Design Consideration for Simultaneous Global Drug Development Program (SGDDP)*

Qin Huang[1](黄钦),Gang Chen[2] (陈刚)
DIA China, May 18, 2011

1. Office of Biostatistics, Center for Drug Evaluation, SFDA, Beijing, P.R. China.
2. Biostatistics and Programming, Johnson & Johnson, New Jersey, USA

*The views expressed in this article are personal opinions of the authors and may not necessarily represent the position of State Food and Drug Administration (SFDA).

Outline

- Challenges in current simultaneous global drug development
- Existing methods and concerns
- A new method for the design of SGDDP
- Future consideration and discussion

Challenges on SGDDP with MRCT Data

- Interpretation of treatment effect between overall and a region
  - "consistency" or "similarity"?
- Validity and quality
- Monitoring for regional differences: conduct and outcomes
- Single large study – need additional study/data?
- Large regional heterogeneity may lead to non-approval in US FDA: 2001-2007
  - 9/22 approvable but more information needed
  - 4/22 non-approval because of regional heterogeneity

法规解读

- 如果不考虑在中国注册上市：样本量无要求

- MRCT数据用于国内上市注册（首先在国外上市）：

  ➢ 如果中国部分结果与总体结果可以判断基本一致，则还需达到进口药品注册的临床试验要求：
  药代+至少100对受试者

  ➢ 如果中国部分结果与总体结果无法判断一致性，则按照适宜的桥接策略或重新进行针对中国人群的临床试验（要求具有统计显著性）

Three Major Issues

- What is scientific justification for "at least 100 pairs"?

- How do we define and conclude the "consistency"?

- If necessary, how do we design an acceptable "bridging trial" for different cases?

Existing Regulatory Guidelines

- ICHE5 provides a general principle for the evaluation of the impact of ethnic factors on treatment effect.

- MHLW (Japan) "Basic Principles on Global Clinical Trials" provides a method for the Japanese subgroup sample size requirement in a MRCT.

Japan Guidance

- The Ministry of Health, Labor and Welfare of Japan (MHLW) (2007) issued a guidance, "Basic Principles on Global Clinical Trials", to deal with the "drug lag" problem in Japan,

- The method in the guidance recommended that the sample size for the Japanese patients in a MRCT should satisfy:

$$\Pr(D_j / D_{all} > \pi) \quad 1 - \beta', \tag{1.1}$$

  where π is 0.5 or greater and β' is 0.2 or less; D_{all} and D_J represent the estimated treatment effect for entire target regions and for Japan, respectively.

Japan Guidance

There are two major concerns regarding the guidance.

1. The selection of π and β' the is arbitrary/subjective.

2. The conclusion of similarity or consistency of treatment effect based on the formula (1.1) is difficult to interpret

Major Concerns

Major issues when claiming the similarity or consistency of the treatment effect between different ethnic groups.

1. What should be the definition of "similarity" or "consistency"?
2. How do we control the overall "type I" or "false positive" error rate for the claim.
3. Are the results of TE subgroup in a MRCT adequate to support local registration and how to interpret the subgroup results?
4. How do we analyze the data combining TE patients in MRCT and a local clinical trial (LCT)?

SGDDP Design

- There is no complete solution to the issue and we need to think beyond current paradigm of using "MRCT" or "bridging" to deal with potential impact of ethnic factors – need a new method

- The proposed design is based the following fundamental assumption:
  - Patients in TE group and in NTE group share some level of biological commonality of humankind

- With the above assumption, it is reasonable to consider that the efficacy information collected from different ethnic groups in a SGDDP can be combined in certain manner.

SGDDP Design

Multi-regional Clinical Trial (MRCT)	→	Local Clinical Trial (LCT)

Simultaneous Global Drug Development Program (SGDDP)
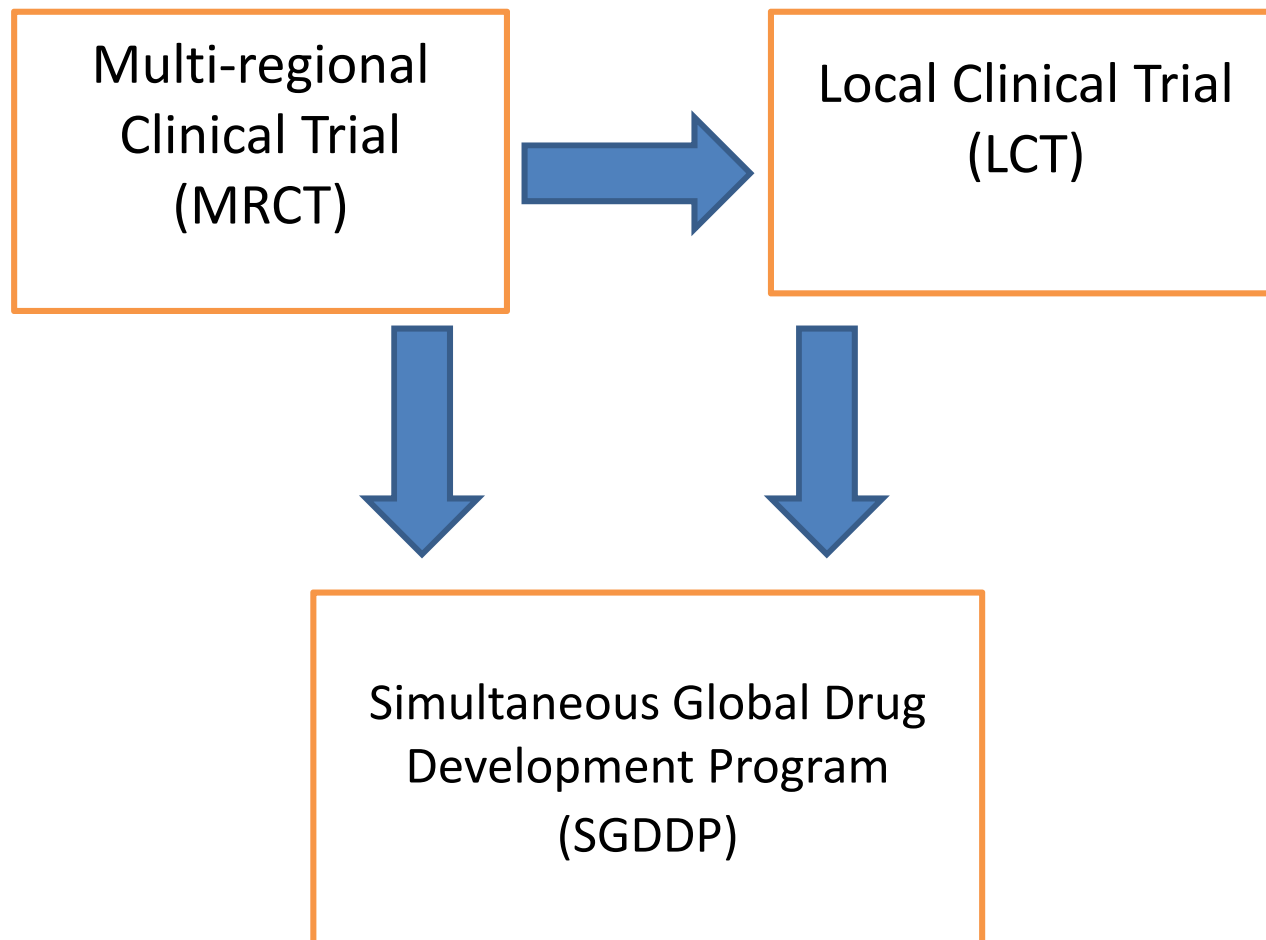
SGDDP Design

The design features for the SGDDP can be outlined below:

- The purpose: to test the new treatment is **effective for** the TE population.

- The SGDDP is a **prospectively designed** program

- The effect size hypothesized for TE group is the **same as or smaller** for the overall group.

- The overall false positive rate ("type I error") controlled at a pre-specified α level.

- Same key design features for the MRCT and the LCT in the SGDDP,
  - e.g., inclusion and exclusion criteria, etc. are same.

- The final analysis of the SGDDP is based on the combined weighted information collected from both TE and NTE groups enrolled in SGDDP. The information collected from NTE group in MRCT will be down-weighted because of the potential impact of ethnic factors.

SGDDP Results Interpretation

Since the design of a SGDDP is prospective and patient enrollment in a SGDDP is sequential, when the results from a MRCT are available, the decision rules regarding the efficacy results of the SGDDP can be described as follows:

- If the efficacy results of the MRCT in the SGDDP are positive (p   α) and data has been used for global regulatory registration, the data collected from the SGDDP will be analyzed at α level again. The result will be considered positive for local regulatory registration if overall p <= α.

SGDDP Results Interpretation

- If the efficacy results of the MRCT in a SGDDP are promising (e.g., $\alpha <$ p $<2\alpha$) and data may not be adequate for US or EU registration, the data collected from the SGDDP may still be analyzed at α level again. The results of the SGDDP may be considered positive if overall p $<= \alpha$. Whether this evidence is adequate for local regulatory registration needs to be considered case by case and pending for more discussion.

- If the efficacy results of the MRCT in a SGDDP are not promising, the SGDDP should be stopped and results of the SGDDP will be considered exploratory.

SGDDP Results Interpretation

Two important notes:

- The criteria/definition for "promising" should be considered case by case

- The MRCT stage plays a role of "futility interim analysis" in SGDDP since SGDDP will stop only if MRCT is negative. Therefore, the overall false positive rate will not be inflated due to MRCT analysis.

Statistical Method

- The stat method proposed for the design of the SGDDP is based on the weighted Z test which was discussed by Lan, Soo, Siu and Wang (2005)*.

- That is, the information collected from NTE patients in the SGDDP will be down-weighted.

- The weight (w) is 0   w   1 and should be pre-specified.

* Lan, K.K. G., Soo, Y., Siu, C., Wang, M. (2005). The use of weighted Z-tests in medical research. *Journal of Biopharmaceutical Statistics*. 15: 625–639.

Statistical Method

- The following assumptions are used in the sample size calculation:
  - Endpoints: Normal, Binary, Time to event
  - The type I error for MRCT=0.05 (2-sided),
  - Overall false positive rate for SGDDP=0.05 (2-sided),
  - Power =80% for both MRCT and SGDDP
  - Sample size for MRCT = 500 (calculated using above assumptions and appropriate effect size).
  - Denote n_{1p} as the total sample size needed for the TE group.

Statistical Method

- Let Z_1 and Z_2 be the test statistics for the TE group and NTE group, respectively in SGDDP. A weighted Z test statistic combining the information collected from both subgroups in SGDDP can be constructed as follows:

$$Z = \sqrt{1-w}Z_1 + \sqrt{w}Z_2$$

where $0 \le w \le 1$.

Sample Size Calculation for LCT

Sample size calculation has been considered for 3 endpoints: continuous, binary and time to event. As example, the sample size for continuous endpoint will be discussed here.

- If the endpoint of a SGDDP can be measured as a continuous random variable with a normal distribution and Let $E_1 = E_A(Z_1)$ be the mean of Z_1 under the alternative hypothesis H_A, the sample size for LCT can be calculated from the following formula.

$$E_1 = \frac{\delta_1}{\sigma_1 \sqrt{\dfrac{1}{n_{1p}/2} + \dfrac{1}{n_{1p}/2}}} = \frac{1}{2}\sqrt{n_{1p}}\,\delta_1 / \sigma_1$$

Table 1: Sample Size for LCT

MRCT size=500 with 10% TE pts, MRCT effect size = 0.25, s.e.=1

Weight (wt)	TE group effect size		
	0.15	0.20	0.25
0%	1346	735	453
10%	683	371	225
20%	493	266	159
30%	**370**	**198**	**117**
40%	277	148	85
50%	203	107	60
60%	142	74	40
70%	90	45	23

Sample Size for LCT

- As shown in Table 1, if 50 patients (10%) are enrolled from TE population in MRCT and we assume the treatment effect in TE group is slightly smaller than the overall effect in MRCT, e.g., = 0.20, then if using w=30%, additional 198 patients are needed for LCT in SGDDP to maintain the power for the detection of smaller effect size in TE.

- Therefore, in this case an approximately 248 (50+198) total TE patients need to be enrolled in the SGDDP.

Sample Size for LCT

- It is important to note that if choosing w=0, in other words, we will completely abandon the information collected from NTE group, then we need enroll 453 total patients in LCT assuming treatment effect size is the same (0.25). By adding the 50 TE patients enrolled in MRCT, a total of 503 TE patients are required for the SGDDP with design power=80% and alpha level of 0.05 (2-sided).

Discussion

- The proposed method provides the design and sample size consideration for a SGDDP to fulfill the requirement of global regulatory registration of a new treatment.

- The method of this design controls the overall false positive rate under a given significance level, e.g., 0.05 with an adequate test power for the program.

Discussion

- There are several important factors need to be considered when designing a SGDDP:
    - weight selection,
    - endpoint selection,
    - the proportion of TE patients enrolled in MRCT,
    - treatment effect size for TE vs. the NTE population,
    - the variance σ for TE and NTE patients.

Discussion

- **Weight selection**:
  - Selecting w=0 means "two ethnic groups are completely different" and no information from the NTE subgroup will be used.
  - The Largest weight (1) should always be less than the actual proportion of NTE patients in the SGDDP.
    - For example, if in a SGDDP, the MRCT enrolled 400 NTE patients and 100 TE patients and the LCT enrolled 100 additional TE patients, then w should be 400/600 = 0.66.

Discussion

- **Weight selection**: The weight selection should be based on statistical, clinical and regulatory considerations including those factors listed in ICHE5, such as,

    1). PK/PD profiles of TE and NTE patients;

    2). The toxicity profile of the new treatment, e.g., for a highly toxic drug, it may be reasonable to require a smaller weight ;

    3). Local medical practice/conduct.

Discussion

- **Endpoint selection**: The sample size requirement may be different when different endpoint used in the design of a SGDDP. For example, comparing to using a continuous (normal) endpoint, the sample size required for the LCT with a binary endpoint is reduced more quickly when increasing the weight.

Discussion

- **Proportion of TE patients enrolled in MRCT**:

  Ideally, the proportion of the TE patients enrolled in MRCT should be reasonable to make sure the robust and consistent efficacy findings between the overall and TE subgroup. However, in a MRCT, it is often to have dozens of countries participated in the trial. It is not practical to enroll a large proportion of patients from one country which may lead to the difficult interpretation of the efficacy results due to the imbalance patient enrolment between countries (regions).

Discussion

Treatment effect size for TE vs. the overall:

- In the design of a SGDDP, we assume that the treatment effect of TE population is same as or smaller than the treatment effect of the overall effect. This assumption is reasonable since the objective of a SGDDP is to test that the treatment is effective.

- If in some cases, one do believe the treatment effect for TE population is greater than the overall effect, then the LCT may not be necessary if there is a reasonable proportion TE patients in the MRCT according to the proposed design.

Discussion

σ for TE and NTE patients:

- For simplicity, the variance σ for normal (continuous) endpoints in this research is assumed same in both TE and NTE patient population. In practice, the variance associated with treatment effect may be different for different patient populations. The results of sample size calculation are also available when using different variance for different ethnic groups.

Discussion

- **Important note**: In this design the results of either LCT or the TE subgroup cannot stand alone and the results should be considered as exploratory (subgroup analyses) and interpreted with caution. The registration decision should be based on the overall results of the SGDDP.

ACKNOWLEDGMENT

Table 1a: Sample Size for LCT

MRCT size=500 with 20% TE pts, MRCT effect size = 0.25, s.e.=1

Weight (wt)	TE group effect size		
	0.15	0.20	0.25
0%	1296	685	403
10%	638	332	188
20%	449	232	127
30%	**327**	**166**	**88**
40%	235	118	59
50%	162	79	37
60%	101	48	20
70%	49	23	8