Strategies in designing interim analyses under Discrete Random-Effects Model in a multiregional trial

Hsiao-Hui Tsou [1,2*], Chi-Tian Chen [1], K. K. Gordon Lan [3]
[1] Institute of Population Health Sciences, National Health Research Institutes, Zhunan, Miaoli County, Taiwan, ROC
[2] Graduate Institute of Biostatistics, College of Public Health, China Medical University, Taichung, Taiwan, ROC
[3]Janssen R & D, Pharmaceutical Companies of Johnson & Johnson, Raritan, NJ, USA

[*]Correspondence to: Hsiao-Hui Tsou, Division of Biostatistics and Bioinformatics, Institute of Population Health Sciences, National Health Research Institutes, 35 Keyan Road, Zhunan, Miaoli County 350, Taiwan.
E-mail: tsouhh@nhri.org.tw

Abstract
When designing a trial, the sample size is usually estimated from limited information. It would be desirable to modify total sample size or adjust sample size allocation using accumulated data for increasing the efficiency of clinical trials. In this research, we focus on the design of a multiregional trial since developing pharmaceutical products via multiregional clinical trials has become a preferred strategy. We consider a discrete random effects model to account for heterogeneous treatment effect across regions. We propose some strategies for sample size re-estimation or sample size re-allocation to increase the efficiency of a clinical trial based on interim data.

Key Words: Multiregional clinical trials, interim analysis, Discrete Random-Effects Model, sample size allocation

1. Introduction

Developing pharmaceutical products via multiregional clinical trials (MRCTs) has become a popular strategy. However, regional differences, such as from ethnic factors, have been observed in many MRCTs and may have an impact on a test drug's effect. The definition of the overall treatment effect is unclear when there is heterogeneity among regions. How to model regional differences and assess influence of regional differences on trial planning are important challenges in the design and evaluation of MRCTs.

According the final concept paper published by ICH E9(R1): Addendum to Statistical Principles for Clinical Trials on Choosing Appropriate Estimands and Defining Sensitivity Analyses in Clinical Trials, we can see that the term "estimand" has recently attracted much attention in the clinical trials community. An estimand in a clinical trial usually reflects what is to be estimated for demonstrating the benefit of a treatment. Dr. Frank Bretz from Novartis also gave a presentation on "Estimands and Their Role in Clinical Trials" in 2015. In this presentation, he mentioned the importance of building an

estimand framework to help distinguishing between target of estimation (estimand) and method of estimation (estimator). His presentation focuses on traditional clinical trials.

In this article, we focus on multiregional clinical trials. We introduce the primary estimand in an MRCT. We review three models used for estimate the overall treatment effect in MRCTs. We further propose strategy for sample size re-estimation or sample size re-allocation to increase the efficiency of a clinical trial based on interim results.

2. Estimand in a Multi-regional Clinical Trial

In an MRCT, overall treatment effect is usually the primary estimand. What is the overall treatment effect in an MRCT? The overall treatment effect is the treatment difference between the test drug and the control for all participating regions. How to combine information from participating regions to estimate the overall treatment effect is an increasingly important topic in MRCTs. The estimate of overall treatment effect depends on the underlying model assumption. Three models are usually used in MRCTs: Fixed-effects model (FEM), Continuous random-effects model (CREM), Discrete random-effects model (DREM).

2.1 Fixed-effects Model

Traditionally, a common treatment effect across regions is assumed for the design and evaluation of an MRCT. The parameter of interest is the overall treatment effect $\theta = \mu_T - \mu_C$. Let treatment difference in region k be $\theta_k = \mu_{Tk} - \mu_{Ck}$, for $k = 1, 2,\ldots, K$, where μ_{Tk} and μ_{Ck} are the treatment effects of treatment group T and control group C, respectively, in region k. Thus, $\theta = \theta_1 = \theta_2 = \ldots = \theta_K$ under FEM. However, regional heterogeneity in MRCTs has been observed and may have an impact on the treatment effect. Some interesting and insightful papers such as Hung *et al*. (2010) and Wang & Hung (2012) had discussed this issue.

2.2 Continuous Random-effects Model

Continuous random-effects model (CREM) was originally proposed by DerSimonian and Laird for meta-analysis (DerSimonian & Laird, 1986). In MRCTs, CREM is a Level II random effects model with a continuous normal distribution (Lan & Pinheiro, 2012). CREM assume that regional treatment effects are random sample from a normal distribution. That is, the estimate of regional treatment effects under CREM can be presented as $\hat{\theta}_k \sim N(\theta_k, 4\sigma^2 / N_k)$ and $\theta_k \sim N(\theta, \tau^2)$, where σ^2 and τ^2 are the within-region and between-regions variations. Under CREM, the estimator of overall treatment effect $\hat{\theta} = \sum r_k \hat{\theta}_k$ is the most efficient estimator, where $r_k = \left[\text{var}(\hat{\theta}_k)\right]^{-1} \Big/ \sum\left[\text{var}(\hat{\theta}_k)\right]^{-1}$. Many researchers intended to use CREM for solving the problem of heterogeneous treatment effects across regions. In this article, we elaborated on the undesirable characteristics of CREM in the Appendix and suggested another random effects model (DREM) as an alternative option.

2.3 Discrete Random-effects Model

Recognizing regional treatment differences are not random samples from a normal distribution, Lan and Pinheiro (2012) proposed a discrete random effects models (DREM) to account for between-region variability. Lan *et al.* (2014) applied DREM from a continuous endpoint to binary and time-to-event endpoints. Liu *et al.* (2016) further

derived an optimal sample size allocation over regions to maximize the power of consistency and provided some guidelines on the design of MRCTs. For simplicity, we focus on the cases of two parallel treatment arms: a test product T and a placebo control C with a 1:1 patient allocation ratio. Under the DREM, patient population is partitioned into K disjoint clinical regions S_1, S_2, …, S_K, with the probability $P(S_k) = W_k$, where the weights W_k is the population weight in region k. The sample size N is the sum of the K regional sample sizes, $N = \sum_{k=1}^{K}(N_k^T + N_k^C)$. We further assume the total sample size $N^T + N^C = N$, and that $N_k^i = N_k$ for $i = T, C$. For the jth patient in treatment i allocated to region k, the treatment effect for a randomly selected patient in the population is ν_j^i and the weight $P(\nu_j^i = \mu_{ik}) = W_k$, for $k = 1, 2, \ldots, K$. The distribution F_i for ν_j^i is then defined by the possible values $\{\mu_{i1}, \mu_{i2}, \ldots, \mu_{iK}\}$ and their respective probabilities $\{W_1, W_2, \ldots, W_K\}$. The overall treatment difference θ is defined as the weighted sum of the individual effect of regional difference, denoted by $\theta = \sum_{k=1}^{K} W_k\theta_k$, where θ_k is the treatment effect in the kth region.

2.3.1 Hypothesis testing

The testing hypothesis of the overall treatment effect is

$$H_0: \theta = 0 \text{ vs. } H_A: \theta \neq 0.$$

We are interested only in $\theta > 0$. For normal or continuous responses, a standard two-sample test statistic is

$$Z_{DREM} = \frac{\hat{\theta}}{\sqrt{\text{Var}(\hat{\theta})}} = \frac{\hat{\theta}}{\sqrt{(4\sigma^2 + 2\tau^2)/N}}$$

which is approximated to normal distribution with a sufficiently large sample size N. The new treatment T would be claimed beneficial to patients if $Z_{DREM} > z_{1-\alpha}$, where $z_{1-\alpha}$ denotes the $(1-\alpha)$th percentile of the standard normal distribution.

3. Adaptive Strategy Based on Interim results

When designing an MRCT, the sample size is usually estimated from limited information. It would be desirable to modify total sample size or adjust sample size allocation to some regions using interim data for increasing the efficiency of the trial. However, the sample size re-estimation or sample size re-allocation to regions based on the interim results is a data-driven adaption. Thus, the estimates of parameters and the Z-score statistic are biased. The bias should be adjusted appropriately. .

Chen *et al* (2017) proposed a statistical approach under the case that the data from the region with the minimal observed treatment effect is excluded from the analysis in order to attain the regulatory approval of the study drug. They considered one-stage analysis in an MRCT and provided a drop-minimum data analysis first formulated within the FEM and then extend it to DREM.

In this article, we borrow their idea to adjust the bias of parameters estimates and the Z-score statistic from adaption based on the interim results. The bias is calculated as $B = E[\tilde{\theta} - \theta] = E[\sum_1^K W_k^*(\hat{\theta}_k - \theta_k)]$, where W_k^* is the new weight after adaption based on the interim data. Let $V = Var[\tilde{\theta}]$. Finally, we have the adjusted test statistic

$$Z_{Adj} = (\tilde{\theta} - B) / \sqrt{V}$$

The approximate normality of Z_{Adj} could be inferred by Chen *et al* (2017).

4. Conclusion

In this article, we introduce the definition of estimand in a traditional clinical trial and further illustrate the primary estimand in a multiregional clinical trial. We review three models used to estimate the overall treatment effect in an MRCT. We proposed an adaptation procedure based on interim data which may increase the efficiency of an MRCT. We illustrated the possible bias of parameters estimates and the Z-score statistic. Moreover, we provided a solution to adjust the biased test statistics for testing the overall treatment effect.

Appendix: Possible problems of CREM

The estimate of regional treatment effects under CREM can be presented as $\hat{\theta}_k \sim N(\theta_k, 4\sigma^2 / N_k)$ and $\theta_k \sim N(\theta, \tau^2)$, where σ^2 and τ^2 are the within-region and between-regions variations. Some possible problems of CREM are listed as follows.

(a) Conditionally, θ_k could be different, but deviations come from random noise. Unconditionally, $\mathrm{Var}(\hat{\theta}_k) = \mathrm{E}[\mathrm{Var}(\hat{\theta}_k|\theta_k)] + \mathrm{Var}[\mathrm{E}(\hat{\theta}_k|\theta_k)] = 4\sigma^2/\mathrm{N_k} + \tau^2$. The weighted estimate is $\hat{\theta} = \sum r_k \hat{\theta}_k \sim N(\theta, 4\sigma^2 / N + \tau^2 \sum r_k^2)$, where the weight $r_k = [\mathrm{Var}(\hat{\theta}_k)]^{-1} / \sum [\mathrm{Var}(\hat{\theta}_k)]^{-1}$. Under CREM, all $\{\theta_k\}$ are i.i.d. As a result, the unconditional regional treatment effects are identical since $\mathrm{E}(\theta_k) = \theta$ for all k. Also, all the errors $\{\theta_k - \theta\}$ have the same distribution; they contend no information on regional heterogeneity.

(b) CREM uses the reciprocal of variance as the weight to average regional treatment effects. Thus, the overall treatment effect under CREM may not be easily interpreted due to the complicated weight structure.

(c) CREM uses the reciprocal of variance as a weight to combine the treatment effects. The reciprocal variance weights, $(4\sigma^2/N_k + \tau^2)^{-1}$, depend on the trial sample size, within-region variation, and between-regions variation. Note that $4\sigma^2/N_k$ becomes negligible compared with τ^2, where $\tau^2 > 0$ and N_k tends to infinity for all k. In large sample cases, all the variances $\mathrm{Var}(\hat{\theta}_k)$ are approximately equal to a constant τ^2 for all $k = 1,\ldots, K$. Thus, CREM pushes the reciprocal variance weights to equal weights $(1/K)$ for all regions in large sample cases.

(d) The CREM design employs weights different from the sample-size weights. As a result, it violates the one-patient-one-weight principle. In an MRCT, it is possible that some regions are pooled at the end of the trial. For example, suppose 5 (K=5) regions participated in an MRCT and CREM was used to combine regional treatment effects for estimating the overall treatment effect when designing the MRCT. The weight for each region was close to 1/5 at the design stage of the MRCT since all weights are pushed to $1/K$ under CREM. Suppose three regions are pooled at the end of the trial for some region. Then there are a total of 3 (K=3) regions at the analysis stage of the MRCT and the reciprocal of variance weight under CREM is close to 1/3 for combining regional treatment effects. Thus, the weights at the end of a trial could be very different from the pre-determined weights (change from 1/5 to 1/3). This violates the one-patient-one-weight principle.

(e) As the regional sample size N_k tends to infinity, the regional variance $\text{Var}(\hat{\theta}_k) = 4\sigma^2/N_k + \tau^2$ will tend to τ^2, where τ^2 is assumed not negligible. Therefore, the regional variance $\text{Var}(\hat{\theta}) \approx \tau^2/K$ does not go to zero when the trial sample size is large. The power of the Z-test will not go to 1 even when all N_k tend to infinity. This means that a large number of regions are also required for sufficient power ($K \to \infty$ is also required for power $\to 1$).

Acknowledgements

This work was supported by grants from the National Health Research Institutes, Zhunan, Taiwan [PH-105-PP-11, PH-106-PP-02] and the Ministry of Science and Technology of Republic of China [MOST 105-2118-M-400-002].

References

Bretz F. Acknowledged with ICH E9(R1) Expert Working Group and Akacha M, Ohlssen D, Schmidli H, Rosenkranz G. Estimands and their role in clinical trials. EFSPI/PSI-UK. 2015. http://www.psiweb.org/docs/default-source/resources/psi-subgroups/scientific/2015/estimands-28-09-2015/frank-bretz.pdf?sfvrsn=2. Accessed on Oct. 1, 2017.

Chen CT, Lan KK Gordon, Tsou HH. Estimation of treatment effect in multiregional clinical trials. *Submitted for publication*. 2017.

Chen F, Li G, Lan KK Gordon. Inconsistency and drop-minimum data analysis. *Statistics in Medicine* 2017; 36:416–425. DOI: 10.1002/sim.7166.

DerSimonian R, Laird N. Meta-analysis in clinical trials. *Controlled Clinical Trials* 1986; **7**:177–188.

Hung HMJ, Wang SJ, O'Neill RT. Consideration of regional difference in design and analysis of multi-regional trials. *Pharmaceutical Statistics*. 2010; 24:173–178. DOI:10.1002/pst.440.

Lan KKG, Pinheiro J. Combined estimation of treatment effects under a discrete random effects model. *Statistics in Biosciences* 2012; 4:235–244.

Lan KKG, Pinheiro, J, Chen F. Designing multiregional trials under the discrete random effects model. *Journal of Biopharmaceutical Statistics* 2014; 24:415–428.

Liu JT, Tsou HH, Lan KK Gordon, Chen CT, Lai YH, Chang WJ, Tzeng CS, Hsiao CF. Assessing the consistency of the treatment effect under the discrete random effects model in multiregional clinical trials. *Statistics in Medicine* 2016; 35:2301–2314. DOI: 10.1002/sim.6869.

Wang SJ, Hung HMJ. Ethnic sensitive or molecular sensitive beyond all regions being equal in multiregional clinical trials. *Journal of Biopharmaceutical Statistics* 2012; **22**:879–893.