Journal of the Chinese Statistical Association
Vol. 53, (2015) 64–78

AN OVERVIEW OF DESIGN AND EVALUATION OF MULTIREGIONAL CLINICAL TRIALS

Yuh-Jenn Wu[1], Chi-Tian Chen[2] and Chin-Fu Hsiao[†23]

[1]Department of Applied Mathematics, Chung Yuan Christian University

[2]Division of Biostatistics and Bioinformatics, Institute of Population Health Sciences, National Health Research Institutes

[3]Division of Clinical Trial Statistics, Institute of Population Health Sciences, National Health Research Institutes

ABSTRACT

In recent years, the design of multiregional clinical trials (MRCTs) has become a new strategy for global drug development. A MRCT is a clinical trial which incorporates subjects from many countries around the world under the same protocol. After showing the overall efficacy of a drug in all global regions, one can concurrently assess the chance of applying the overall trial results to each region and consequently support new drug approval in each region. To achieve the goal of conducting MRCT, several regulatory guidelines and statistical methods have been established. In this paper, we will briefly describe the regulatory requirements for MRCTs. Also in this article, two statistical approaches for design and evaluation of MRCTs were reviewed. One assumes that the treatment effect is uniform across regions, while the other uses the random effect model to address the heterogeneous treatment effect across regions. Special consideration was also placed on the determination of the number of subjects in a specific region to establish the consistency of treatment effects between the specific region and the entire group.

Key words and phrases: multiregional clinical trial, bridging study, consistent trend.
JEL classification: C12, C13.

[†]Correspondence to: Chin-Fu Hsiao
E-mail: chinfu@nhri.org.tw.

1. Introduction

Taiwan government has identified biotechnology as one of the key technologies for Taiwan in the 21st century and biopharmaceutical industry as the most vital and important industry. Therefore, Taiwan government has spent efforts to promote Taiwan's participation in global development and international clinical study. After a new pharmaceutical product has been commercially brought for marketing in one region (e.g., the United States or European Union) on the basis of its proven efficacy and safety, the pharmaceutical sponsor might look for registration of the product in a new region, e.g., Taiwan. However, to extrapolate the original clinical data to new populations, the differences on race, diet, environment, culture, and medical practice among regions might cause influence upon a medicine's effect. Recently, the International Conference on Harmonisation (ICH) published a guideline entitled "*Ethnic Factors in the Acceptability of Foreign Clinical Data*" known as ICH E5 to provide a general framework for evaluation of the impact of ethnic factors on the efficacy, safety, dosage, and dose regimen (ICH, 1998). More specifically, the ICH E5 guideline suggests that a bridging study be conducted in the new region to provide pharmacodynamic or clinical data on efficacy, safety, dosage, and dose regimen to allow extrapolation of the foreign clinical data to the population of the new region.

However, conducting a bridging study in a new region may cause duplication of clinical evaluation in the new region. Doing so not only demands valuable development resources but also postpones availability of the pharmaceutical product to the needed patients in the new regions. To cut down the drug lag (i.e., time lag for drug approval), simultaneous drug development, submission, and approval in the world is therefore desirable. In recent years, the design of MRCTs has led to a new strategy for global drug development. A MRCT is defined to be a clinical trial which recruits subjects from many countries around the world under the same protocol. After showing the overall efficacy of a drug in all regions, we can concurrently evaluate the possibility of applying the overall trial results to each region and consequently support registration in each region. In many Asian countries, the trend for participating in simultaneous clinical development with clinical trials conducted in Europe and the United States has been quickly growing. In particular, Taiwan, Korea, Hong-Kong, and Singapore have already had much experience in joining the MRCTs.

To evaluate the possibility of applying the overall trial results in a MRCT to each region, the Japanese Ministry of Health, Labour and Welfare (MHLW) published the "*Basic Principles on Global Clinical Trials*" guidance to provide two methods as examples for determining the number of Japanese subjects required for assessing consistency in treatment effect between the Japanese group and the entire group. For Method 1, it is recommended that the sample size for Japan should meet

$$P(D_{\text{Japan}}/D_{\text{All}} > \pi) > \gamma,$$

where D_{Japan} and D_{All} are the observed treatment effects for the Japanese group and the entire group, respectively. For Method 2, suppose there are three regions participated in the MRCT, the sample size should be determined such that

$$P(D_1 > 0, D_2 > 0, D_3 > 0) > \gamma,$$

where D_i represents the observed treatment effect for region $i, i = 1, \ldots, 3$. Note that the Japanese MHLW suggests that π should be chosen to be 0.5 or greater, and γ should be chosen to be 0.80 or greater. On the other hand, the 11^{th} Q & A (ICH, 2006) for the ICH E5 guidance also discusses the concept of a MRCT and says that "It may be desirable in certain situations to achieve the goal of bridging by conducting a multi-regional trial under a common protocol that includes sufficient numbers of patients from each of multiple regions to reach a conclusion about the effect of the drug in all regions. ..." Both guidelines have established a framework for how to demonstrate the efficacy of a drug in all involved regions while also evaluating the possibility of applying the overall trial results to each region by conducting a MRCT.

Based on Method 1 from the MHLW guidance, Ko et al. (2010) established four statistical criteria for assessing consistency between the region of interest and overall results. Based on Method 2, Kawai et al. (2008) developed an approach to rationalize partitioning the total sample size among the regions so that a high probability of showing 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 MRCT. In addition to Japanese guidance, other consistency criteria have also been established in the literature. Uesaka (2009) proposed new statistical criteria of consistency between regional and overall results which do not require impractical sample sizes and discussed methods of sample size allocation to regions. Tsou et al. (2010) provided proposals of statistical consideration to evaluation of results for a specific region in MRCT from an Asian perspective. Tsou et al. (2011) then compared the observed outcome for the specific region with other regions in a MRCT to examine whether the overall results from the entire group can be applied to the region of interest.

In most of the recent approaches for sample size determination in MRCTs, a common treatment effect and an equal variability of the primary endpoint across regions are usually assumed in the literature. In practice, it might be anticipated that there exists a difference in treatment effect due to regional difference (e.g., ethnic difference). Thus, the sample size calculation for MRCTs based on the assumption that the effect size is uniform across regions might be impractical. Consequently, Hung et al. (2010) presented a number of useful statistical analysis tools for exploration of regional differences and a method that may be worth consideration in planning a MRCT. Huang, Chang, and Hsiao (2013) and Wu et al. (2014) then used random effect models to address the issues of heterogeneous treatment effect and variability across regions for design and evaluation of MRCTs.

This paper is organized as follows. In Section 2, we briefly describe Ko et al. (2010) for the design and evaluation of MRCTs in which we assume that treatment effect is uniform across regions. In Section 3, a random effect model for heterogeneous treatment effect across regions proposed by Chen, Hung and Hsiao (2012) is introduced. Discussions are given in Section 4.

2. Effect Size is Uniform Across Regions

Assume that the MRCT is conducted to compare a test product and a placebo control. Let M be the number of regions. Let X_{ij} and Y_{ik} be the efficacy responses for the j^{th} subject and the k^{th} subject in the i^{th} region receiving the test product and the placebo control respectively, $i = 1, \ldots, M, j = 1, \ldots, n_i^T$, and $k = 1, \ldots, n_i^C$. Assume that $X_{ij} \sim N(\mu^T, \sigma^2)$ and $Y_{ik} \sim N(\mu^C, \sigma^2)$, where $N(\mu, \zeta^2)$ represents a normal distribution with mean μ and variance ζ^2. For convention, we assume that σ^2 is known, although in actual practice, σ^2 is unknown and must be estimated from some data. Let $\Delta = \mu^T - \mu^C$. The hypothesis of testing for the overall treatment effect is given as

$$H_0 : \Delta \leqq 0 \text{ vs. } H_A : \Delta > 0. \tag{1}$$

Although the hypothesis is one-sided, the method proposed below can be straightforwardly extended to the two-sided hypothesis. Let Z the test statistic for (1). It can be derived that

$$Z = \frac{(\sum_{i=1}^{M}\sum_{j=1}^{n_i^T} X_{ij})/\sum_{i=1}^{M} n_i^T - (\sum_{i=1}^{M}\sum_{j=1}^{n_i^C} Y_{ij})/\sum_{i=1}^{M} n_i^C}{\sigma\sqrt{\frac{1}{\sum_{i=1}^{M} n_i^T} + \frac{1}{\sum_{i=1}^{M} n_i^C}}}.$$

2.1 Sample Size Determination

Let N denote the total sample size for each group planned for detecting an expected treatment difference $\Delta = \delta$ at the desired significance level α and with power $1 - \beta$. Thus,

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

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

2.2 Applying the Results of the MRCT to the Specific Region

The same consistency criterion of Method 1 in the guidance "*Basic Principles on Global Clinical Trials*" published by the Japanese MHLW can be used to examine whether the overall results from the MRCT can be applied to a specific region.

Let D_i be the observed mean difference in the i^{th} region and D the observed mean difference from all regions. That is,

$$D_i = \bar{X}_{i\cdot} - \bar{Y}_{i\cdot} \text{ and } D = \frac{\sum_{i=1}^{K}\sum_{j=1}^{n_i^T} X_{ij}}{\sum_{i=1}^{K} n_i^T} - \frac{\sum_{i=1}^{K}\sum_{j=1}^{n_i^C} Y_{ij}}{\sum_{i=1}^{K} n_i^C}$$

where $\bar{X}_{i\cdot} = (\sum_{j=1}^{n_i^T} X_{ij})/n_i^T$, and $\bar{Y}_{i\cdot} = (\sum_{j=1}^{n_i^C} Y_{ij})/n_i^C$. Suppose that we are interested in judging whether the treatment is effective in a specific region, say the s^{th} region, where $1 \leqq s \leqq M$. Given that the overall result is significant at α level, we will judge whether the treatment is effective in a specific region by the following criterion:

$$D_s \geqq \rho D, \tag{2}$$

with some pre-specified $0 < \rho < 1$. Here, ρ is recommended to be greater than or equal to 0.5 by the Japanese MHLW.

2.3 Sample Size Determination for the Specific Region

Let p_i represent the proportion of patients out of $2N$ in the i^{th} region, $i = 1, \ldots, M$, where $\Sigma p_i = 1$. We now describe the method of determination of p_s to ensure that the assurance probabilities of criterion (2) given $\Delta = \delta$ are maintained at a desired level, say 80%. The assurance probability of criterion (i) given $\Delta = \delta$ can be expressed as

$$AP = P_\delta(D_s > \rho D | Z > Z_{1-\alpha}) = [\int_{a_1}^{\infty} (\Phi(\frac{u - c_5}{c_4}) - \Phi(\frac{c_3 - c_1 u}{c_2}))\varphi(u)du]/(1-\beta), \tag{3}$$

where $c_1 = \sqrt{p_s}$, $c_2 = \sqrt{1-p_s}$, $c_3 = -Z_{1-\beta}$, $c_4 = \frac{\rho\sqrt{p_s(1-p_s)}}{1-\rho p_s}$,
$c_5 = \frac{(\rho-1)\sqrt{p_s}}{1-\rho p_s}(Z_{1-\alpha} + Z_{1-\beta})$, and $a_1 = c_5 + \frac{c_4(c_3 - c\ c_1)}{c_4 c_1 + c_2}$.

Without loss of generality, we assume that $s = 1$. In other words, we want to see if we can bridge the overall results to the first region. Given $\alpha = 0.025$ and $\rho = 0.5$, Tables 1 shows the assurance probabilities of criterion in (2) for $\beta = 0.2$ and $\beta = 0.1$ with various values of p_1. As it can be seen from Table 1, for $\beta = 0.2$, the sample size for the first region has to be around 30% of the overall sample size to maintain the

assurance probability at 80%. As β decreases to 0.1, if sample size for the first region is around 20% of the overall sample size, the assurance probability for criterion (2) can easily achieve 80% level.

3. Heterogeneous Treatment Effect Across Regions

Again, we focus on the MRCT for comparing a test product with a placebo control on a continuous efficacy endpoint. In this section, we shall continue to use much of the notation developed in Section 2. Different from Section 2, here we assume that $X_{ij} \sim N(\mu_i^T, \sigma_i^2)$ and $Y_{ik} \sim N(\mu_i^C, \sigma_i^2)$. Let θ_i be the treatment difference in the i^{th} region. That is, $\theta_i = \mu_i^T - \mu_i^C$. We assume that $\theta_i \sim N(\theta, \tau^2), i = 1, \ldots, M$ to consider the heterogeneous treatment effect across regions as a random effect. In the MRCT, we are desired to test the following hypotheses for assessing the overall treatment effect

$$H_0 : \theta \leqq 0 \text{ vs. } H_A : \theta > 0. \tag{4}$$

Let $\hat{\theta}_i$ represent the estimator of θ_i. It leads to $\hat{\theta}_i = \bar{X}_{i.} - \bar{Y}_{i.}$, where $\bar{X}_i = (\sum_{j=1}^{n_i^T} X_{ij})/n_i^T$ and $\bar{Y}_i = (\sum_{k=1}^{n_i^C} Y_{ik})/n_i^C$. The general random effect model can accordingly be re-expressed as

$$\hat{\theta}_i = \theta + \nu_i + \varepsilon_i \ ,$$

where $\nu_i \sim N(0, \tau^2)$, $\varepsilon_i \sim N(0, \xi_i^2)$, and ν_i and ε_i are assumed to be independent. It follows that

$$\xi_i^2 = \frac{\sigma_i^2}{n_i^T} + \frac{\sigma_i^2}{n_i^C} = \sigma_i^2(\frac{1}{n_i^T} + \frac{1}{n_i^C}).$$

Consequently, it shows that $\hat{\theta}_i \sim N(\theta, \xi_i^2 + \tau^2)$. Let $\hat{\sigma}_i^2$, $\hat{\tau}^2$, and $\hat{\xi}_i^2$ indicate the estimators of σ_i^2, τ^2, ξ_i^2 and respectively. We can obtain that

$$\hat{\sigma}_i^2 = \frac{(n_i^T - 1)S_{iX}^2 + (n_i^T - 1)S_{iY}^2}{n_i^T + n_i^C - 2} \quad \text{and} \quad \hat{\xi}_i^2 = \hat{\sigma}_i^2(\tfrac{1}{n_i^T} + \tfrac{1}{n_i^C}),$$

where $S_{iX}^2 = \sum_{j=1}^{n_i^T}(X_{ij} - \bar{X}_{i.})^2/(n_i^T - 1)$ and $S_{iY}^2 = \sum_{k=1}^{n_i^C}(Y_{ik} - \bar{Y}_{i.})^2/(n_i^C - 1)$.
Set $w_i^{-1} = \xi_i^2$. Let $\hat{w}_i$ represent the estimator of w_i. By DerSimonian and Laird (1986), the method of moments estimator $\hat{\tau}^2$ for τ^2 is

$$\hat{\tau}^2 = \frac{\hat{w}}{\hat{w}^2 - \Sigma_{i=1}^M \hat{w}_i^2}\{\Sigma_{i=1}^M \hat{w}_i(\hat{\theta}_i - \Sigma_{i=1}^M \frac{\hat{w}_i\hat{\theta}_i}{\hat{w}})^2 - (M-1)\}, \ ,$$

where $\hat{w} = \sum_{i=1}^M \hat{w}_i$. If a negative estimate for τ^2 is obtained, the estimate $\hat{\tau}^2$ should be set to 0. In light of the distributional assumption, we can obtain

$$\hat{\theta}_i \sim N(\theta, w_i^{-1} + \tau^2).$$

Let $w_i^* = (w_i^{-1} + \tau^2)^{-1}$ and $\hat{w}_i^*$ indicate the estimate of w_i^*. Consequently,

$$\hat{w}_i^* = (\hat{w}_i^{-1} + \hat{\tau}^2)^{-1}.$$

Here, we treat the term w_i^{*-1} if it were the true variance of $\hat{\theta}_i$. Doing so gives that the maximum likelihood estimate of θ can be provided by $\hat{\theta}^*$, where

$$\hat{\theta}^* = \sum_{i=1}^{M} \hat{\theta}_i w_i^* / \sum_{i=1}^{M} w_i^*. \tag{5}$$

It should be noted that $\hat{\theta}^*$ is asymptotically unbiased for θ, with variance approximately equal to $1/\sum_{i=1}^{M} \hat{w}_i^*$ and also the number of regions, M, should be small. Let $S = [\sum_{i=1}^{M} \hat{w}_i^*(\hat{\theta}_i - \hat{\theta}^*)^2]/(\sum_{i=1}^{M} \hat{w}_i^*)$. Consequently, by Hartung (1999), $(\sum_{i=1}^{M} \hat{w_i^*})S$ can be estimated by a (central) χ^2-distribution with (M-1) degrees of freedom. By Hartung (1999) again, we can show that $\hat{\theta}^*$ and S are stochastically independent. Therefore, the test statistic for (4) under H_0 is given as

$$T = \frac{\hat{\theta}^* \sqrt{\Sigma_{i=1}^{M} \hat{w}_i^*}}{\sqrt{(\Sigma_{i=1}^{M} \hat{w}_i^*) S/(M-1)}} = \frac{\hat{\theta}^*}{\sqrt{S/(M-1)}} \sim t_{M-1}, \tag{6}$$

where t_n represents the t distribution with degrees of freedom n.

3.1 Sample Size Determination

Given the alternative hypothesis that $\theta = \Delta$, the test statistic T in (6) is distributed as a noncentral t distribution with (M-1) degrees of freedom and noncentrality parameter $\sqrt{(M-1)\sum_{i=1}^{M} w_i^*}\Delta$. Let N indicate the total sample size per group required for discovering an expected treatment difference $\theta = \Delta$ at the desired nominal significance level α and with power $1-\beta$ for the MRCT. Subsequently, once σ_i^2 and τ^2 are given, N requires to satisfy

$$\begin{aligned} 1-\beta &= P(T > t_{\alpha,M-1} | \theta = \Delta) \\ &= 1 - F_{M-1}(t_{\alpha,M-1} | \sqrt{(M-1)(\Sigma_{i=1}^{M} w_i^*)\Delta}\ ) \\ &= 1 - F_{M-1}(t_{\alpha,M-1} | \Delta \sqrt{(M-1)\Sigma_{i=1}^{M} (\frac{2\sigma_i^2}{p_i N} + \tau^2)}\ ), \end{aligned} \tag{7}$$

where $t_{1-\alpha,n}$ indicates the upper α_{th} percentile of the t distribution with degrees of freedom n, and $F_{M-1}(\cdot|\delta)$ is the cumulative distribution function of the noncentral t-distribution with $(M-1)$ degrees of freedom and noncentral parameter δ. If $\sigma_i^2 = \sigma^2$, the equation (7) becomes

$$\beta = F_{M-1}(t_{\alpha,M-1} | \theta \sqrt{(M-1)\Sigma_{i=1}^{M} (\frac{2\sigma^2}{p_i N} + \tau^2)}\ ), \tag{8}$$

Given τ^2, σ^2, p_i and θ_i, we can estimate the overall treatment effect Δ by equation (5), and N can thus be evaluated by the above equation with specification of α and β. It should be noted that τ^2 should not be large. If τ^2 is large, this means that some regions may have strong treatment effect, while other regions may not demonstrate any treatment efficacy at all. In this case, the overall treatment may not be meaningful or interpretable for all regions.

3.2 Applying the Results of the MRCT to the Specific Region

Suppose that we wish to see whether the treatment is effective in a specific region, say the s^{th} region, where $1 \leqq s \leqq M$. Once the overall result is demonstrated significance at α level, again by the Method 1 in the guidance published by Japanese MHLW, we will judge whether the treatment effect in a specific region is consistent with the overall results by the following criterion:

$$\hat{\theta}_s > \rho\hat{\theta}^* \ , \tag{9}$$

for some pre-specified $0 < \rho < 1$. Note that the Japanese MHLW suggests that ρ be 0.5 or greater.

3.3 Sample Size Determination for the Specific Region

Here we need to determine the values of p_s to assure that the assurance probability of the consistency criterion (9) under a specific alternative hypothesis $\theta = \Delta$ can achieve a desired level, say 80%. In particular, the assurance probability of criterion (9) can be expressed as

$$AP = P(\hat{\theta}_s > \rho\hat{\theta}^* | T > t_{1-\alpha, M-1}), \tag{10}$$

where P represents the probability measure with respect to $\theta = \Delta, \sigma^2$ and τ^2.

For the purpose of illustration, suppose that a MRCT will be conducted in 3 regions, say Taiwan, European Union (EU), and United States (USA)). Let p_1, p_2, and p_3 be the proportion of patients for those three regions respectively. Here we consider three designs: $p_1 = p_2 < p_3, p_1 < p_2 = p_3$, and $p_1 < p_2 < p_3$. In other words, Taiwan is the smallest region, EU is the second smallest region, while USA is the largest region. For this trial, the sponsor also wish to evaluate whether the overall results from the MRCT can be applied to Taiwan after the overall result is statistically significant.

Set $\alpha = 0.025, \beta = 0.2$, and $\sigma = 20$, Tables 2-4 display the total sample size and the assurance probability for different combinations of design parameters with $(\theta_1, \theta_2, \theta_3) = (10, 15, 15), (\theta_1, \theta_2, \theta_3) = (8, 10, 15), (\theta_1, \theta_2, \theta_3) = (15, 10, 8)$, and $\tau^2 = 4$ respectively. The first panel in tables corresponds to $p_1 = p_2 < p_3$, the second panel corresponds to $p_1 < p_2 = p_3$, and the third panel corresponds to $p_1 < p_2 < p_3$. For instances, the first line in Table 1 corresponds to a design with $(p_1, p_2, p_3) =$

$(0.1, 0.1, 0.8), (\theta_1, \theta_2, \theta_3) = (10, 15, 15)$, and $\tau^2 = 4$. As we can see, in this case, the total sample size required per group in the MRCT is 76, and the assurance probability for Taiwan is 0.80.

When $(\theta_1, \theta_2, \theta_3) = (10, 15, 15)$ or $(\theta_1, \theta_2, \theta_3) = (8, 10, 15)$, it can be seen that the required sample size per group for the MRCT increases as p_1 increases. This makes intuitive sense since the overall treatment effect becomes smaller as p_1 increases. On the other hand, if $(\theta_1, \theta_2, \theta_3) = (15, 10, 8)$, the required sample size per group for the MRCT decreases as p_1 increases. In all cases, the assurance probability for the specific region increases as p_1 increases. This phenomenon occurs since the observed overall result $\hat{\theta}^*$ will be increasingly dominated by the observed result $\hat{\theta}_1$ from the specific region as p_1 increases.

4. Discussions

A MRCT recruits subjects from many countries/regions around the world under the same protocol. After showing the overall efficacy of a drug in all global regions, the pharmaceutical sponsor may concurrently wish to assess the possibility of applying the overall trial results to each region and consequently support new drug registration in each region. In this paper, we describe two strategies for design and evaluation of MRCTs. One assumes that the treatment effect is uniform across regions, while the other one uses the random effect model to address the heterogeneous treatment effect across regions.

Discovering the ethnic difference for a given pharmaceutical product is rather important. Recently, it has been shown by many studies that ethnic difference may mediate variability among persons in the response to a therapeutic agent (Eichebaum (1979), Mahgoab et al. (1997), and Caraco (2004)). In other words, after the intake of identical doses of a given agent, some patients may have strong therapeutic response, while others may not have any response or have clinically significant side effects. Since conducting MRCTs will incorporate subjects from many countries around the world, a difference in treatment effect due to ethnic differences might be reasonably expected. In planning a MRCT, the regions involved are strongly encouraged to join in the global development from the stage of exploration. Doing so may be able to estimate the effect size for each region prior to planning a phase III confirmatory trial. The ethnic differences may thus be possibly discovered from exploratory studies.

When the treatment effect only exists in some regions, care should be taken on whether or not the overall treatment effect in (4) is meaningful or interpretable for all regions. In other words, if there exist regional differences for a given therapeutic agent, then only regions that possibly satisfy the minimal clinical meaningfulness requirement can be chosen to participate in the MRCT. Note that if regional differences are due to trial conduct / data quality among regions, the results of the entire trial might not be interpretable.

MRCTs might have benefits on decreasing Taiwan patients' exposures on un-marketed drugs, reducing drug lag, and increasing available treatment options. However, it should be noted that more management skills and preparation time due to various cultures, languages, religions, and medical practices may be required for conducting MRCTs.

Table 1. The assurance probability for observing similarity criterion given α =0.025, β=0.1 and 0.2, and ρ = 0.50

	AP	
p_1	β=0.2	β=0.1
0.10	0.70	0.72
0.20	0.78	0.80
0.30	0.84	0.86
0.40	0.89	0.91
0.50	0.93	0.94
0.60	0.96	0.97
0.70	0.98	0.99
0.80	1.00	1.00
0.90	1.00	1.00

Table 2. The required sample size per group and the assurance probability for the consistency criterion for θ_1=10, θ_2=15, θ_3=15, τ^2=4, α=0.025, β=0.2, and ρ= 0.5 with $\sigma_i = \sigma$ = 20.

p_1	p_2	p_3	*N*	AP
		$p_1 = p_2 < p_3$		
0.1	0.1	0.8	76	0.80
0.15	0.15	0.7	76	0.88
0.2	0.2	0.6	76	0.93
0.25	0.25	0.5	78	0.97
0.3	0.3	0.4	80	0.99
		$p_1 < p_2 = p_3$		
0.1	0.45	0.45	70	0.79
0.15	0.425	0.425	72	0.86
0.2	0.4	0.4	74	0.92
0.25	0.375	0.375	77	0.97
0.3	0.35	0.35	80	0.99
		$p_1 < p_2 < p_3$		
0.1	0.25	0.65	72	0.82
0.15	0.3	0.55	73	0.86
0.2	0.35	0.45	74	0.92

Table 3. The required sample size per group and the assurance probability for the consistency criterion for θ_1=8, θ_2=10, θ_3=15, τ^2=4, α=0.025, β=0.2, and ρ= 0.5 with $\sigma_i = \sigma = 20$.

p_1	p_2	p_3	N	AP
		$p_1 = p_2 < p_3$		
0.1	0.1	0.8	86	0.80
0.15	0.15	0.7	91	0.87
0.2	0.2	0.6	98	0.92
0.25	0.25	0.5	107	0.97
0.3	0.3	0.4	119	0.99
		$p_1 < p_2 = p_3$		
0.1	0.45	0.45	108	0.80
0.15	0.425	0.425	111	0.88
0.2	0.4	0.4	115	0.93
0.25	0.375	0.375	119	0.97
0.3	0.35	0.35	125	0.99
		$p_1 < p_2 < p_3$		
0.1	0.25	0.65	92	0.81
0.15	0.3	0.55	99	0.87
0.2	0.35	0.45	109	0.93

Table 4. The required sample size per group and the assurance probability for the consistency criterion for θ_1=15, θ_2=10, θ_3=8, τ^2=4, α=0.025, β=0.2, and ρ= 0.5 with $\sigma_i = \sigma = 20$.

p_1	p_2	p_3	N	AP
		$p_1 = p_2 < p_3$		
0.1	0.1	0.8	301	0.98
0.15	0.15	0.7	226	0.99
0.2	0.2	0.6	184	0.99
0.25	0.25	0.5	157	0.99
0.3	0.3	0.4	138	0.99
		$p_1 < p_2 = p_3$		
0.1	0.45	0.45	194	0.9576
0.15	0.425	0.425	173	0.9815
0.2	0.4	0.4	158	0.9937
0.25	0.375	0.375	145	0.998
0.3	0.35	0.35	135	0.9995
		$p_1 < p_2 < p_3$		
0.1	0.25	0.65	231	0.97
0.15	0.3	0.55	189	0.98
0.2	0.35	0.45	162	0.99

References

Caraco, Y. (2004). Genes and the response to drugs. *The New England Journal of Medicine*, **351(27)**, 2867-2869.

Chen, C. T., Hung, H. M. J., and Hsiao, C. F. (2012). Design and evaluation of Multiregional trials with heterogeneous treatment effect across region. *Journal of Biopharmaceutical Statistics*, **22**, 1037-1050.

DerSimonia, R. and Laird, N. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials*, **7**, 177-188.

Eichebaum, M., Spannbrucker, N., and Steincke, B. et al. (1979). Defective N-oxidation of sparteine in man: a new pharmacogenetic defect. *European Journal of Clinical Pharmacology* , **16**, 183-7.

Hung, H. M. J., Wang, S. J., and O'Neill, R. (2010). Consideration of regional difference in design and analysis of multi-regional trials. *Pharmaceutical Statistics*, **9(3)**, 173-178.

ICH. (1998). International Conference on Harmonisation. *Tripartite Guidance E5 Ethnic Factors in the Acceptability of Foreign Data.* The U.S. Federal Register **83**, 31790-31796.

ICH. (2006). International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. *Q&A for the ICH E5 Guldeline on Ethnic Factors in the Acceptability of Foreign Data.* Available at: http:// www.ich.org/LOB/media/MEDIA 1194.pdf.

Hartung, J. (1999). An alternative method for meta-analysis. *Biometrical Journal*, **41(8)**, 901-916.

Huang, Y. F., Chang, W. J., and Hsiao, C. F. (2013). An empirical Bayes approach to evaluation of results for a specific region in multi-regional clinical trials. *Pharmaceutical Statistics*, **11(4)**, 295-299.

Kawai, N., Chuang-Stein, C., Komiyama, O., and Li, Y. (2008). An approach to rationalize partitioning sample size into individual regions in a multiregional trial. *Drug Information Journal*, **42**, 139-147.

Ko, F. S., Tsou, H. H., Liu, J. P., and Hsiao, C. F. (2010). Sample size determination for a specific region in a multi-regional trial. *Journal of Biopharmaceutical Statistics*, **20**, 870-885.

Mahgoab, A., Idle, J. R., and Drig, L. G. et al. (1977). Polymorphic hydroxylation of debrisoquine in man. *Lancet*, **2**, 584-586.

Ministry of Health, Labour and Welfare of Japan (MHLW). (2007). *Basic Principles on Global Clinical Trials*. Tokyo:MHLW.

Tsou, H. H., Chow, S. C., Lan, K. K. G., Liu, J. P., Wang, M., Chern, H. D., Ho, L. T., Hsiung, C. A., and Hsiao, C. F*. (2010). Proposals of statistical consideration to evaluation of results for a specific region in multi-regional trials - Asian perspective. *Pharmaceutical Statistics*, **9**, 201-206.

Tsou, H. H., Chien, T. Y., Liu, J. P., and Hsiao, C. F. (2011). A consistency approach to evaluation of bridging studies and multiregional trials. *Statistics in Medicine*, **17**, 2171-2186.

Uesaka, H. (2007). Sample size allocation to regions in a multiregional trial. *Journal of Biopharmaceutical Statistics*, **19**, 580-594.

Wu, Y. J., Tan, T. S., Chow, S. C., and Hsiao, C. F. (2014). Sample size estimation of multiregional clinical trials with heterogeneous variability across regions. *Journal of Biopharmaceutical Statistics*, **24**, 254-271.

[Received June 2014; accepted November 2014.]

Journal of the Chinese Statistical Association
Vol. 53, (2015) 64–78

多區域臨床試驗之設計與評估之概觀

吳裕振[1] 陳啓天[2] 蕭金福[23]

[1]中原大學應用數學系
[2]國家衛生研究院群體科學研究所生物統計與生物資訊研究組
[3]國家衛生研究院群體科學研究所臨床試驗統計組

摘 要

近年來, 多區域臨床試驗的設計已成爲全球性生技藥品研發一項嶄新的策略。所謂多區域臨床試驗意旨同時在全球多個國家且在同一臨床試驗計畫書所執行之臨床試驗。其主要目的爲當全球資料顯示藥物具有效能性時, 亦能夠同時在臨床試驗所涵蓋的國家中, 同步申請許可、以及同步通過上市。爲達到多區域臨床試驗的目標, 許多法規與統計方法已陸續建立。本文首先會簡要地探討一些多區域臨床試驗的法規。此外, 我們亦將介紹兩種多區域臨床試驗的評估與設計。其一假設藥物效能性無區域間之差異; 其二則假設區域間具藥物效能性之差異。最後, 我們亦針對多區域臨床試驗的執行, 提出合理分配各區域樣本數之方法, 期能使各區域新藥效能性達成一致性之機率爲最高。

關鍵詞: 多區域臨床試驗, 銜接性試驗, 一致性。
JEL classification: C12, C13.