

## On Sample Size Determination of Bioequivalence Trials<sup>1)</sup>

Sang-Gue Park<sup>2)</sup>

### Abstract

Sample size determination plays an important role in designing a bioequivalence trial. Formulae of sample sizes based on Schuirmann's two one-sided tests procedures are given for bioequivalence studies with the 2x2 crossover design and two-sample parallel design. A practical discussion for the relationship among these formulae is given.

**Keywords** : Bioequivalence, Crossover Design, Inter-Variability, Intra-Variability, Parallel Design, Power, Sample Size

### 1. 서론

생물학적 동등성(이하 생동성) 시험에 대한 정의는 식품의약품안전청고시로 관리되는 생물학적 동등성 시험기준 제2조에 명시되어 있다. 고시에 따르면 생동성 입증을 위한 생체내 시험으로서 주성분이 전신순환혈에 흡수되어 약효를 나타내는 의약품에 대하여 동일 주성분을 함유한 동일 투여경로의 두 제제가 생체이용률에 있어서 통계학적으로 동등하다는 것을 입증하기 위해 실시하는 시험을 말한다.

가장 일반적인 생동성 입증 방법은 오리지널 제제와 시험제제의 약물동태학적 평가 파라미터를 비교하는 약물동태학적 시험 방법이며, 이 외에 약물동력학적 시험방법, 비교임상 시험방법, 생체의 시험방법이 있다. 국내의 경우는 현행 생동성 시험기준 제2조의 정의에 따라 약물동태학적 시험방법으로 수행되고 있다.

제제간의 생동성 시험을 설계할 때 가장 중요하게 결정되어지는 사항 중 하나가 각 군당 피험자의 크기를 결정하는 것이다. 제제간의 동등성 여부를 판단하기 위해 이루어지는 생동성 시험에서 시험 결과에 영향을 줄 수 있는 많은 변수들이 있고 비교적 소수의 피험자를 대상으로 이루어지는 시험임을 감안한다면 적절한 검정력을 확보할 수 있는 시험 크기를 결정하는 것은 매우 중요한 일이다. 일반적으로 오리지널 제제

---

1) This research is supported in 2006 Chung-Ang Research Grant

2) Professor, Department of Statistics, Chung-Ang University, Seoul, 156-756, Korea  
E-mail : spark@cau.ac.kr

와 시험 제제를 비교 평가하는 생동성 시험을 설계할 때 두 제제 간에 실제 차이가 존재하지 않을 경우 두 제제가 동등한 제제로 평가되는 확률인 검정력이 최소한 80%를 넘도록 피험자의 크기를 결정하는 것으로 사용되어 오고 있다.

KFDA에 의하면 제제간의 동등성을 비교하는데 사용하는 통계적 설계는 교차설계법이나 평행설계법을 사용하고 있는데 제제의 반감기가 길지 않을 경우에는 일반적으로 2x2 교차설계법을 권하고 있다.

제제간의 동등성을 입증 하는 방법은 크게 두 가지 방법이 있는데 Schuirmann(1987)이 제안한 5% 유의수준에서 생체이용률의 평균 차이를 two one-sided test(TOST)로 수행하거나 90% 신뢰구간을 계산하여 미리 설정된 생동성 관리 한계인 ( $\log(0.80)$ ,  $\log(1.25)$ ) 안에 포함되면 두 제제의 생동성을 인정하도록 하고 있다. 이 두 가지 방법은 통계적으로 동일한 결과를 제시한다.

본 연구 논문에서는 생동성 시험에서 사용될 수 있는 두 가지 설계 방법에 대한 표본 설계 방법을 논의하고자 한다. 생동성 시험을 설계하는데 가장 기본적인 적절한 검정력을 확보할 수 있는 피험자수를 계산하는 연구는 Phillips(1990), Liu & Chow(1992) 등에 의해서 시작되었고, 최근 대부분의 국가 규정이 로그변환된 생체이용률을 사용하는 것으로 바뀌면서 Chow & Wang(2001), Blood(2002) 등에 의해서 다시 연구되었다. 그 중 Chow & Wang의 연구가 가장 종합적으로 생동성 실험에서 피험자 수를 계산하는 통계적 방법을 제안하고 있지만 그 논문에 전개 과정과 편집에 오류가 많아 생동성 실험을 설계하는 생명과학자들에게 적절한 피험자수 계산에 매우 어려움을 주고 있다. 또한 생동성 실험을 설계하는 생명과학자들에게 생동성 실험 설계인 교차설계법도 제제의 반감기가 길게 예상되는 경우에 교차설계법을 대신하는 하나의 선택으로 활용될 수 있는데 이 때 피험자 수 계산에 대한 논의가 역시 필요한 실정이다.

2장에서는 일반적으로 사용되는 생동성 시험 설계인 2x2 교차설계법이 사용되었을 경우 로그변환된 생체이용률 자료에 대한 피험자수 계산 공식을 유도하고, 3장에서는 반감기가 긴 제제에 사용될 수 있는 평행설계법이 사용되었을 경우 역시 로그변환된 생체이용률 자료에 대한 피험자수 계산 공식을 유도한 후, 4장에서는 피험자수의 크기 관점에서 교차설계법과 평행설계법에 관련된 논의를 하고자 한다.

## 2. 2x2 교차설계법에 의한 피험자수 공식 유도

2x2 교차설계법 하에서 실시된 생동성 시험에서  $i$ -번째 순서의  $j$ -번째 피험자가 복용한  $k$ -번째 제제의 생체이용률 자료를  $Y_{ijk}$  ( $i=1,2; j=1,2,\dots,n_i; k=T,R$ )라 하자. 편의상 첫 번째 순서는  $TR$  순서로 제제를 복용한 집단을 의미하고, 두 번째 순서는  $RT$  순서로 제제를 복용한 집단을 의미한다. 이 때 충분한 휴약기간(washout period)이 있어 잔존효과(carryover effect)가 존재 하지 않음은 제제간의 생동성을 입증하는 통계적 추론에 필수적이다.

두 제제간의 생동성을 판단하기 위한 가설을 다음과 같이 설정하자.

$$H_0: \frac{\mu_T}{\mu_R} \leq \Delta_1, \frac{\mu_T}{\mu_R} \geq \Delta_2, H_1: \Delta_1 < \frac{\mu_T}{\mu_R} < \Delta_2. \quad (2.1)$$

여기서  $\mu_R$ 과  $\mu_T$ 는 오리지널 제제와 시험 제제의 평균치를,  $\Delta_1$ 과  $\Delta_2$ 는 KFDA에서 미리 설정된 관리 기준값인 0.8과 1.25를 의미한다.

KFDA를 포함하여 국제적으로 두 제제간의 생동성을 평가하기 위해 실험에서 얻어진 생체이용률을 로그 변환해서 사용하도록 추천하고 있다. 생체이용률 자료에 로그를 취하는 이유는 크게 비율에 관한 추론보다는 차이에 관한 추론이 더 타당하다는 임상학적 이유, 각 중 모수들의 설명이 더 적절하다는 약물동태학적 이유, 마지막으로 경험적으로 생체이용률 자료가 로그 정규분포를 많이 따른다는 통계학적 이유 때문이다(Rani & Pargai, 2004).

로그 변환된 생체이용률 반응치  $Y_{ijk}^*$ 의 통계적 모형은 다음과 같이 표현할 수 있다.

$$Y_{ijk}^* = \mu_k^* + \gamma_{ik} + \epsilon_{ijk}^* \quad (2.2)$$

여기서  $\mu_R^*$ 와  $\mu_T^*$ 는 로그변환된 오리지널 제제와 시험 제제의 평균이고,  $\gamma_{ik}$ 는 순서 효과, 기간효과 그리고 잔류효과를 의미하는 모수, 확률벡터  $\epsilon_{ij} = (\epsilon_{ijT}, \epsilon_{ijR})$ 의 분산 공분산 행렬은 제제 분산(intra-variability)을 의미하는  $\sigma_e^2$ , 피험자 분산(inter-variability)을 의미하는  $\sigma_s^2$ 을 포함하여 다음과 같이 표현할 수 있다.

$$\begin{pmatrix} \sigma_e^2 + \sigma_s^2 & \sigma_s^2 \\ \sigma_s^2 & \sigma_e^2 + \sigma_s^2 \end{pmatrix}.$$

모형 (2.2)를 이용하면 가설 (2.1)은 다음과 같이 표현할 수 있다.

$$\begin{aligned} H_0 &: \mu_T^* - \mu_R^* \leq \log(\Delta_1), \mu_T^* - \mu_R^* \geq \log(\Delta_2), \\ H_1 &: \log(\Delta_1) < \mu_T^* - \mu_R^* < \log(\Delta_2) \end{aligned} \quad (2.3)$$

가설 (2.3)은 다음과 같이 두 개의 가설로 다시 나누어 표현할 수 있다.

$$\begin{aligned} H_{01} &: \mu_T^* - \mu_R^* \leq -\Delta^*, \quad H_{11} : \mu_T^* - \mu_R^* > -\Delta^*, \\ H_{02} &: \mu_T^* - \mu_R^* \geq \Delta^*, \quad H_{12} : \mu_T^* - \mu_R^* < \Delta^*, \end{aligned}$$

여기서  $\Delta^* = \log(1.25)$ 이다.

제제간의 생동성을 판단하기 위해 가설 (2.3)에 Schuirmann(1987)의 검정 TOST(two one-sided test)를 적용할 수 있다. 즉, 가설  $H_{01}$ 과  $H_{11}$ 을 검정하는 검정 통계량  $S_L$ 을 이용한 판단기준

$$S_L = \frac{\overline{Y}_T^* - \overline{Y}_R^* + \Delta^*}{\hat{\sigma}_d \sqrt{1/n_1 + 1/n_2}} > t(\alpha, n_1 + n_2 - 2)$$

과 가설  $H_{02}$ 과  $H_{12}$ 을 검정하는 검정통계량  $S_U$ 을 이용한 판단기준

$$S_U = \frac{\overline{Y}_T^* - \overline{Y}_R^* - \Delta^*}{\widehat{\sigma}_d^* \sqrt{1/n_1 + 1/n_2}} < -t(\alpha, n_1 + n_2 - 2),$$

이 동시에  $H_{11}$  과  $H_{12}$  를 채택한다면 유의수준  $\alpha$ 에서 제제간의 생동성을 판단할 수 있다. 여기서  $t(\alpha, n_1 + n_2 - 2)$ 는 자유도가  $n_1 + n_2 - 2$ 인  $t$ -분포표에서 상위  $\alpha$ 에 해당되는 백분위수이고

$$\begin{aligned} \overline{Y}_T^* &= \frac{1}{2}(\overline{Y}_{1.T}^* + \overline{Y}_{2.T}^*), \quad \overline{Y}_R^* = \frac{1}{2}(\overline{Y}_{1.R}^* + \overline{Y}_{2.R}^*), \\ d_{1j} &= \frac{1}{2}(Y_{1jT}^* - Y_{1jR}^*), \quad d_{2j} = \frac{1}{2}(Y_{2jT}^* - Y_{2jR}^*), \quad \overline{d}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} d_{ij}, \\ \widehat{\sigma}_d^* &= \sqrt{\frac{1}{n_1 + n_2 - 2} \sum_{i=1}^2 \sum_{j=1}^{n_i} (d_{ij} - \overline{d}_i)^2} \end{aligned}$$

이다.

Schuirmann이 제안한 유의수준 5%의 TOST는 현재 FDA에서 사용하고 있는 기준인 90% 신뢰구간을 만들어 미리 정해진 관리 한계선과 비교하는 방법과 동일한 방법이다. 대립가설의 한 값  $\theta^* = \mu_T^* - \mu_R^* \in (-\Delta^*, \Delta^*)$ 에서 Schuirmann의 TOST 검정의 검정력은 다음과 같이 계산 될 수 있다.

$$\begin{aligned} 1 - \beta &= P \left\{ t(\alpha, n_1 + n_2 - 2) - \frac{\Delta^* + \theta^*}{\widehat{\sigma}_d^* \sqrt{1/n_1 + 1/n_2}} < \frac{\overline{Y}_T^* - \overline{Y}_R^* - \theta^*}{\widehat{\sigma}_d^* \sqrt{1/n_1 + 1/n_2}} \right. \\ &\quad \left. < \frac{\Delta^* - \theta^*}{\widehat{\sigma}_d^* \sqrt{1/n_1 + 1/n_2}} - t(\alpha, n_1 + n_2 - 2) \right\}. \quad (2.4) \end{aligned}$$

식 (2.3)의 직관적인 의미를 보기위해 군당 피험자의 크기  $n_1 = n_2 = n$ 이라 하고 만약 두 제제의 생체이용률의 차이  $\theta^*$ 가 없다고 하면 식(2.4)은 다음과 같이 표현 될 수 있다.

$$1 - \beta = P \left\{ |T_{(2n-2)}| < \frac{\Delta^*}{\sqrt{\frac{2}{n} \widehat{\sigma}_d^*}} - t(\alpha, 2n - 2) \right\}, \quad (2.5)$$

여기서  $T_{(2n-2)}$ 는 자유도가  $2n - 2$ 인  $t$ -분포를 따르는 확률변수이다. 식 (2.5)로부터 유의수준  $\alpha$ 에서 검정력  $1 - \beta$ 를 확보할 수 있는 피험자의 크기  $n$ 을 계산할 수 있다. 즉,

$$n \geq \{t(\alpha, 2n - 2) + t(\beta/2, 2n - 2)\}^2 \left( \frac{\widehat{\sigma}_e^*}{\Delta^*} \right)^2, \quad \theta^* = 0, \quad (2.6)$$

여기서  $\widehat{\sigma}_e^* = \sqrt{2} \widehat{\sigma}_d^*$  이고,  $\widehat{\sigma}_e^{*2}$ 는 생동성 시험에서 흔히 계산되는 분산분석표의 제제 분산(intra-subject variability)으로 추정될 수 있다.

두 제제간의 생동성 시험을 설계하는데 두 제제간의 생체이용률의 값이 비슷할 것으로 예측은 되지만  $\theta^*$ 가 0이 아닌 다른 값이라면 식 (2.4)는 다음과 같이 표현된다.

$$1 - \beta = P \left\{ T_{(2n-2)} < \frac{\Delta^* - \theta^*}{\sqrt{\frac{2}{n} \widehat{\sigma}_d^*}} - t(\alpha, 2n-2) \right\} - \left\{ T_{(2n-2)} < t(\alpha, 2n-2) - \frac{\Delta^* + \theta^*}{\sqrt{\frac{2}{n} \widehat{\sigma}_d^*}} \right\} \quad (2.7)$$

Chow & Wang(2001)은 이 경우에  $\sqrt{n} \theta^*$ 가 비교적 크다는 가정을 하고 식 (2.7) 원편의 두 항 중 두 번째 항이 0으로 근사한다는 가정 하에 근사적인 공식을 다음과 같이 제안하였다.

$$n \geq \{t(\alpha, 2n-2) + t(\beta, 2n-2)\}^2 \left( \frac{\widehat{\sigma}_e^*}{\Delta^* - |\theta^*|} \right)^2, \quad \theta^* \neq 0. \quad (2.8)$$

하지만 식 (2.8)의 정확성에 관해 Blood(2002)는  $\theta^*$ 가 10%나 15% 수준에서도 정확하지 못하다고 평가하였다. 따라서  $\sqrt{n} \theta^*$ 이 크지 않을 때는 피험자수를 계산하는 데는 (2.7)을 사용하여 계산하는 것이 좋다고 생각된다.

식 (2.7)과 (2.8)은 주어진 검정력을 만족하는 최소의  $n$ 을 찾게 되는데  $t$ -분포의 자유도가 역시 피험자수에 의존하므로 반복적인 프로그램 계산식에 의해서 계산될 수 있다. 일부 빈번하게 나타날 수 있는 모수에 대한 요약은 [표 2.1]에 정리되어 있다.

[표 2.1] 군당 피험자수 계산

$\widehat{\sigma}_e^* \setminus \theta^*$	$1 - \beta = 0.80$				$1 - \beta = 0.90$			
	0	5	10	15	0	5	10	15
10	3	4	5	13	4	4	7	17
12	4	5	7	18	5	6	10	24
14	5	6	9	24	6	7	12	33
16	6	7	12	31	7	9	16	42
18	7	8	15	39	9	11	20	53
20	8	10	18	47	12	13	24	65
22	10	12	21	57	14	15	29	79
24	11	14	25	68	15	18	34	94
26	13	16	29	79	16	21	40	110
28	15	18	33	92	19	24	46	127
30	17	20	38	106	21	27	52	145
32	19	23	43	120	24	31	59	165
34	21	26	48	135	27	35	67	186
36	24	29	54	151	30	39	75	209
38	26	32	60	168	33	43	83	232
40	29	35	67	186	36	47	92	257

지금은 대다수의 국가의 FDA에서 생동성 시험에서 얻어진 생체이용률을 로그변환해서 사용하지만 그 전에는 원자료로 생동성 시험 분석을 수행하였고, 일부 국가의

경우 지금까지도 원자료를 이용해서 생동성을 분석하는 경우가 있다. 생동성 실험을 설계할 때는 과거 유사 생동성 실험에서 얻어진 분석 결과에서 평균이나 분산을 추측하는 경우가 많기 때문에 원자료로 분석된 결과에서도 생동성 실험의 피험자 수를 계산할 수 있어야 한다.

원자료에 대한 생동성 검정 가설 (2.1)에 위의 논의를 그대로 적용시키려면 다음과 같은 간단한 조작으로 근사적인 검정을 완성할 수 있다.

$$\begin{aligned} H_{01}: \mu_T - \mu_R &\leq -\Delta\mu_R, & H_{11}: \mu_T - \mu_R &> -\Delta\mu_R, \\ H_{02}: \mu_T - \mu_R &\geq \Delta\mu_R, & H_{12}: \mu_T - \mu_R &< \Delta\mu_R. \end{aligned}$$

위 가설에서  $\Delta\mu_R$ 을 주어진 것으로 가정한다면 지금까지 논의한 표본 결정 공식을 사용할 수 있다. 즉,  $\Delta$ 를 0.2로,  $\mu_R$ 을 이전 자료에서 얻어진 오리지널 제제의 평균 생체이용률로 계산한다. 가설 자체가 근사적인 값을 이용하기 때문에 Chow와 Wang(2001)의 근사적인 방법으로도 충분한 계산을 할 수 있을 것이다. 피험자 수를 계산하면 다음과 같다.

$$n \geq \{t(\alpha, 2n-2) + t(\beta/2, 2n-2)\}^2 \left( \frac{CV}{0.2} \right)^2, \quad \theta' = 0, \quad (2.9)$$

$$n \geq \{t(\alpha, 2n-2) + t(\beta, 2n-2)\}^2 \left( \frac{CV}{0.2 - |\theta'|} \right)^2, \quad \theta' \neq 0. \quad (2.10)$$

여기서  $\theta' = (\mu_T - \mu_R)/\mu_R$ 이고,  $CV = \sigma_e/\mu_R$ 이다.

### 3. 평행설계법에 의한 피험자수 공식 유도

일반적으로 생동성 시험을 2x2 교차설계법을 이용하여 입증할 때 가장 중요한 사항 중의 하나가 제제의 잔존효과가 존재하지 않도록 실험하는 것이다. 일반적으로 실험에서 휴약기간을 설정할 때 제제의 4배 혹은 5배 정도의 반감기(half life)로 휴약기간을 설정한다. 하지만 이렇게 설정된 반감기가 매우 길 때 경우에 따라서는 교차설계법보다 평행설계법이 더 실험의 정도를 가져올 수 있다. 하지만 피험자 분산( $\sigma_s^2$ )과 제제 분산( $\sigma_e^2$ )을 분리할 수 있는 교차설계법과 달리 평행설계법에서는 이를 분리할 수 없어 총변동( $\sigma_s^2 + \sigma_e^2$ )으로 피험자수를 계산하여야 하기 때문에 커진 분산 때문에 더 많은 피험자가 필요하게 된다.

평행설계법 하에서 실시된 생동성 시험은 일반적으로 시행되는 두 집단 문제로 보면 된다.  $Y_{ij}$  ( $i = T, R; j = 1, 2, \dots, n_i$ )는  $i$ -번째 제제를 복용한  $j$ -번째 피험자의 생체이용률 자료를 의미한다.

로그 변환된 생체이용률 반응치  $Y_{ij}^*$ 의 통계적 모형은 다음과 같이 표현할 수 있다.

$$Y_{ij}^* = \mu_i + \epsilon_{ij}^*, \quad (3.1)$$

여기서  $\mu_R^*$ 와  $\mu_T^*$ 는 오리지널 제제와 시험 제제의 평균이고,  $\epsilon_{ij}^*$ 는 제제분산과 피험자분산이 포함된 변동을 의미하는 확률변수이다.

두 제제간의 생동성을 판단하기 위해 (2.1), (2.2)와 같이 가설을 설정할 수 있다. 일

반적인 두 집단 문제에서  $t$ -검정 통계량을 사용하면 가설  $H_{01}$ 과  $H_{11}$ 을 검정하는 검정통계량  $T_L$ 을 이용한 판단기준

$$T_L = \frac{\overline{Y}_T^* - \overline{Y}_R^* + \Delta^*}{\widehat{\sigma}^* \sqrt{1/n_1 + 1/n_2}} > t(\alpha, n_1 + n_2 - 2)$$

과 가설  $H_{02}$ 과  $H_{12}$ 을 검정하는 검정통계량  $T_U$ 을 이용한 판단기준

$$T_U = \frac{\overline{Y}_T^* - \overline{Y}_R^* - \Delta^*}{\widehat{\sigma}^* \sqrt{1/n_1 + 1/n_2}} < -t(\alpha, n_1 + n_2 - 2),$$

이 동시에  $H_{11}$ 과  $H_{12}$ 를 채택한다면 유의수준  $\alpha$ 에서 제제간의 생동성을 판단할 수 있다. 여기서

$$\overline{Y}_T^* = \frac{1}{n_1} \sum_{j=1}^{n_1} Y_{ij}^*, \quad \overline{Y}_R^* = \frac{1}{n_2} \sum_{j=1}^{n_2} Y_{ij}^*,$$

$$\widehat{\sigma}^* = \sqrt{\frac{1}{n_1 + n_2 - 2} \left\{ \sum_{j=1}^{n_1} (Y_{Tj}^* - \overline{Y}_T^*)^2 + \sum_{j=1}^{n_2} (Y_{Rj}^* - \overline{Y}_R^*)^2 \right\}}$$

이다. 2장과 비슷하게 계산해보면 군당 피험자의 크기  $n_1 = n_2 = n$ 이라 하고 만약 두 제제의 생체이용률의 차이  $\theta^*$ 가 없다고 하면  $T_L$ 과  $T_U$ 에 기초한 검정의 검정력 함수는 다음과 같이 표현 될 수 있다.

$$1 - \beta = P \left\{ |T_{(2n-2)}| < \frac{\Delta^*}{\sqrt{\frac{2}{n} \widehat{\sigma}^*}} - t(\alpha, 2n-2) \right\}, \quad (3.2)$$

식 (3.2)로부터 유의수준  $\alpha$ 에서 검정력  $1 - \beta$ 를 확보할 수 있는 피험자 크기  $n$ 을 계산할 수 있다. 즉,

$$n \geq \{t(\alpha, 2n-2) + t(\beta/2, 2n-2)\}^2 \left( \frac{\widehat{\sigma}^*}{\Delta^*} \right)^2, \quad \theta^* = 0, \quad (3.3)$$

두 제제간의 차이  $\theta^*$ 가 0이 아니면 검정력 함수는 다음과 같이 표현된다.

$$1 - \beta = P \left\{ T_{(2n-2)} < \frac{\Delta^* - \theta^*}{\sqrt{\frac{2}{n} \widehat{\sigma}^*}} - t(\alpha, 2n-2) \right\} - \left\{ T_{(2n-2)} < t(\alpha, 2n-2) - \frac{\Delta^* + \theta^*}{\sqrt{\frac{2}{n} \widehat{\sigma}^*}} \right\} \quad (3.4)$$

식 (3.3)과 (3.4)는 주어진 검정력을 만족하는 최소의  $n$ 을 찾게 되는데  $t$ -분포의 자유도가 역시 피험자수에 의존하므로 반복적인 프로그램 계산식에 의해서 계산될 수 있다.

원자료로 분석된 결과에서도 생동성 실험의 피험자 수를 근사적으로 계산하면 다음과 같다.

$$n \geq \{t(\alpha, 2n-2) + t(\beta/2, 2n-2)\}^2 \left( \frac{CV}{0.2} \right)^2, \quad \theta' = 0 \quad (2.9)$$

$$n \geq \{t(\alpha, 2n-2) + t(\beta, 2n-2)\}^2 2 \left( \frac{CV}{0.2 - |\theta'|} \right)^2, \theta' \neq 0 \quad (2.10)$$

여기서  $CV = \sigma/\mu_R$ 이다.

#### 4. Concluding Remarks

대부분의 국가에서 생동성 실험을 수행할 때 기본으로 채택하고 있는 설계 방법인 2x2 교차설계법은 피험자간 변동과 제제간 변동을 분리할 수 있어 적은 수의 피험자로 제제간의 동등성을 확보할 수 있는 좋은 실험 방법이지만 실험하는 과정에서 피험자의 확률화나 순서효과 혹은 잔존효과가 존재하지 않게 해야 하는 등 기술적으로 어려운 점이 많은 방법이다.

이에 반해 평행설계법은 피험자 변동과 제제 변동을 분리할 수 없기 때문에 교차설계법보다는 많은 피험자로 실험을 수행해야 하지만 잔존효과 등과 같은 어려운 모수를 고려하지 않아도 되기 때문에 실험이 간단할 수 있다. 따라서 평행설계법으로 생동성을 평가할 때는 교차설계법보다는 늘어난 피험자 변동만큼 더 많은 수의 피험자를 고려해야 한다는 것을 시험 책임자는 알아야 한다.

이렇게 계산된 피험자 수는 적절한 검정력을 확보하고 있기 때문에 적절하게 실험된 경우에 제제간의 생동성을 이론적으로 확보하는데 큰 문제가 없겠지만 생동성 시험 자체가 많은 변동을 가지고 있기 때문에 실험 수행에 신경을 써야 한다. 특히 생체이용률을 측정하는 과정에서 많은 측정오차가 발생되고 있기 때문에 이렇게 이론적으로 계산되었다 하더라도 피험자나 제제변동이 이론치보다 크게 얻어져 생동성을 확보하는데 실패하는 경우도 많기 때문이다.

본 연구논문은 생동성 시험을 설계하는데 필요한 통계적 피험자수를 계산하는 방법을 제시하였다. 이와 관련된 논문으로 Chow & Wang(2001)가 있으나 그들 논문에서 제시한 근사적인 방법이 보다 더 정확한 값을 제시하고 있으며, 그들 논문에서 나타나고 있는 전개과정에서 많은 편집 오류를 수정하여 제시하고 있기 때문에 실제 실험 설계자들에게 도움을 줄 수 있으리라 생각된다. 최근 생동성 평가의 새로운 기준으로 자리 잡고 있는 모집단 생동성 평가, 개인 생동성 평가 등에 맞는 피험자 수 계산 공식을 유도하는 분야는 아직 본격적인 연구가 진행되고 있지 않아 시급히 연구되어야 하는 문제라고 생각된다.

#### References

1. Blood, P. (2002). Sample size calculation in bioequivalence trials. *Journal of Pharmacokinetics and Pharmacogenetics*, 29, 1, 95-97.
2. Chow, S.C. and Wang, H. (2001). On sample size calculation in bioequivalence trials. *Journal of Pharmacokinetics and Pharmacodynamics*, 28, 2, 155-169.
3. Liu, J.P. and Chow, S.C. (1992). Sample size determination for the two one-sided tests procedure in bioequivalence. *Journal of Pharmacokinetics and Biopharmaceutics*, 20, 1, 101-104.
4. Phillips, K.F. (1990). Power of the two one-sided tests procedure in



- bioequivalence. *Journal of Pharmacokinetics and Biopharmaceutics*, 18, 137-144.
5. Rani, S. and Pargal, A. (2004). Bioequivalence : An overview of statistical concepts. *Indian Journal of Pharmacology*, 36, 4, 209-216.
  6. Schuirmann, D.J. (1987). A comparison of two one-sided tests procedures and the power approach for assessing the bioequivalence trials. *Journal of Pharmacokinetics and Biopharmaceutics*, 15, 657-680.

[ 2007년 3월 접수, 2007년 5월 채택 ]