Journal of Biopharmaceutical Statistics, 18: 483–493, 2008

ISSN: 1054-3406 print/1520-5711 online
DOI: 10.1080/10543400801995452

Taylor & Francis
Taylor & Francis Group

ON POWER AND SAMPLE SIZE CALCULATION FOR QT STUDIES WITH RECORDING REPLICATES AT GIVEN TIME POINT

Shein-Chung Chow[1], Bin Cheng[2], and Dennis Cosmatos[3]
[1]*Department of Biostatistics and Bioinformatics, Duke University School of Medicine, Durham, North Carolina, USA*
[2]*Department of Biostatistics, Columbia University, New York, New York, USA*
[3]*Biostatistics for Translational Development, Wyeth Research, Collegeville, Pennsylvania, USA*

The problem of the impact on power and sample size calculation for routine QT studies with ECG recording replicates under a parallel-group design and a crossover design is examined. Replicate ECGs are defined as single ECG recorded within several minutes of a nominal time (PhRMA, 2003). Formulas for sample size calculations with and without adjustment for covariates such as some pharmacokinetic responses (e.g., AUC or C_{max}), which are known to be correlated to the QT intervals, were derived under both the parallel-group design and the crossover design. The results indicate that the approach of replicates may require a smaller sample size for achieving the same power when the correlation coefficient between the recording replicates (or repeated measures) is close to 0 (i.e., these replicate ECGs are almost independent). On the other hand, if the correlation coefficient is close to 1, then there is not much gain regardless of whether replicate ECGs are considered. In this paper, an approach to identifying optimal allocation between the number of subjects and the number of replicates per subject is proposed for achieving the maximum power under a fixed budget constraint. The proposed approach can also be applied to minimize the cost for a given power.

Key Words: Correlation coefficient; Crossover design; Measurement error; Parallel-group design; QT Studies.

1. INTRODUCTION

In clinical trials, a 12-lead electrocardiogram (ECG) is usually conducted for assessment of potential cardiotoxicity induced by the treatment under study. On an ECG tracing, the QT interval is measured from the beginning of the Q wave to the end of the T wave. QT interval is often used to indirectly assess the delay in cardiac repolarization, which can predispose one to the development of life-threatening cardiac arrhythmias such as torsade de pointes (Moss, 1993). *QTc interval* refers to

Received May 8, 2007; Accepted February 10, 2008
Address correspondence to Bin Cheng, Department of Biostatistics, Columbia University, New York, NY, USA; E-mail: bc2159@columbia.edu

the QT interval corrected by heart rate. In clinical practice, it is recognized that the prolongation of the QT/QTc interval is related to increased risk of cardiotoxicity such as a life-threatening arrhythmia (Temple, 2003). Thus it is suggested that a careful evaluation of potential QT/QTc prolongation be assessed for potential drug-induced cardiotoxicity.

For development of a new pharmaceutical entity, most regulatory agencies such as the United States Food and Drug Administration (FDA) require the evaluation of proarrhythmic potential (see, e.g., CPMP, 1997; FDA/TPD, 2003; ICH, 2005). As a result, a draft guidance on the clinical evaluation of QT/QTc interval prolongation and proarrythmic potential for non-antiarrythmic drugs is being prepared by the ICH (ICH E14). This draft guidance calls for a placebo-controlled study in normal healthy volunteers with a positive control to assess cardiotoxicity by examining QT/QTc prolongation. Under a valid study design (e.g., a parallel-group design or a crossover design), ECGs will be collected at baseline and at several time points posttreatment for each subject. Malik and Camm (2001) recommend that it would be worthwhile to consider 3 to 5 replicate ECGs at each time point within a 2- to 5-minute period. Replicate ECGs are then defined as a single ECG recorded within several minutes of a nominal time (PhRMA, 2003). Along this line, Strieter et al. (2003) studied the effect of replicate ECGs on QT variability in health subjects. In practice, it is then of interest to investigate the impact of recording replicates on power and sample size calculation in routine QT studies.

In clinical trials, a prestudy power analysis for sample size calculation is usually performed to ensure that the study will achieve a desired power (or the probability of correctly detecting a clinically meaningful difference if such a difference truly exists). For QT studies, the following information is necessarily obtained prior to the conduct of the prestudy power analysis for sample size calculation. The information includes (1) the variability associated with the primary study endpoint such as the QT intervals (or the QT interval endpoint change from baseline); (2) the maximal difference in QT interval between treatment groups; and (3) the number of time points where QT measurements are taken. Under the above assumptions, the procedures as described in Longford (1993); Chow et al. (2003) can then be applied for sample size calculation under the study design (e.g., a parallel-group design or a crossover design). Although QT/QTc studies involve multiple time points, we will consider in this paper the simplified case with only one time point. And we argue that considering one time point, though conservative, is reasonable for sample size determination purposes. This is particularly true if we focus on the time point where the maximal QT difference between treatments is expected.

The remainder of this article is organized as follows. In the next section, commonly used study designs such as a parallel-group design or a crossover design for routine QT studies with recording replicates are briefly described. Power analyses and the corresponding sample size calculations under a parallel-group design and a crossover design are derived in Section 3. Extensions to the designs with covariates (PK responses) are considered in Section 4. The sample size allocation optimization is discussed in Section 5. Section 6 provides some concluding remarks.

2. STUDY DESIGNS AND MODELS

A typical study design for QT studies is either a parallel-group design or a crossover design depending on the PK profile of the investigational drug. Statistical models under a parallel-group design and a crossover design are briefly outlined below.

Under a parallel-group design, qualified subjects will be randomly assigned to receive either treatment A or treatment B. ECGs will be collected at baseline and at several time points posttreatment. Subjects will fast at least 3 hours and rest at least 10 minutes prior to the scheduled ECG measurements. Identical lead-placement and the same ECG machine will be used for all measurements. As recommended by Malik and Camm (2001), 3 to 5 recording replicate ECGs at each time point will be obtained within a 2- to 5-minute period.

Let y_{ijk} be the QT interval observed from the kth recording replicate of the jth subject who receives treatment i, where $i = 1, 2$; $j = 1, \ldots, n$; and $k = 1, \ldots, K$. Consider the following model:

$$y_{ijk} = \mu_i + e_{ij} + \epsilon_{ijk}, \tag{1}$$

where e_{ij} are independent and identically distributed as normal random variables with mean 0 and variance σ_s^2 (between-subject or intersubject variability) and ϵ_{ijk} are independent and identically distributed as normal random variables with mean 0 and variance σ_e^2 (within-subject or intrasubject variability or measurement error variance). Thus, we have $\mathrm{Var}(y_{ijk}) = \sigma_s^2 + \sigma_e^2$.

Under a crossover design, qualified subjects will be randomly assigned to receive one of the two sequences of test treatments under study. In other words, subjects who are randomly assigned to sequence 1 will receive treatment 1 first and then be crossed over to receive treatment 2 after a sufficient period of washout. Let y_{ijkl} be the QT interval observed from the kth recording replicate of the jth subject in the lth sequence who receives the ith treatment, where $i = 1, 2$; $j = 1, \ldots, n$; $k = 1, \ldots, K$; and $l = 1, 2$. We consider the following model:

$$y_{ijkl} = \mu_i + \beta_{il} + e_{ijl} + \epsilon_{ijkl}, \tag{2}$$

where β_{il} are independent and identically distributed normal random period effects (period uniquely determined by sequence l and treatment i) with mean 0 and variance σ_p^2, e_{ijl} are independent and identically distributed normal subject random effects with mean 0 and variance σ_s^2, and ϵ_{ijkl} are independent and identically distributed normal random errors with mean 0 and variance σ_e^2. Thus, $\mathrm{Var}(y_{ijkl}) = \sigma_p^2 + \sigma_s^2 + \sigma_e^2$.

To ensure a valid comparison between the parallel design and the crossover design, we assume that μ_i, σ_s^2, and σ_e^2 are the same as those given in (1) and (2) and consider an extra variability σ_p^2, which is due to the random period effect for the crossover design.

3. POWER AND SAMPLE SIZE CALCULATION

3.1. Parallel-Group Design

Under the parallel-group design as described in the previous section, to evaluate the impact of recording replicates on power and sample size calculation, for simplicity, we will consider only one time point posttreatment. The results for recording replicates at several posttreatment intervals can be similarly obtained. Under model (1), considering the sample mean of QT intervals of the jth subject who receives the ith treatment, $\mathrm{Var}(\bar{y}_{ij.}) = \sigma_s^2 + \frac{\sigma_e^2}{K}$. The hypotheses of interest regarding treatment difference in QT interval are given by

$$H_0 : \mu_1 - \mu_2 \geq 10, \quad \text{versus} \quad H_1 : \mu_1 - \mu_2 < 10. \tag{3}$$

Under the null hypothesis of no treatment difference, the following statistic can be derived:

$$T = \frac{\bar{y}_{1..} - \bar{y}_{2..} - 10}{\sqrt{\frac{2}{n}\left(\hat{\sigma}_s^2 + \frac{\hat{\sigma}_e^2}{K}\right)}},$$

where

$$\hat{\sigma}_e^2 = \frac{1}{2n(K-1)} \sum_{i=1}^{2} \sum_{j=1}^{n} \sum_{k=1}^{K} (y_{ijk} - \bar{y}_{ij.})^2,$$

and

$$\hat{\sigma}_s^2 = \frac{1}{2(n-1)} \sum_{i=1}^{2} \sum_{j=1}^{n} (\bar{y}_{ij.} - \bar{y}_{i..})^2 - \frac{1}{2nK(K-1)} \sum_{i=1}^{2} \sum_{j=1}^{n} \sum_{k=1}^{K} (y_{ijk} - \bar{y}_{ij.})^2.$$

Under the null hypothesis in (3), T has a central t-distribution with $2n-2$ degrees of freedom. Let $\sigma^2 = \mathrm{Var}(y_{ijk}) = \sigma_s^2 + \sigma_e^2$ and $\rho = \frac{\sigma_s^2}{\sigma_s^2 + \sigma_e^2}$; then, under a given alternative $\mu_1 - \mu_2 = d < 10$ in (3), the power of the test can be approximated as follows:

$$1 - \beta \approx \Phi\left(- z_\alpha + \frac{\delta}{\sqrt{\frac{2}{n}\left(\rho + \frac{1-\rho}{K}\right)}} \right), \tag{4}$$

where $\delta = (10 - d)/\sigma$ is the relative effect size and Φ is the cumulative distribution of a standard normal. To achieve the desired power of $1 - \beta$ at the α level of significance, the sample size needed per treatment is

$$n = \frac{2(z_\alpha + z_\beta)^2}{\delta^2} \left(\rho + \frac{1 - \rho}{K} \right). \tag{5}$$

3.2. Crossover Design

Under a crossover model (2), it can be verified that $\bar{y}_{i...}$ is an unbiased estimator of μ_i with variance $\frac{\sigma_p^2}{2} + \frac{\sigma_s^2}{2n} + \frac{\sigma_e^2}{2nK}$. Thus, we used the following test statistic

to test the hypotheses in (3):

$$T = \frac{\bar{y}_{1\ldots} - \bar{y}_{2\ldots} - 10}{\sqrt{\hat{\sigma}_p^2 + \frac{1}{n}\left(\hat{\sigma}_s^2 + \frac{\hat{\sigma}_e^2}{K}\right)}},$$

where

$$\hat{\sigma}_e^2 = \frac{1}{4n(K-1)} \sum_{i=1}^{2} \sum_{j=1}^{n} \sum_{k=1}^{K} \sum_{l=1}^{2} (y_{ijkl} - \bar{y}_{ij.l})^2,$$

$$\hat{\sigma}_s^2 = \frac{1}{4(n-1)} \sum_{i=1}^{2} \sum_{j=1}^{n} \sum_{l=1}^{2} (\bar{y}_{ij.l} - \bar{y}_{i..l})^2 - \frac{1}{4nK(K-1)} \sum_{i=1}^{2} \sum_{j=1}^{n} \sum_{k=1}^{K} \sum_{l=1}^{2} (y_{ijkl} - \bar{y}_{ij.l})^2,$$

and

$$\hat{\sigma}_p^2 = \frac{1}{2} \sum_{i=1}^{2} \sum_{l=1}^{2} (\bar{y}_{i..l} - \bar{y}_{....})^2 - \frac{1}{4n(n-1)} \sum_{i=1}^{2} \sum_{j=1}^{n} \sum_{l=1}^{2} (\bar{y}_{ij.l} - \bar{y}_{i..l})^2.$$

Under the null hypothesis in (3), T has a central t-distribution with $2n-4$ degrees of freedom. Let σ^2 and ρ be defined as in the previous section, and $\gamma = \sigma_p^2/\sigma^2$; then $\mathrm{Var}(y_{ijkl}) = \sigma^2(1+\gamma)$. Under a given alternative $\mu_1 - \mu_2 = d < 10$ in (3), the power of the test can be approximated as follows:

$$1 - \beta \approx \Phi\left(-z_\alpha + \frac{\delta}{\sqrt{\gamma + \frac{1}{n}\left(\rho + \frac{1-\rho}{K}\right)}} \right), \tag{6}$$

where $\delta = (10-d)/\sigma$. To achieve the desired power of $1-\beta$ at the α level of significance, the sample size needed per treatment is

$$n = \frac{(z_\alpha + z_\beta)^2}{\delta^2 - \gamma(z_\alpha + z_\beta)^2} \left(\rho + \frac{1-\rho}{K} \right). \tag{7}$$

3.3. Remarks

Let n_{old} be the sample size with $K = 1$ (i.e., there is single measure for each subject). Then, we have $n = \rho n_{\text{old}} + (1-\rho) n_{\text{old}}/K$. Thus, sample size (with recording replicates) required for achieving the desired power is a weighted average of n_{old} and n_{old}/K. Note that this relationship holds under both a parallel and a crossover design. Table 1 provides sample sizes required under a chosen design (either parallel or crossover) for achieving the same power with single recording ($K = 1$), three recording replicates ($K = 3$), and five recording replicates ($K = 5$).

Note that if ρ closes to 0, then these repeated measures can be treated as independent replicates. As it can be seen from the above, if $\rho \approx 0$, then $n \approx n_{\text{old}}/K$. In other words, sample size is indeed reduced when the correlation coefficient between recording replicates is close to 0 (in this case, the recording replicates are almost independent). Table 2 shows the sample size reduction for different values

Table 1 Sample size for achieving the same power with K recording replicates

	K		
ρ	1	3	5
1.0	n	$1.00n$	$1.00n$
0.9	n	$0.93n$	$0.92n$
0.8	n	$0.86n$	$0.84n$
0.7	n	$0.80n$	$0.76n$
0.6	n	$0.73n$	$0.68n$
0.5	n	$0.66n$	$0.60n$
0.4	n	$0.60n$	$0.52n$
0.3	n	$0.53n$	$0.44n$
0.2	n	$0.46n$	$0.36n$
0.1	n	$0.40n$	$0.28n$
0.0	n	$0.33n$	$0.20n$

of ρ under the parallel design. However, in practice, ρ is expected to be close to 1. In this case, we have $n \approx n_{\text{old}}$. In other words, there is not much gain for considering recording replicates in the study.

In practice it is of interest to know whether the use of a crossover design can further reduce the sample size when other parameters such as d, σ^2, and ρ, remain the same. Comparing formulas (5) and (7), we conclude that the sample size reduction by using a crossover design depends on the parameter $\gamma = \sigma_p^2/\sigma^2$, which is a measure of the relative magnitude of period variability with respect to the within-period subject marginal variability. Let $\theta = \frac{\gamma}{(z_\alpha + z_\beta)^2}$; then, by (5) and (7), the sample size n_{cross} under the crossover design and the sample size n_{parallel} under the parallel-group design satisfy $n_{\text{cross}} = \frac{n_{\text{parallel}}}{2(1-\theta)}$. When the random period effect is negligible, that is, $\gamma \approx 0$ and hence $\theta \approx 0$, we have $n_{\text{cross}} = \frac{n_{\text{parallel}}}{2}$. This indicates that the use of a crossover design could further reduce the sample size by half as compared to a parallel-group design when the random period effect is negligible (based on the comparison of the above formula and the formula given in [5]). However, when the random period effect is not small, the use of a crossover design may not result in sample size reduction. Table 3 shows the sample size under

Table 2 Sample sizes required under a parallel-group design

	Power = 80%					Power = 90%				
	ρ					ρ				
(K, δ)	0.2	0.4	0.6	0.8	1.0	0.2	0.4	0.6	0.8	1.0
(3, 0.3)	81	105	128	151	174	109	140	171	202	233
(3, 0.4)	46	59	72	85	98	61	79	96	114	131
(3, 0.5)	29	38	46	54	63	39	50	64	73	84
(5, 0.3)	63	91	119	147	174	84	121	159	196	233
(5, 0.4)	35	51	67	82	98	47	68	89	110	131
(5, 0.5)	23	33	43	53	63	30	44	57	71	84

Table 3 Sample sizes required under a crossover design with $\rho = 0.8$

	Power = 80%					Power = 90%				
	γ					γ				
(K, δ)	0.000	0.001	0.002	0.003	0.004	0.000	0.001	0.002	0.003	0.004
(3, 0.3)	76	83	92	102	116	101	115	132	156	190
(3, 0.4)	43	45	47	50	53	57	61	66	71	77
(3, 0.5)	27	28	29	30	31	36	38	40	42	44
(5, 0.3)	73	80	89	99	113	98	111	128	151	184
(5, 0.4)	41	43	46	48	51	55	59	64	69	75
(5, 0.5)	26	27	28	29	30	35	37	39	40	42

different values of γ. It is seen that the possibility of sample size reduction under a crossover design depends on whether the carryover effect of the QT intervals could be avoided. As a result, it is suggested that a sufficient length of washout period be applied between dosing periods to wear off the residual (or carryover) effect from one dosing period to another. For a fixed sample size, the possibility of power increase by crossover design also depends on parameter γ. Figure 1 shows that crossover design results in power increase when γ is close to 0 but may result in considerable power loss when γ is not close to 0.

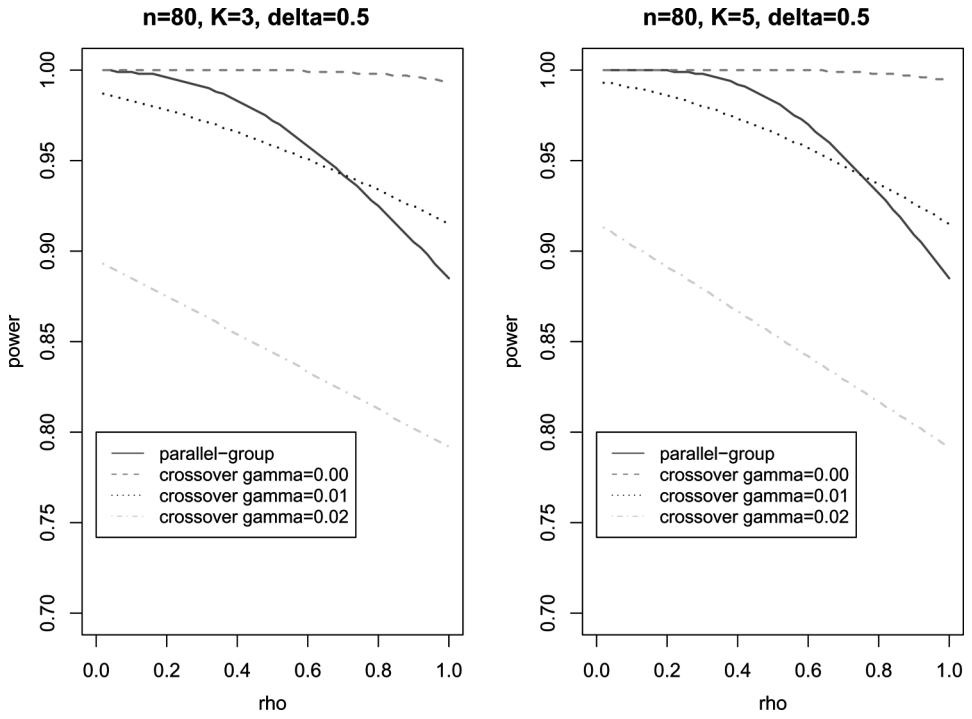

Figure 1 Power comparison under parallel-group and crossover designs.

4. EXTENSION

In the previous section, we consider models without covariates. In practice, additional information—such as some pharmacokinetic (PK) responses, e.g., area under the blood or plasma concentration time curve (AUC), and the maximum concentration (C_{max}), which are known to be correlated to the QT intervals—may be available, for example, in active-controlled QT studies. In this case, models (1) and (2) are necessarily modified to include the PK responses as covariates for a more accurate and reliable assessment of power and sample size calculation (Cheng and Shao, 2007).

4.1. Parallel-Group Design

After the inclusion of the PK response as covariate, model (1) becomes

$$y_{ijk} = \mu_i + \eta x_{ij} + e_{ij} + \epsilon_{ijk},$$

where x_{ij} is the PK response for subject j. The least square estimate of η is given by

$$\hat{\eta} = \frac{\sum_{i=1}^{2}\sum_{j=1}^{n}(\bar{y}_{ij.} - \bar{y}_{i..})(x_{ij} - \bar{x}_{i.})}{\sum_{i=1}^{2}\sum_{j=1}^{n}(x_{ij} - \bar{x}_{i.})^2}.$$

Then $(\bar{y}_{1..} - \bar{y}_{2..}) - \hat{\eta}(\bar{x}_{1.} - \bar{x}_{2.})$ is an unbiased estimator of $\mu_1 - \mu_2$ with variance

$$\left[\frac{(\bar{x}_{1.} - \bar{x}_{2.})^2}{\sum_{ij}(x_{ij} - \bar{x}_{i.})^2/n} + 2\right]\left(\rho + \frac{1-\rho}{K}\right)\frac{\sigma^2}{n},$$

which can be approximated by

$$\left[\frac{(v_1 - v_2)^2}{\tau_1^2 + \tau_2^2} + 2\right]\left(\rho + \frac{1-\rho}{K}\right)\frac{\sigma^2}{n},$$

where $v_i = \lim_{n\to\infty} \bar{x}_{i.}$, and $\tau_i^2 = \lim_{n\to\infty} \sum_{j=1}^{n}(x_{ij} - \bar{x}_{i.})^2/n$. Similar to Section 3.1, to achieve the desired power of $1-\beta$ at the α level of significance, the sample size needed per treatment group is given by

$$n = \frac{(z_\alpha + z_\beta)^2}{\delta^2}\left[\frac{(v_1 - v_2)^2}{\tau_1^2 + \tau_2^2} + 2\right]\left(\rho + \frac{1-\rho}{K}\right). \tag{8}$$

In practice, v_i and τ_i^2 are estimated by the corresponding sample mean and sample variance from the pilot data. Note that if there are no covariates or the PK responses are balanced across treatments (i.e., $v_1 = v_2$), then formula (8) reduces to (5).

4.2. Crossover Design

After the PK response is taken into consideration as a covariate, model (2) becomes

$$y_{ijkl} = \mu_i + \eta x_{ijl} + \beta_{il} + e_{ijl} + \epsilon_{ijkl}.$$

Then $(\bar{y}_{1\cdots} - \bar{y}_{2\cdots}) - \hat{\eta}(\bar{x}_{1\cdot\cdot} - \bar{x}_{2\cdot\cdot})$ is an unbiased estimator of $\mu_1 - \mu_2$ with variance

$$\left[\gamma + \left(\frac{(\bar{x}_{1\cdot\cdot} - \bar{x}_{2\cdot\cdot})^2}{\sum_{ijl}(x_{ijl} - \bar{x}_{i\cdot\cdot})^2/n} + 1\right)\left(\rho + \frac{1-\rho}{K}\right)\right]\sigma^2,$$

which can be approximated by

$$\left[\gamma + \left(\frac{(v_1 - v_2)^2}{\tau_1^2 + \tau_2^2} + 1\right)\left(\rho + \frac{1-\rho}{K}\right)\right]\sigma^2,$$

where $v_i = \lim_{n\to\infty} \bar{x}_{i\cdot\cdot}$, and $\tau_i^2 = \lim_{n\to\infty} \sum_{jl}(x_{ijl} - \bar{x}_{i\cdot\cdot})^2/n$. To achieve the desired power of $1-\beta$ at the α level of significance, the sample size required per treatment group is

$$n = \frac{(z_\alpha + z_\beta)^2}{\delta^2 - \gamma(z_\alpha + z_\beta)^2}\left[\frac{(v_1 - v_2)^2}{\tau_1^2 + \tau_2^2} + 1\right]\left(\rho + \frac{1-\rho}{K}\right). \tag{9}$$

When there are no covariates or PK responses that satisfy $v_1 = v_2$, formula (9) reduces to (7).

Formulas (8) and (9) indicate that under either a parallel-group or a crossover design, a larger sample size is required to achieve the same power if the covariate information is to be incorporated.

5. ALLOCATION OPTIMIZATION

For optimization of the allocation of n (the number of subjects) and K (the number of recording replicates) in routine QT studies with recording replicates, we may consider two approaches, namely, the fixed-power approach and the fixed-budget approach. The fixed-power approach is to find optimal allocation of n and K for achieving a desired (fixed) power in the way that the total budget is minimized. For the fixed-budget approach, the purpose is to find optimal allocation of n and K for achieving maximum power.

In this section, for simplicity, we will describe only the solution under a parallel-group design. The results under a crossover design can be similarly obtained. Let C_1 be the cost for recruiting a subject and C_2 be the associated cost for each QT recording replicate. Finding n and K for achieving a desired (fixed) power of $1-\beta$ under the minimal budget is equivalent to minimizing $C = nC_1 + nKC_2$ under the constraint of $2(z_\alpha + z_\beta)^2(\rho K + 1 - \rho) - nK\delta^2 = 0$. Under the given constraint, the total cost can be expressed as a function of K

$$C(K) = \frac{2(z_\alpha + z_\beta)^2}{\delta^2}\left(\rho C_2 K + \frac{(1-\rho)C_1}{K} + \rho C_1 + (1-\rho)C_2\right),$$

which attains its minimum at $K = \left[\sqrt{\frac{C_1(1-\rho)}{C_2\rho}}\right] + 1$, where function $[t]$ denotes the integer part of t. In practice, we may consider choosing the K value among $K = 1, 3,$ and 5 that results in the smallest C.

When the total budget is fixed, say, $nC_1 + nKC_2 = C_0$, where C_0 is a known constant, the power function (4) becomes a function of K only

$$H(K) = \Phi\left(-z_\alpha + \frac{\delta}{\sqrt{\frac{2(C_1+C_2K)}{C_0}\left(\rho + \frac{1-\rho}{K}\right)}}\right),$$

whose maximal value also occurs at $K = \left[\sqrt{\frac{C_1(1-\rho)}{C_2\rho}}\right] + 1$.

Note that for any fixed ρ, both the fixed-power approach (for achieving a desired power but minimizing the total budget) and the fixed-budget approach (for achieving the minimal power under a fixed total budget) result in the same optimal choice of K (the number of replicates), which is given by $K = \left[\sqrt{\frac{C_1(1-\rho)}{C_2\rho}}\right] + 1$.

6. CONCLUDING REMARKS

Under a parallel-group design, the possibility that the sample size can be reduced depends on the parameter ρ, the correlation between the QT recording replicates. As indicated earlier, when ρ closes to 0, these recording repeats can be viewed as (almost) independent replicates. As a result, $n \approx n_{\text{old}}/K$. When ρ is close to 1, we have $n \approx n_{\text{old}}$. Thus, there is not much gain for considering recording replicates in the study. On the other hand, assuming that all other parameters remain the same, the possibility of further reducing the sample size by a crossover design depends on the parameter γ, which is a measure of the magnitude of the relative period effect.

When analyzing QT intervals with recording replicates, we may consider change from baseline. It is, however, not clear which baseline should be used when there are also recording replicates at baseline. Strieter et al. (2003) proposed the use of the so-called time-matched change from baseline, which is defined as measurement at a time point on the postbaseline day minus measurement at same time point on the baseline. The statistical properties of this approach, however, are not clear. In practice, it may be of interest to investigate relative merits and disadvantages among the approaches using (1) the most recent recording replicates, (2) the mean recording replicates, or (3) time-matched recording replicates as the baseline. This requires further research.

REFERENCES

Cheng, B., Shao, J. (2007). Exact tests for negligible interaction in two-way linear models. *Statistica Sinica* 17:1441–1455.

Chow, S. C., Shao, J., Wang, H. (2003). *Sample Size Calculation in Clinical Research*. New York: Marcel Dekker.

CPMP (1997). Points to consider: The assessment of the potential for QT interval prolongation by non-cardiovascular products. Available at: www.coresearch.biz/regulations/cpmp.pdf.

FDA/TPD (2003). Preliminary concept paper: The clinical evaluation of QT/QTc interval prolongation and proarrythmic potential for non-arrythmic drug products. Released on November 15, 2002. Revised on February 6, 2003.

ICH (2005). ICH E14 Guidance on the Clinical Evaluation of QT/QTc Interval Prolongation and Proarrythmic Potential for Non-Antiarrythmic Drugs. Geneva, Switzerland: International Conference on Harmonisation, May 2005.

Longford, N. T. (1993). *Random Coefficient Models*. New York: Oxford University Press.

Malik, M., Camm, A. J. (2001). Evaluation of drug-induced QT interval prolongation. *Drug Safety* 24:323–351.

Moss, A. J. (1993). Measurement of the QT interval and the risk associated with QT interval prolongation. *American Journal of Cardiology* 72:23B–25B.

PhRMA (2003). Investigating drug-induced QT and QTc prolongation in the clinic: Statistical design and analysis considerations. Report from the Pharmaceutical Research and Manufacturers of America QT Statistics Expert Team, August 14, 2003.

Strieter, D., Wu, W., Agin, M. (2003). Assessing the effects of replicate ECGs on QT variability in healthy subjects. Presented at Midwest Biopharmaceutical Workshop, May 21, 2003.

Temple, R. (2003). Overview of the concept paper, history of the QT/TdP concern; regulatory implications of QT prolongation. Presentations at Drug Information Agency/FDA Workshop. Available at: www.diahome.org.