.1234	0.0714	0.0502	0.0348	0.0230		0.1420	0.0764	0.0526	0.0358	0.0236
		0.4	0.60				0.4876	0.2598	0.1586						
			0.65			0.2234	0.1500	0.1012	0.0658					0.1380	0.0766

ρ	$\beta_{1,s}$ or $\beta_{2,s}$	$\frac{\triangle_1}{\sigma}$	$\frac{\triangle_{1c}}{\sigma}$	$N: D_1 \geq \rho D_{1c}$						$N: D_1 \geq \rho D$					
				100	300	500	700	1000	1500	100	300	500	700	1000	1500
			0.70		0.2062	0.1150	0.0800	0.0552	0.0364			0.1520	0.0934	0.0606	0.0384
			0.75		0.1234	0.0714	0.0502	0.0348	0.0230		0.1584	0.0794	0.0538	0.0364	0.0238
			0.80	0.3012	0.0832	0.0488	0.0346	0.0240	0.0160		0.0934	0.0518	0.0360	0.0248	0.0162
			0.85	0.2012	0.0602	0.0356	0.0252	0.0176	0.0118		0.0644	0.0370	0.0260	0.0180	0.0118
			0.90	0.1474	0.0456	0.0270	0.0192	0.0134	0.0090	0.1844	0.0478	0.0278	0.0196	0.0136	0.0090
	0.25	0.2	0.35			0.3782	0.2312	0.1512	0.0968						0.1186
			0.40		0.2442	0.1324	0.0918	0.0630	0.0414			0.1680	0.1048	0.0682	0.0434
			0.45		0.1234	0.0712	0.0502	0.0348	0.0230		0.1454	0.0770	0.0528	0.0360	0.0236
			0.50	0.3024	0.0762	0.0448	0.0318	0.0222	0.0148		0.0820	0.0466	0.0326	0.0226	0.0148
		0.4	0.60			0.3760	0.2312	0.1512	0.0968						
			0.65		0.2422	0.1324	0.0918	0.0630	0.0414				0.1186	0.0726	0.0450
			0.70		0.1234	0.0712	0.0502	0.0348	0.0230		0.1710	0.0812	0.0546	0.0368	0.0238
			0.75	0.2690	0.0762	0.0448	0.0318	0.0222	0.0148		0.0858	0.0476	0.0332	0.0228	0.0150
			0.80	0.1708	0.0522	0.0308	0.0220	0.0154	0.0102	0.2880	0.0556	0.0320	0.0224	0.0156	0.0104
			0.85	0.1206	0.0380	0.0226	0.0160	0.0112	0.0076	0.1438	0.0396	0.0230	0.0164	0.0114	0.0076
			0.90	0.0902	0.0288	0.0172	0.0124	0.0086	0.0058	0.1002	0.0298	0.0176	0.0124	0.0086	0.0058
	0.3	0.2	0.30						0.2676						
			0.35		0.4122	0.1868	0.1270	0.0864	0.0564				0.1808	0.1024	0.0620
			0.40		0.1328	0.0762	0.0536	0.0372	0.0246		0.1692	0.0844	0.0572	0.0388	0.0252
			0.45	0.3008	0.0712	0.0418	0.0298	0.0206	0.0138		0.0770	0.0436	0.0306	0.0212	0.0140
			0.50	0.1538	0.0448	0.0266	0.0190	0.0132	0.0088	0.1970	0.0466	0.0272	0.0192	0.0134	0.0088
		0.4	0.55						0.2676						
			0.60		0.3752	0.1866	0.1270	0.0864	0.0564					0.1300	0.0680
			0.65		0.1322	0.0762	0.0536	0.0372	0.0246			0.0920	0.0602	0.0400	0.0258
			0.70	0.2478	0.0712	0.0418	0.0298	0.0206	0.0138		0.0810	0.0448	0.0310	0.0214	0.0140
			0.75	0.1444	0.0448	0.0266	0.0190	0.0132	0.0088	0.2076	0.0476	0.0274	0.0194	0.0134	0.0090
			0.80	0.0966	0.0308	0.0184	0.0132	0.0092	0.0062	0.1114	0.0320	0.0188	0.0134	0.0092	0.0062
			0.85	0.0696	0.0226	0.0134	0.0096	0.0068	0.0046	0.0756	0.0230	0.0136	0.0098	0.0068	0.0046

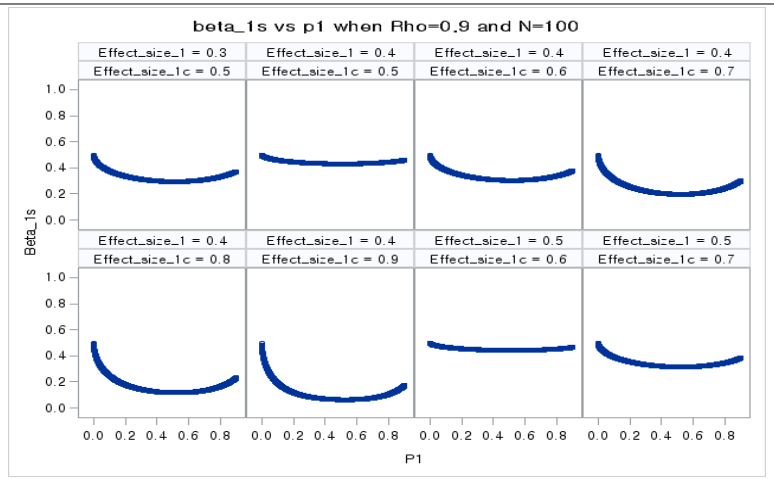
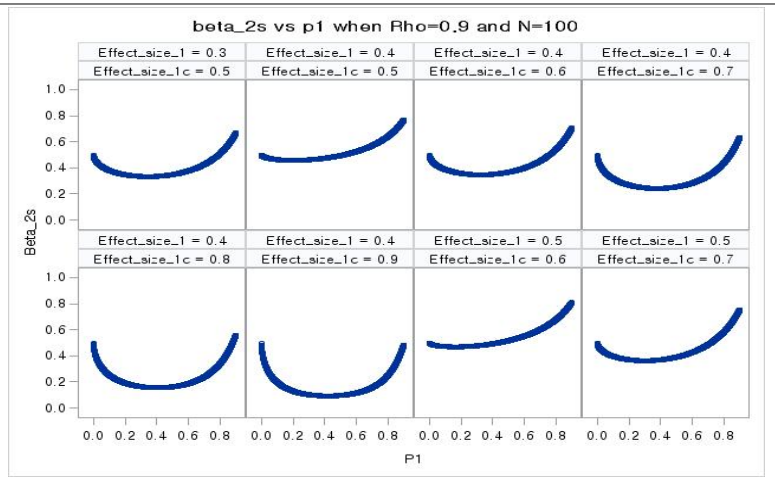
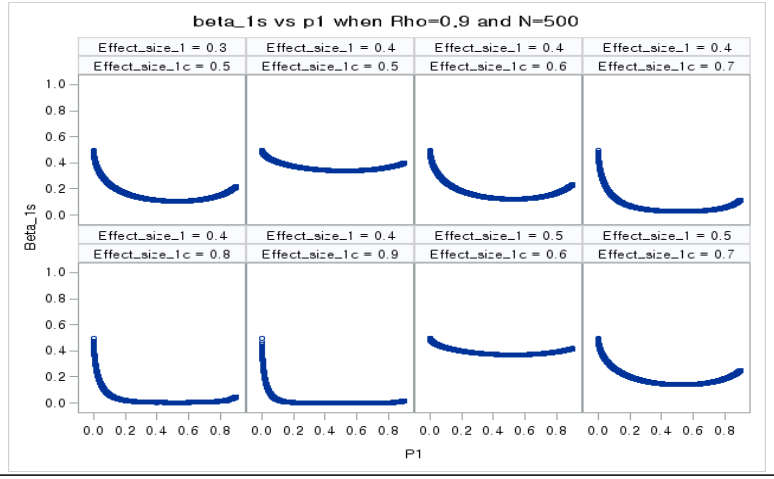
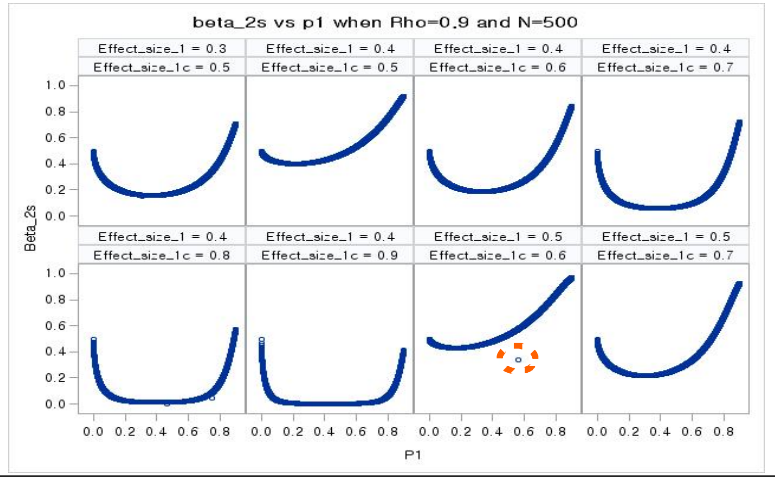
ρ	$\beta_{1,s}$ or $\beta_{2,s}$	$\frac{\triangle_1}{\sigma}$	$\frac{\triangle_{1c}}{\sigma}$	$N: D_1 \geq \rho D_{1c}$						$N: D_1 \geq \rho D$					
				100	300	500	700	1000	1500	100	300	500	700	1000	1500
			0.90	0.0528	0.0172	0.0104	0.0074	0.0052	0.0034	0.0556	0.0176	0.0104	0.0074	0.0052	0.0036
	0.35	0.2	0.30				0.3230	0.2008	0.1260						
			0.35		0.1658	0.0922	0.0646	0.0446	0.0294			0.1114	0.0724	0.0480	0.0308
			0.40	0.3372	0.0672	0.0396	0.0280	0.0196	0.0130		0.0734	0.0414	0.0290	0.0200	0.0132
			0.45	0.1316	0.0370	0.0220	0.0156	0.0110	0.0074	0.1674	0.0384	0.0224	0.0160	0.0110	0.0074
			0.50	0.0762	0.0234	0.0140	0.0100	0.0070	0.0048	0.0826	0.0240	0.0142	0.0102	0.0070	0.0048
		0.4	0.55				0.3220	0.2008	0.1260						
			0.60		0.1620	0.0922	0.0646	0.0446	0.0294			0.1534	0.0812	0.0510	0.0320
			0.65	0.2330	0.0670	0.0396	0.0280	0.0196	0.0130		0.0784	0.0428	0.0296	0.0202	0.0134
			0.70	0.1178	0.0370	0.0220	0.0156	0.0110	0.0074	0.1586	0.0392	0.0228	0.0160	0.0112	0.0074
			0.75	0.0728	0.0234	0.0140	0.0100	0.0070	0.0048	0.0814	0.0242	0.0142	0.0102	0.0072	0.0048
			0.80	0.0498	0.0162	0.0098	0.0070	0.0050	0.0034	0.0528	0.0166	0.0098	0.0070	0.0050	0.0034
			0.85	0.0362	0.0120	0.0072	0.0052	0.0036	0.0024	0.0376	0.0120	0.0072	0.0052	0.0036	0.0024
			0.90	0.0276	0.0092	0.0056	0.0040	0.0028	0.0020	0.0284	0.0092	0.0056	0.0040	0.0028	0.0020
0.9	0.1	0.2	0.35					0.3344	0.1882						0.2294
			0.40			0.3552	0.2142	0.1388	0.0884				0.2604	0.1518	0.0928
			0.45		0.3670	0.1754	0.1186	0.0802	0.0522			0.1956	0.1260	0.0832	0.0534
			0.50		0.2022	0.1108	0.0768	0.0526	0.0346		0.2288	0.1166	0.0792	0.0538	0.0350
		0.4	0.55						0.3212						
			0.60				0.3142	0.1912	0.1186					0.2932	0.1360
			0.65			0.2276	0.1500	0.1002	0.0648				0.1752	0.1090	0.0680
			0.70		0.2522	0.1338	0.0920	0.0628	0.0410			0.1488	0.0980	0.0652	0.0422
			0.75		0.1600	0.0898	0.0626	0.0432	0.0284		0.1810	0.0948	0.0650	0.0442	0.0288
			0.80		0.1130	0.0650	0.0456	0.0316	0.0210		0.1206	0.0672	0.0466	0.0320	0.0212
			0.85	0.3304	0.0846	0.0492	0.0348	0.0242	0.0160		0.0882	0.0504	0.0354	0.0244	0.0162
			0.90	0.2330	0.0660	0.0388	0.0274	0.0190	0.0126	0.2836	0.0680	0.0394	0.0278	0.0192	0.0128
	0.15	0.2	0.35				0.3000	0.1846	0.1150					0.2228	0.1252
			0.40		0.4486	0.1922	0.1290	0.0868	0.0564			0.2242	0.1396	0.0910	0.0580

ρ	$\beta_{1,s}$ or $\beta_{2,s}$	$\frac{\triangle_1}{\sigma}$	$\frac{\triangle_{1c}}{\sigma}$	$N : D_1 \geq \rho D_{1c}$						$N : D_1 \geq \rho D$					
				100	300	500	700	1000	1500	100	300	500	700	1000	1500
			0.45		0.1960	0.1078	0.0748	0.0512	0.0338		0.2240	0.1138	0.0774	0.0524	0.0342
			0.50		0.1224	0.0700	0.0492	0.0340	0.0224		0.1296	0.0722	0.0502	0.0344	0.0226
		0.4	0.55					0.3126	0.1792						
			0.60			0.2750	0.1764	0.1164	0.0748				0.2392	0.1330	0.0804
			0.65		0.2572	0.1358	0.0934	0.0636	0.0416			0.1548	0.1006	0.0668	0.0428
			0.70		0.1482	0.0838	0.0586	0.0404	0.0266		0.1682	0.0886	0.0608	0.0414	0.0270
			0.75	0.4704	0.0990	0.0572	0.0404	0.0280	0.0186		0.1054	0.0592	0.0412	0.0284	0.0188
			0.80	0.2582	0.0714	0.0418	0.0296	0.0206	0.0136		0.0740	0.0426	0.0300	0.0208	0.0138
			0.85	0.1828	0.0540	0.0318	0.0226	0.0158	0.0106	0.2076	0.0554	0.0324	0.0228	0.0158	0.0106
			0.90	0.1388	0.0424	0.0252	0.0178	0.0124	0.0084	0.1496	0.0432	0.0254	0.0180	0.0126	0.0084
	0.2	0.2	0.30						0.2292						
			0.35			0.2656	0.1712	0.1134	0.0728				0.2014	0.1232	0.0764
			0.40		0.2170	0.1176	0.0812	0.0556	0.0366		0.2666	0.1262	0.0850	0.0572	0.0372
			0.45		0.1196	0.0686	0.0482	0.0334	0.0220		0.1272	0.0708	0.0492	0.0338	0.0222
			0.50	0.3050	0.0774	0.0452	0.0320	0.0222	0.0148		0.0800	0.0460	0.0324	0.0224	0.0148
		0.4	0.55				0.2840	0.1766	0.1104						0.1322
			0.60		0.3182	0.1600	0.1088	0.0738	0.0482			0.2026	0.1226	0.0792	0.0502
			0.65		0.1514	0.0854	0.0596	0.0412	0.0272		0.1772	0.0914	0.0624	0.0424	0.0276
			0.70	0.3932	0.0928	0.0538	0.0380	0.0264	0.0174		0.0988	0.0556	0.0388	0.0268	0.0176
			0.75	0.2210	0.0632	0.0372	0.0264	0.0184	0.0122	0.2840	0.0654	0.0378	0.0266	0.0184	0.0122
			0.80	0.1520	0.0460	0.0272	0.0194	0.0136	0.0090	0.1682	0.0470	0.0276	0.0196	0.0136	0.0090
			0.85	0.1124	0.0350	0.0208	0.0148	0.0104	0.0070	0.1196	0.0356	0.0210	0.0150	0.0104	0.0070
			0.90	0.0870	0.0276	0.0164	0.0118	0.0082	0.0056	0.0908	0.0280	0.0166	0.0118	0.0082	0.0056
	0.25	0.2	0.30				0.3780	0.2170	0.1328						0.1566
			0.35		0.2974	0.1502	0.1026	0.0698	0.0456			0.1708	0.1104	0.0730	0.0468
			0.40		0.1266	0.0722	0.0506	0.0350	0.0232		0.1368	0.0750	0.0520	0.0356	0.0234
			0.45	0.2914	0.0734	0.0430	0.0304	0.0212	0.0140		0.0758	0.0438	0.0308	0.0212	0.0142
			0.50	0.1650	0.0484	0.0286	0.0202	0.0142	0.0094	0.1820	0.0492	0.0288	0.0204	0.0142	0.0094

ρ	$\beta_{1,s}$ or $\beta_{2,s}$	$\frac{\triangle_1}{\sigma}$	$\frac{\triangle_{1c}}{\sigma}$	$N: D_1 \geq \rho D_{1c}$						$N: D_1 \geq \rho D$					
				100	300	500	700	1000	1500	100	300	500	700	1000	1500
		0.4	0.50						0.3196						
			0.55			0.2422	0.1584	0.1054	0.0680				0.2450	0.1246	0.0744
			0.60		0.1726	0.0964	0.0670	0.0462	0.0304		0.2298	0.1064	0.0714	0.0480	0.0312
			0.65	0.3868	0.0916	0.0532	0.0374	0.0260	0.0172		0.0986	0.0552	0.0384	0.0264	0.0174
			0.70	0.1972	0.0576	0.0338	0.0240	0.0168	0.0112	0.2464	0.0596	0.0346	0.0244	0.0168	0.0112
			0.75	0.1286	0.0396	0.0236	0.0168	0.0116	0.0078	0.1406	0.0404	0.0238	0.0168	0.0118	0.0078
			0.80	0.0918	0.0290	0.0172	0.0124	0.0086	0.0058	0.0966	0.0294	0.0174	0.0124	0.0086	0.0058
			0.85	0.0692	0.0222	0.0132	0.0094	0.0066	0.0044	0.0716	0.0224	0.0134	0.0096	0.0066	0.0044
			0.90	0.0542	0.0176	0.0106	0.0076	0.0052	0.0036	0.0554	0.0176	0.0106	0.0076	0.0052	0.0036
	0.3	0.2	0.30			0.2810	0.1786	0.1178	0.0756				0.2452	0.1346	0.0812
			0.35		0.1512	0.0846	0.0592	0.0408	0.0270		0.1728	0.0896	0.0614	0.0418	0.0274
			0.40	0.3026	0.0722	0.0422	0.0298	0.0208	0.0138		0.0750	0.0430	0.0302	0.0210	0.0138
			0.45	0.1472	0.0428	0.0254	0.0180	0.0126	0.0084	0.1622	0.0436	0.0256	0.0182	0.0126	0.0084
			0.50	0.0918	0.0286	0.0170	0.0122	0.0084	0.0056	0.0958	0.0288	0.0170	0.0122	0.0086	0.0056
		0.4	0.50					0.2708	0.1604						
			0.55		0.2420	0.1292	0.0890	0.0608	0.0398			0.1644	0.1012	0.0656	0.0418
			0.60	0.4542	0.0962	0.0558	0.0392	0.0272	0.0180		0.1064	0.0586	0.0406	0.0278	0.0184
			0.65	0.1798	0.0530	0.0314	0.0222	0.0154	0.0104	0.2260	0.0552	0.0320	0.0226	0.0156	0.0104
			0.70	0.1082	0.0338	0.0200	0.0144	0.0100	0.0066	0.1172	0.0344	0.0204	0.0144	0.0100	0.0068
			0.75	0.0734	0.0234	0.0140	0.0100	0.0070	0.0048	0.0766	0.0238	0.0142	0.0100	0.0070	0.0048
			0.80	0.0532	0.0172	0.0104	0.0074	0.0052	0.0034	0.0546	0.0174	0.0104	0.0074	0.0052	0.0036
			0.85	0.0404	0.0132	0.0080	0.0058	0.0040	0.0028	0.0412	0.0134	0.0080	0.0058	0.0040	0.0028
			0.90	0.0318	0.0106	0.0064	0.0046	0.0032	0.0022	0.0324	0.0106	0.0064	0.0046	0.0032	0.0022
	0.35	0.2	0.30		0.2472	0.1266	0.0870	0.0594	0.0390			0.1474	0.0948	0.0628	0.0402
			0.35	0.3726	0.0748	0.0434	0.0308	0.0214	0.0142		0.0786	0.0446	0.0312	0.0216	0.0144
			0.40	0.1308	0.0372	0.0220	0.0158	0.0110	0.0074	0.1448	0.0380	0.0224	0.0158	0.0110	0.0074
			0.45	0.0726	0.0224	0.0134	0.0096	0.0068	0.0046	0.0754	0.0226	0.0134	0.0096	0.0068	0.0046
			0.50	0.0468	0.0150	0.0090	0.0064	0.0046	0.0030	0.0478	0.0152	0.0090	0.0064	0.0046	0.0030

ρ	$\beta_{1,s}$ or $\beta_{2,s}$	$\frac{\triangle_1}{\sigma}$	$\frac{\triangle_{1c}}{\sigma}$	$N: D_1 \geq \rho D_{1c}$						$N: D_1 \geq \rho D$					
				100	300	500	700	1000	1500	100	300	500	700	1000	1500
		0.4	0.50			0.2970	0.1876	0.1234	0.0790						0.0984
			0.55		0.1128	0.0648	0.0456	0.0316	0.0208		0.1358	0.0704	0.0482	0.0328	0.0214
			0.60	0.1654	0.0490	0.0290	0.0206	0.0144	0.0096	0.2162	0.0512	0.0298	0.0210	0.0146	0.0096
			0.65	0.0874	0.0276	0.0164	0.0118	0.0082	0.0056	0.0938	0.0282	0.0166	0.0118	0.0082	0.0056
			0.70	0.0548	0.0178	0.0106	0.0076	0.0054	0.0036	0.0568	0.0180	0.0108	0.0076	0.0054	0.0036
			0.75	0.0378	0.0124	0.0074	0.0054	0.0038	0.0026	0.0386	0.0124	0.0074	0.0054	0.0038	0.0026
			0.80	0.0278	0.0092	0.0056	0.0040	0.0028	0.0020	0.0280	0.0092	0.0056	0.0040	0.0028	0.0020
			0.85	0.0212	0.0070	0.0042	0.0030	0.0022	0.0014	0.0214	0.0070	0.0042	0.0030	0.0022	0.0014
			0.90	0.0168	0.0056	0.0034	0.0024	0.0018	0.0012	0.0168	0.0056	0.0034	0.0024	0.0018	0.0012

Appendix F : Figure A1. Scatter plot of the regional type II error rate and p_1 with changes in N when $\rho = 0.9$

1) Scatter plot of $\beta_{1,s}$ vs p_1	2) Scatter plot of $\beta_{2,s}$ vs p_1
	
	

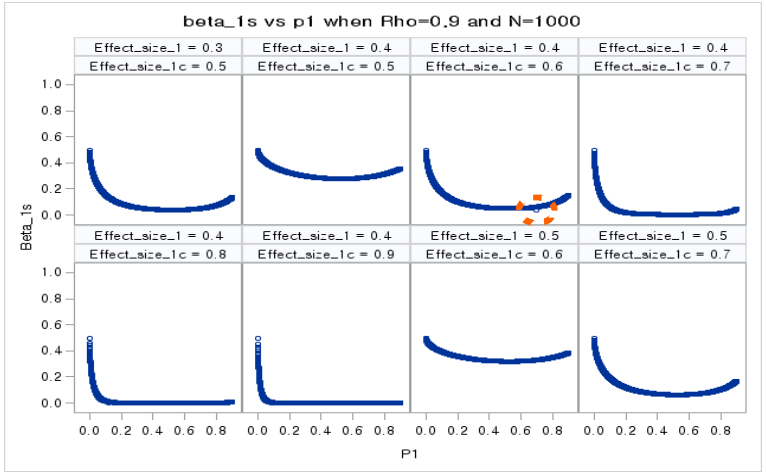
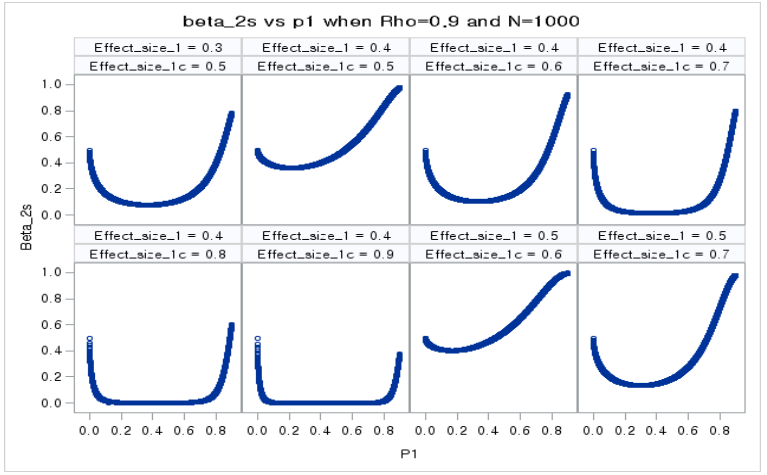
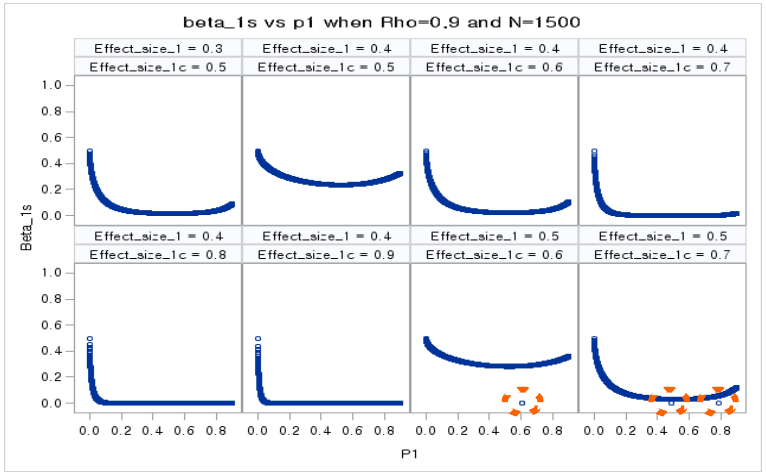
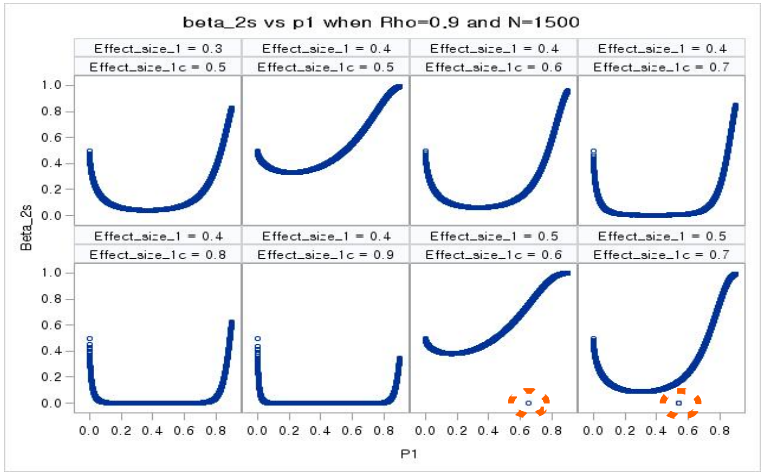
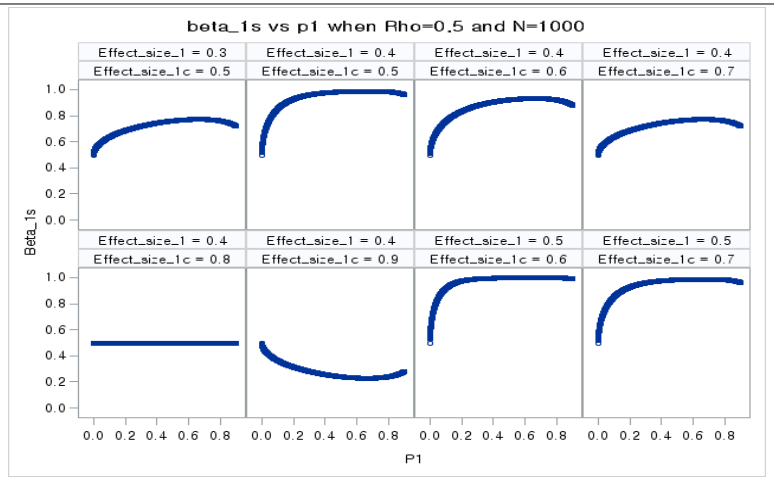
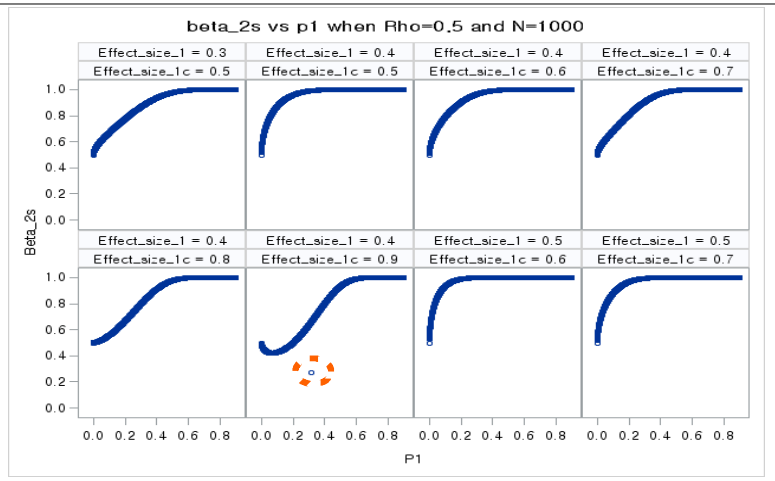
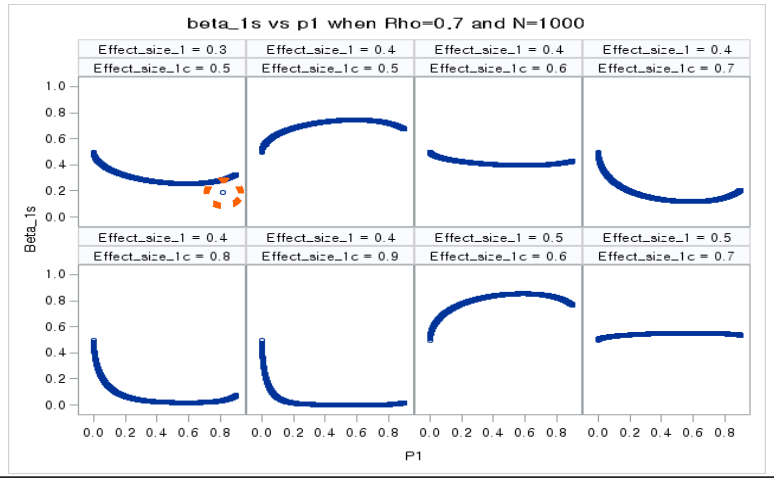
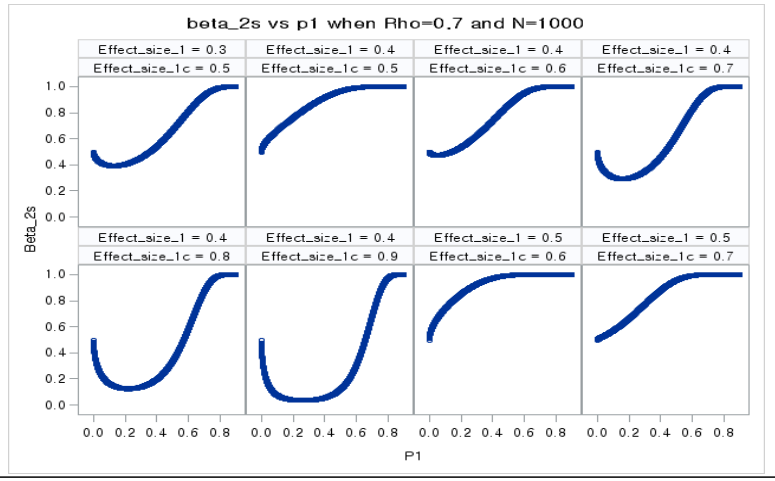
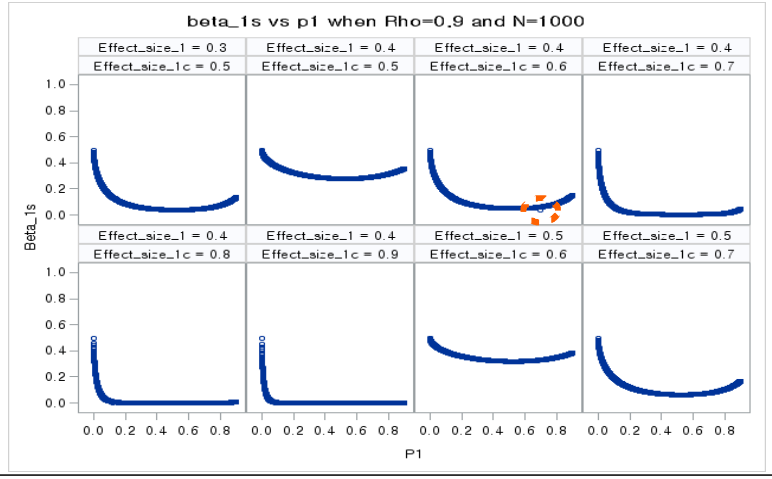
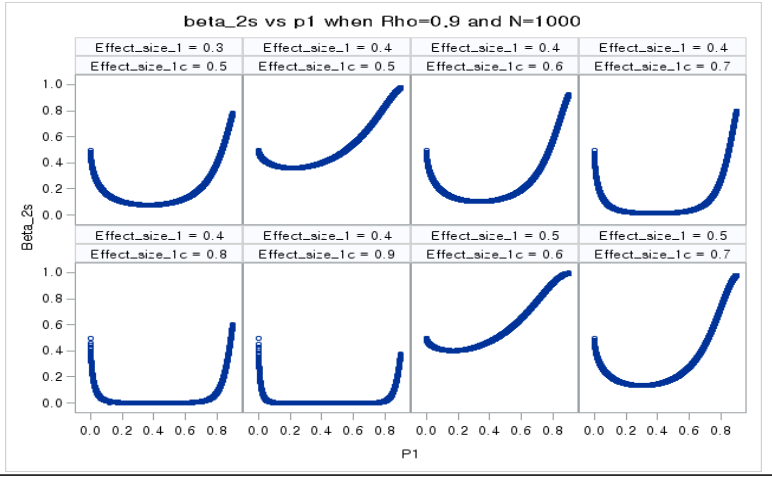
1) Scatter plot of $\beta_{1,s}$ vs p_1	2) Scatter plot of $\beta_{2,s}$ vs p_1
	
	

Figure A2. Scatter plot of the regional type II error rate and p_1 with changes in ρ when $N=1{,}000$

1) Scatter plot of $\beta_{1,s}$ vs p_1	2) Scatter plot of $\beta_{2,s}$ vs p_1
	
	

1) Scatter plot of $\beta_{1,s}$ vs p_1	2) Scatter plot of $\beta_{2,s}$ vs p_1
	

Appendix G : Relationship between the regional type I error and the assurance probability

$$\alpha_{1,s} = P_\delta\left(D_1 < \rho D_{1c} \mid Z > z_{1-\alpha},\ H_{so} : \triangle_1/\triangle_{1c} \geq 1\right)$$

$$= 1 - P_\delta\left(D_1 \geq \rho D_{1c} \mid Z > z_{1-\alpha},\ H_{so} : \triangle_1/\triangle_{1c} \geq 1\right)$$

$$= 1 - P_\delta\left(Z_1 \geq \rho\sqrt{\frac{p_1}{1-p_1}}\, Z_{1c} \mid \sqrt{p_1}\, Z_1 + \sqrt{1-p_1}\, Z_{1c} > z_{1-\alpha},\ H_{so}\right)$$

$$= 1 - P_0\left(Z_1 \geq \rho\sqrt{\frac{p_1}{1-p_1}}\, Z_{1c} + \frac{\sqrt{p_1}}{\sigma\sqrt{2/N}}(\rho\triangle_{1c} - \triangle_1) \mid \sqrt{p_1}\, Z_1 + \sqrt{1-p_1}\, Z_{1c} > z_{1-\alpha} - \frac{1}{\sigma\sqrt{2/N}}(p_1\triangle_1 + (1-p_1)\triangle_{1c})\right)$$

if $\triangle_1 = \triangle_{1c} = \delta$,

$$= 1 - P_0\left(Z_1 \geq \rho\sqrt{\frac{p_1}{1-p_1}}\, Z_{1c} + (\rho-1)\sqrt{p_1}\,\frac{\delta}{\sigma\sqrt{2/N}} \mid \sqrt{p_1}\, Z_1 + \sqrt{1-p_1}\, Z_{1c} > z_{1-\alpha} - \frac{\delta}{\sigma\sqrt{2/N}}\right)$$

$$= 1 - P_0\left(Z_1 \geq \rho\sqrt{\frac{p_1}{1-p_1}}\, Z_{1c} + (\rho-1)\sqrt{p_1}(z_{1-\alpha} + z_{1-\beta}) \mid \sqrt{p_1}\, Z_1 + \sqrt{1-p_1}\, Z_{1c} > -z_{1-\beta}\right)$$

$$\because \frac{\delta}{\sigma\sqrt{2/N}} = Z_{1-\alpha} + Z_{1-\beta}$$

$$= 1 - \frac{P_0\left(Z_1 \geq c_1 Z_{1c} + c_2,\ c_3 Z_1 + c_4 Z_{1c} > c_5\right)}{P_0\left(Z > -Z_{1-\beta}\right)}$$

$$= 1 - AP_1$$

where $c_1 = \rho\sqrt{\frac{p_1}{1-p_1}}$, $c_2 = (\rho-1)\sqrt{p_1}(Z_{1-\alpha} + Z_{1-\beta})$, $c_3 = \sqrt{p_1}$,

$c_4 = \sqrt{1-p_1}$, $c_5 = -Z_{1-\beta}$

Similarly,

$$\alpha_{2,s} = P_\delta\left(D_1 < \rho D \mid Z > z_{1-\alpha},\ H_{so} : \triangle_1/\triangle \geq 1\right)$$

$$= 1 - P_\delta\left(D_1 \geq \rho D \mid Z > z_{1-\alpha},\ H_{so} : \triangle_1/\triangle \geq 1\right)$$

$$= 1 - P_\delta\left(Z_1 \geq \frac{\rho\sqrt{p_1(1-p_1)}}{1-\rho p_1} Z_{1c} \mid \sqrt{p_1}\, Z_1 + \sqrt{1-p_1}\, Z_{1c} > z_{1-\alpha},\ H_{so}\right)$$

$$= 1 - P_0\left(Z_1 \geq \frac{\rho\sqrt{p_1(1-p_1)}}{1-\rho p_1} Z_{1c} + \frac{\rho\sqrt{p_1(1-p_1)}}{\sigma\sqrt{2/N}} \frac{\triangle_{1c}}{\sigma\sqrt{2/N}} \sqrt{1-p_1} - \frac{\triangle_1}{\sigma\sqrt{2/N}} \sqrt{p_1} \mid \sqrt{p_1}\, Z_1 + \sqrt{1-p_1}\, Z_{1c} > z_{1-\alpha} - \frac{1}{\sigma\sqrt{2/N}}(p_1\triangle_1 + (1-p_1)\triangle_{1c})\right)$$

if $\triangle_1 = \triangle_{1c} = \delta$,

$$= 1 - P_0\left(Z_1 \geq \frac{\rho\sqrt{p_1(1-p_1)}}{1-\rho p_1} Z_{1c} + \frac{(\rho-1)\sqrt{p_1}}{1-\rho p_1} \frac{\delta}{\sigma\sqrt{2/N}} \mid \sqrt{p_1}\, Z_1 + \sqrt{1-p_1}\, Z_{1c} > z_{1-\alpha} - \frac{\delta}{\sigma\sqrt{2/N}}\right)$$

$$= 1 - P_0\left(Z_1 \geq \frac{\rho\sqrt{p_1(1-p_1)}}{1-\rho p_1} Z_{1c} + \frac{(\rho-1)\sqrt{p_1}}{1-\rho p_1}(Z_{1-\alpha} + Z_{1-\beta}) \mid \sqrt{p_1}\, Z_1 + \sqrt{1-p_1}\, Z_{1c} > -Z_{1-\beta}\right)$$

$\because \dfrac{\delta}{\sigma\sqrt{2/N}} = Z_{1-\alpha} + Z_{1-\beta}$

$$= 1 - \frac{P_0\left(Z_1 \geq c_6 Z_{1c} + c_7,\ c_3 Z_1 + c_4 Z_{1c} > c_5\right)}{P_0\left(Z > -Z_{1-\beta}\right)}$$

$$= 1 - AP_2$$

where $c_6 = \dfrac{\rho\sqrt{p_1(1-p_1)}}{1-\rho p_1}$, $c_7 = \dfrac{(\rho-1)\sqrt{p_1}}{1-\rho p_1}(Z_{1-\alpha} + Z_{1-\beta})$, $c_5 = -Z_{1-\beta}$

Appendix H : R program for calculating the regional type II error rate

(i) R program code for $\beta_{1,s}$ calculation

```
rm(list=ls())
alpha<-0.025
effsz_1<-seq(0.1,1.5,0.1)
len.effsz_1<-length(effsz_1)
effsz_1c<-seq(0.1,1.5,0.1)
len.effsz_1c<-length(effsz_1c)
N<-seq(100,1000,100)
len.N<-length(N)
rho<-seq(0.5,0.9,0.1)
len.rho<-length(rho)
p1<-seq(0.05,0.95,0.05)
len.p1<-length(p1)

output<-matrix(NA,ncol=16,nrow=(len.effsz_1)*(len.effsz_1c)*(len.N)*(len.rho)*(len.p1))

colnames(output)<-c("Effect_size_1", "Effect_size_1c", "Alpha", "N", "P1", "Rho", "D1", "D2", "D3", "D4", "D5", "B1", "Numerator", "Denominator", "Output_NO", "Beta_1s")
output

for(i in 1:len.effsz_1){
 for(j in 1:len.effsz_1c){
  for(k in 1:len.N){
   for(l in 1:len.rho){
    for(m in 1:len.p1){
     d1=rho[l]*sqrt(p1[m]/(1-p1[m]))
     d2=sqrt(p1[m])/sqrt(2/N[k])*(rho[l]*effsz_1c[j]-effsz_1[i])
     d3=sqrt(p1[m])
```

```
        d4=sqrt(1-p1[m])
        d5=qnorm(1-alpha)-(1/sqrt(2/N[k]))*(p1[m] * effsz_1[i]
            +(1-p1[m])*effsz_1c[j])
        b1=d2+d1*(d5-d2*d3)/(d1*d3+d4)
        f1=function(u){
            y=(pnorm((u-d2)/d1)-pnorm((d5-d3*u)/d4))*dnorm(u)
            return(y)
                        }
        f2=function(u){
            y=(1-pnorm((d5-d3*u)/d4))*dnorm(u)
            return(y)
                        }
        num=integrate(f1,b1,Inf)[[1]]
        denom=integrate(f2,-Inf,Inf)[[1]]
        beta_1s=round(num/denom,4)
        outno<-m+len.p1*(l-1)+(len.p1*len.rho)*(k-1)
                +(len.p1*len.rho*len.N)*(j-1)
                +(len.p1*len.rho*len.N*len.effsz_1c)*(i-1)
        output[outno,]<-c(effsz_1[i], effsz_1c[j], alpha, N[k], p1[m], rho[l],
                          d1, d2, d3, d4, d5, b1, num, denom, outno, beta_1s)
      }
     }
    }
   }
  }
output
```

(ii) R program code for $\beta_{2,s}$ calculation

```
rm(list=ls())
alpha<-0.025
effsz_1<-seq(0.1,1.5,0.1)
len.effsz_1<-length(effsz_1)
effsz_1c<-seq(0.1,1.5,0.1)
len.effsz_1c<-length(effsz_1c)
N<-seq(100,1000,100)
len.N<-length(N)
rho<-seq(0.5,0.9,0.1)
len.rho<-length(rho)
p1<-seq(0.05,0.95,0.05)
len.p1<-length(p1)

output<-matrix(NA,ncol=16,nrow=(len.effsz_1)*(len.effsz_1c)*(len.N)*(len.rho)*(le
n.p1))

colnames(output)<-c("Effect_size_1", "Effect_size_1c", "Alpha", "N", "P1", "Rho",
"D6", "D7", "D8", "D9", "D10", "B2", "Numerator", "Denominator", "Output_NO",
"Beta_2s")
output

for(i in 1:len.effsz_1){
 for(j in 1:len.effsz_1c){
  for(k in 1:len.N){
   for(l in 1:len.rho){
    for(m in 1:len.p1){
     d6=rho[l] * sqrt(p1[m]*(1-p1[m]))/(1-rho[l]*p1[m])
     d7=d6*effsz_1c[j]*sqrt(1-p1[m])/sqrt(2/N[k])-effsz_1[i]*sqrt(p1[m])
        /sqrt(2/N[k])
     d8=sqrt(p1[m])
     d9=sqrt(1-p1[m])
```

```
        d10=qnorm(1-alpha)-p1[m]*effsz_1[i]/sqrt(2/N[k])
             -(1-p1[m])*effsz_1c[j]/sqrt(2/N[k])
        b2=d7 + d6*(d10 - d7*d8)/(d6*d8+d9)
        f1=function(u){
           y=(pnorm((u-d7)/d6)-pnorm((d10-d8*u)/d9))*dnorm(u)
           return(y)
                       }
        f2=function(u){
           y=(1-pnorm((d10-d8*u)/d9))*dnorm(u)
           return(y)
                       }
        num=integrate(f1,b2,Inf)[[1]]
        denom=integrate(f2,-Inf,Inf)[[1]]
        beta_2s=round(num/denom,4)

        outno<-m+len.p1*(l-1)+(len.p1*len.rho)*(k-1)+(len.p1*len.rho*len.N)
                *(j-1)+(len.p1*len.rho*len.N*len.effsz_1c)*(i-1)
        output[outno,]<-c(effsz_1[i], effsz_1c[j], alpha, N[k], p1[m], rho[l],
                          d6, d7, d8, d9, d10, b2, num, denom, outno, beta_2s)
      }
     }
    }
   }
  }
output
```

References

1. Ando, Y. Looking beyond ICH-E9 in the Era of global drug development. PMDA Biostatistics summer workshop. 2012; Available at: http://www.pmda.go.jp/files/000163980.pdf
2. Ando, Y. and Hamasaki, T. Practical issues and lessons learned from multi-regional clinical trials via case examples: a Japanese perspective. Pharmaceutical Statistics. 2010;9:190-200
3. Bae, KS. International trend and current Korean status of clinical trial technology. Korea National Enterprise for Clinical Trials. 2010; Available at: http://konect.or.kr/datainfo/RandD1/Read.htm?paramNttID=2251&param Page=1
4. Chen, J., Quan, H., Binkowitz, B., Quyang, S. P., Tanaka, Y., Li, G., Menjoge, S. and Ibia, E. Assessing consistent treatment effect in a multi-regional clinical trial: a systematic review. Pharmaceutical Statistics. 2010;9:242-253.
5. Chow, SC., Shao, J. and Hu, OYP. Assessing sensitivity and similarity in bridging studies. Journal of Biopharmaceutical Statistics. 2002;12: 385-400.
6. Chow, SC., Shao, J. and Wang, H. Sample size calculations in clinical research. CRC press:Boca Raton. 2003;55-56.
7. Committee for Medicinal Products for Human Use. Guideline on the investigation of subgroups in confirmatory clinical trials. 2014.

8. Fukunaga, S., Kusama, M. and Ono, S. The effect size, study design and development experience in commercially sponsored studies for new drug applications in approved drugs. SpringerPlus. 2014;3:740. Available at: http://www.springerplus.com/contents/3/1/740
9. Hsiao, CF., Hsu, YY., Tsou, HH. and Liu, JP. Use of prior information for Bayesian evaluation of bridging studies. Journal of Biopharmaceutical statistics. 2005;15(4):625-639.
10. Hung, HMJ., Wang, SJ. and O'Neill, RT. Consideration of regional difference in design and analysis of multi-regional trials. Pharmaceutical Statistics. 2010;9:173-178.
11. Ikeda, K. and Bretz, F. Sample size and proportion of Japanese patients in multi-regional trials. Pharmaceutical Statistics. 2010;9:207-216.
12. International Conference on Harmonization. Q&A for the ICH-E5 Guideline on Ethnic Factors in the Acceptability of Foreign Clinical Data. 2006.
13. International Conference on Harmonization. The ICH E9 Guideline on Statistical Principles for Clinical Trials. 1998.
14. International Conference on Harmonization. The ICH E17 Draft Guideline on General Principles for Planning and Design of Multi-Regional Clinical Trials. Current Step 2 version. 2016.
15. Kang, SH. and Bae, JH. Regional type II error rate for testing the consistency of treatment effects across regions in multi-regional Clinical trial. 2016; Submitted.
16. Kawai, N., Stein, C., Komiyama, O. and Li, Y. An approach to

rationalize partitioning sample size into individual regions in a multi-regional trial. Drug Informational Journal. 2008;42:139-147.

17. Ko, FS., Tsou, HH., Liu, JP. and Haias, CF. Sample size determination for a specific region in a multi-regional trial. Journal of Biopharmaceutical Statistics. 2010;24:870-885.
18. Lan, GKK., Soo, YW., Siu, CT. and Wang, M. The use of weighted Z-tests in medical research. Journal of Biopharmaceutical statistics. 2005;15(4):625-639.
19. Liu, JP., Hsiao, CF. and Hsueh, HM. Bayesian approach to evaluation of bridging studies. Journal of Biopharmaceutical statistics. 2002;12(3):401-408.
20. Liu, JP., Hsuech, HM. and Chen, JJ. Sample size requirement for evaluation of bridging evidence. Biometrical Journal. 2002;44:969-981.
21. Ministry of Food and Drug Safety. 2015 Clinical Trial Protocol Approval Status and Annual Inspection Result Presentation. Press release. 2016.
    (Press release for clinical trial approval status and annual inspection result in 2015)
22. Ministry of Health, Labor, and Welfare of Japan. Basic Principles on Global Clinical Trials. 2007; Available at: http://www.pmda.go.jp/english/service/pdf/notifications/0928010-e.pdf
23. Quan, H., Zhao, PL., Roessner, M. and Aizawa, K. Sample size considering for Japanese patients in a multi-regional trial based on

MHLW guidance. Pharmaceutical Statistics. 2010;9:100-112.

24. Shao, J. and Chow, SC. Two-stage sampling with pharmaceutical applications. Statistics in Medicine. 1993;21:1727-1742.
25. Shih, WJ. Clinical trials for drug registration in Asian-Pacific countries: Proposal for a new paradigm from a statistical perspective. Controlled Clinical trials. 2001;22:357-366.
26. Uesaka, H. Sample size allocation to regions in a multi-regional trial. Journal of Biopharmaceutical Statistics. 2009;19:580-594.
27. Wang, SJ. Bridging study versus pre-specified nested in global trials. Drug Information Journal. 2009;43:27-34.

국 문 요 약

다지역 임상시험은 다양한 지역에서 동일한 계획서를 가지고 진행함으로써 모든 지역에서 의약품 허가 및 등록을 용이하기 할 뿐만 아니라 각 지역에서 유사한 임상시험을 중복하여 수행하지 않음으로써 시간과 비용을 줄일 수 있는 기회를 제공하기 때문에 그 관심도가 지속적으로 증가하고 있다. 다지역 임상시험의 목적을 달성할 수 있는 능력은 전체 치료 효과를 입증한 이후 참역한 개별 지역에서 효과가 얼마나 유사한지에 달려 있다. 하지만 치료 효과의 지역 간 유사성을 보이기 위한 유사성 기준에 대한 정의와 이러한 유사성을 보이기 위한 특정 관심 지역에 필요한 임상시험 대상자 산출에 대한 기준에 대한 정의가 이뤄지지 않고 있다. 2007년 일본 후생성은 다지역 임상시험에서 유사성 기준과 이러한 기준을 근거로 일본에 필요한 임상시험 대상자수 산출에 대해 가이드라인을 제공하고 있으나, 통계적인 관점에서는 제안하고 있지는 않다.

Ko et al.(2010)은 후생성의 가이드라인을 기반으로 보장 확률개념을 도입하여 특정 관심 지역의 임상시험 대상자수 결정에 대해 제안했다. 하지만 다지역 임상시험의 두 번째 목적인 지역 간 치료 효과의 유사성 대한 가설에 중점을 두고 평가하지 않았다는 점에서 적절하지 않다. 본 논문에서는 다지역 임상시험의 두 번째 목적을 통계적 가설 검정 절차로 활용한 Kang et al.(2016)의 방법을 효과 크기로 표준화 하여 소개하였다. 또한 제안된 방법을 통하여 유사성에 대한 가설 검정 위한 임계값 및 특정 관심지역에 할당되어야 하는 임상시험 대상자수를 산출하기 위해 지역 제 2종 오류율을 활용하는 방법에 대해 논의 및 비교 평가 하였다.

유사성 기준에 따라 지역 제 2종 오류율을 계산하여 살펴본 결과, 특정 관심지역의 효과 크기와 해당지역을 제외한 나머지 지역의 효과크기의 차이가 크거

나, 유사성에 대한 임계값이 커질수록 지역 제 2종 오류율이 더 잘 조정 되는 것 알 수 있다. 사전 정해진 수준의 지역 제 2종 오류율을 만족하면서 효과크기 및 해당 관심 지역의 환자 할당 비율 등과 같은 모수가 동일하다면, 유사성 기준 $D_1 \geq \rho D$의 임계값이 유사성 기준 $D_1 \geq \rho D_{1c}$의 임계값 더 보수적이라고 할 수 있다. 특정 관심 지역의 환자 비율의 경우 지역 제 2종 오류율과 단조 감소의 형태를 보이지 않는다. 이러한 바람직하지 않은 특성은 새로운 유사성 기준의 개발을 통해 개선이 필요하겠다. 또한 지역 제 1종 오류율을 조절하는 방법은 Ko et al.(2010)의 방법의 확장된 형태로 지역 간의 치료 효과가 이질적이라는 가정에서도 활용할 수 있다는 것을 확인 할 수 있었다.

최근 우리나라도 다지역 임상시험을 통해 새로운 의약품을 개발하는 글로벌화 추세가 더욱 활성화되고 있다. 이렇게 다지역 임상시험의 수행 빈도가 높아질수록 특정 관심 지역 예를 들어 우리나라에 할당되어야하는 임상시험 대상자 수 산출 및 치료 효과의 민족 간 유사성을 평가하기 위한 임계값 설정이 더욱 중요해 질 수밖에 없다. 이러한 임상개발 환경에서 본 논문에 소개된 지역 제 2종 오류를 이용한 방법은 다지역 임상시험 환경에 유용하게 활용 될 것이며, 후속 연구를 통해 개선된 방법을 개발 할 필요가 있겠다.

핵심어 : 다지역 임상시험, 유사성 기준, 임계값, 임상시험 대상자수 결정, 보장 확률, 이차 가설, 지역 오류율, 지역 제 1종 오류, 지역 제 2종 오류