

## 2×3 이중 설계에서 생물학적 동등성 평가

우화형<sup>1</sup> · 정규진<sup>2</sup> · 박상규<sup>3</sup>

<sup>1</sup>중앙대학교 응용통계학과 · <sup>2</sup>한남대학교 비즈니스통계학과

접수 2017년 6월 8일, 수정 2017년 7월 3일, 게재확정 2017년 7월 8일

### 요약

두 제제의 생체이용률을 비교하여 동등성을 입증하는 생물학적 동등성 시험은 표준 2×2 교차설계를 원칙으로 하고 있으나, 최근에는 제제의 특성이 고변동성으로 인해 표준 2×2 교차설계를 사용할 경우 지나치게 많은 피험자가 필요하게 되면서, EMA나 MFDS 등에서는 2×2 교차설계를 확장한 2×4, 4×2 또는 4×4 등의 고차원 설계를 권장하고 있다. 2×3 교차설계는 표준 2×2 교차설계에서 1기간을 확장한 설계로 불균형 설계이기는 하지만 2×4 교차설계와 비교해서 경제적, 윤리적 측면에서 장점을 갖고 있는 설계라 할 수 있다. 본 연구에서는 2×3 이중설계를 활용하는 생물학적 동등성 시험에서 설계의 통계적 특성을 고찰해 보고, 생물학적 동등성 평가를 위한 통계적 설계 및 추론 절차를 새롭게 개정된 의약품동등성시험기준을 적용하여 논의한다.

주요용어: 2×3 듀얼 설계, 2×3 추가 대조군 설계, 고변동성 제제, 생동성 시험.

### 1. 서론

오리지널 (original) 의약품과 주성분·함량 및 제형이 동일한 제네릭 (generic) 의약품 (복제 의약품)은 생물학적 동등성 (bioequivalence, 이하 생동성) 시험을 통해 두 제제간의 동등성을 평가받은 후 오리지널 의약품의 대체제로 사용된다. 양질의 제네릭 의약품은 환자의 입장이나 보건 당국의 입장에서 갈수록 높아져가는 의약품 비용을 줄여 가는데 핵심적인 역할을 하고 있다.

식품의약품안전처 (ministry of food and drug safety, MFDS) 의약품동등성시험기준 (제2017-28호, 2017)에 따르면 생동성 시험은 2×2 교차설계를 원칙으로 하며, 두 제제의 생체이용률 (AUCt와 Cmax)을 로그변환한 후 계산된 평균치 차의 90% 신뢰구간이 생동성 허용구간인 (log0.8, log1.25) 이 내이면 제제간의 동등성을 인정한다.

최근 제제의 특성에 기인하는 고변동성 오리지널 제제가 다수 특허 만료가 되어 생동성 시험 대상으로 떠오르게 되고, 이러한 제제들의 생동성 시험을 기존의 2×2 교차설계로 진행하게 될 경우 지나치게 많은 개체가 필요하게 되면서 자연스럽게 2×4, 4×2, 4×4 등의 고차원 설계를 권장하고 있다 (Jeong과 Park, 2011). 또한 MFDS 의약품동등성시험기준에서도 제17조를 새롭게 추가하면서 고차원 교차설계로 생동성 시험을 진행할 경우 Cmax는 확장된 생동성 기준 구간을 활용할 수 있게 해주고 있어, 고차원 교차설계에 대한 제약업계의 연구 요구가 많아지고 있다. 이러한 개정 방향은 2010년 이후 유럽의약품청 (European medicines agency, EMA)을 중심으로 국제적으로도 함께 이루어지고 있다(EMA, 2010).

<sup>1</sup> (06974) 서울특별시 동작구 흑석로 84, 중앙대학교 응용통계학과, 강사.

<sup>2</sup> (34430) 대전광역시 대덕구 한남로 70, 한남대학교 비즈니스통계학과, 교수.

<sup>3</sup> 교신저자 : (06974) 서울특별시 동작구 흑석로 84, 중앙대학교 응용통계학과, 교수.

E-mail: spark@cau.ac.kr

고차원 교차설계 중 가장 많이 활용되고 있는 2×4 교차설계는 2×2 교차설계를 두 번 반복하는 것이다. 2×4 교차설계는 각 개체에게 비교하는 두 제제를 두 번씩 반복해서 투여하므로 각 제제들의 변동성을 측정할 수 있는 장점이 있지만, 4기간에 걸쳐 진행되기 때문에 개체들에게는 피로감을 줄 수 있는 단점이 있다. 이에 기간을 하나 단축시킨 2×3 교차설계가 보다 효율적이고, 또한 윤리적일 수 있다는 논문들이 제안되고 있다(Chow 등, 2002; Noh와 Park, 2013; Park과 Park, 2016).

2×3 교차설계는 시험제제 (T)와 대조제제 (R)가 2순서, 3기간에 걸쳐 교차 투여되는 설계이다. 즉, 2×3 교차설계에서는 한 개체가 3 기간에 걸쳐서 한 가지 제제를 두 번, 다른 한 가지 제제를 한 번 투여 받게 되기 때문에 (TRR, RTT), (TRR, RTR), (TRT, RTT) 등과 같이 3가지 형태로 설계될 수 있다. 여기에서 (TRR, RTT)와 같은 교차설계는 제제의 측면에서 봤을 때 총 6회의 투여에서 시험제제와 대조제제가 각각 3회씩 균형 있게 교차 투여되는 설계로 이를 2×3 이중 설계 (two-sequence dual design)라 한다 (Chow와 Liu, 2008). 그리고 (TRR, RTR)과 같은 설계는 대조제제가 시험제제보다 각 순서에 1회씩 더 투여되는 설계로 2×3 추가 대조군 설계 (extra reference design)라 하며, 추가 대조군 설계와 상반되게 (TRT, RTT)와 같이 시험제제가 대조제제보다 1회씩 더 투여되는 설계는 2×3 추가 시험군 설계 (extra test design)라 할 수 있겠다.

본 연구에서는 2장에서 2×3 이중 교차설계를 활용한 평균 생동성 평가에 대한 통계적 추론을 추가대조군 교차설계와 비교하며 논의하고, 3장에서 최소 군당 개체 크기를 계산하며, 제4장에서는 사례를 통해 2×3 이중 설계를 통한 평균 생동성 평가 결과를, 5장에서 연구결과의 의의를 제시한다.

## 2. 2×3 이중 설계에서 통계적 추론

### 2.1. 2×3 이중 설계에서 통계적 추론

2×3 이중 설계는 제제의 측면에서 봤을 때 시험제제와 대조제제가 투여 횟수에 있어서 균형을 이루고 있으며, 기간에 배치되는 순서에 따라 (TRR, RTT), (TRT, RTR) 또는 (TTR, RRT) 등으로 설계될 수 있다. 어떤 순서로 배치되더라도 제제간의 동등성 평가는 동일하게 이루어지기 때문에 Table 2.1과 같이 (TRR, RTT)의 경우로 평균 생동성을 추론해 본다.

Table 2.1 2×3 dual design

		Period				
		1	2	3	4	5
Sequence	1	T	washout	R	washout	R
	2	R	washout	T	washout	T

충분한 휴약기간을 두어 제제에 대한 잔류효과는 존재하지 않는다는 가정 하에 Table 2.1의 2×3 이중 설계에 대한 통계 모형은 다음과 같다.

$$y_{ijk} = \mu + g_k + S_{ik} + p_j + \pi_{(j,k)} + \varepsilon_{ijk}, \quad (2.1)$$

$$S_{ik} \sim iidN(0, \sigma_S^2), \quad \varepsilon_{ijk} \sim iidN(0, \sigma_e^2),$$

$$i = 1, 2, \dots, n_k, \quad j = 1, 2, 3, \quad k = 1, 2,$$

여기서  $y_{ijk}$  는  $k$ 번째 순서,  $j$ 번째 기간에  $(j, k)$ 번째 제제를 투여한  $i$ 번째 개체에게서 얻은 생체이용률 값을 로그변환한 값이고,  $\mu$ 는 전체 평균,  $g_k$ 는  $k$ 번째 순서효과,  $p_j$ 는  $j$ 번째 기간효과,  $\pi_{(j,k)}$ 는  $(j, k)$ 번째 제제효과이다. 순서효과, 기간효과, 제제효과는 다음의 조건을 만족한다.

$$\sum_{k=1}^2 g_k = 0, \sum_{j=1}^3 p_j = 0, \sum_{k=1}^2 \sum_{j=1}^3 \pi_{(j,k)} = 0, \pi_{(j,k)} = \begin{cases} \pi_T, & (j,k) = (1, 1), (2, 2), (3, 2), \\ \pi_R, & (j,k) = (1, 2), (2, 1), (3, 1). \end{cases} \quad (2.2)$$

$S_{ik}$ 는  $k$ 번째 순서,  $i$ 번째 개체의 변동을 나타내는 확률변수이고,  $\varepsilon_{ijk}$ 는  $k$ 번째 순서,  $j$ 번째 기간,  $i$ 번째 개체 내에서 제제의 변동을 나타내는 확률변수이다.  $S_{ik}$ 의 분산  $\sigma_S^2$ 은 개체에 의한 변동,  $\varepsilon_{ijk}$ 의 분산  $\sigma_e^2$ 은 제제에 의한 변동을 의미하고,  $\sigma_S^2$ 와  $\varepsilon_{ijk}$ 는 서로 독립을 가정한다.

2x3 이중 설계는 한 순서에 시험제제 혹은 대조제제가 2회 투여되므로 제제의 분산 추정이 가능하지만, 설계의 특성을 파악하기 위해 각 제제별 개체 변동성 ( $\sigma_{BT}^2 = \sigma_{BR}^2 = \sigma_S^2$ )과 제제 변동성 ( $\sigma_{WT}^2 = \sigma_{WR}^2 = \sigma_e^2$ )은 동일하다고 가정하였다.

3기간에 걸친 관측값  $y_{ik} = (y_{i1k}, y_{i2k}, y_{i3k})'$  ( $i = 1, 2, \dots, n_k$ )는  $k$ 번째 순서,  $i$ 번째 개체에게서 얻은 생체이용률 값으로 식 (2.1)의 통계 모형에 근거하여 다음과 같이 나타낼 수 있다.

$$y_{i1} \sim N((\mu + g_1 + p_1 + \pi_T, \mu + g_1 + p_2 + \pi_R, \mu + g_1 + p_3 + \pi_R)', \Sigma), \quad (2.3)$$

$$y_{i2} \sim N((\mu + g_2 + p_1 + \pi_R, \mu + g_2 + p_2 + \pi_T, \mu + g_2 + p_3 + \pi_T)', \Sigma), \quad (2.4)$$

여기서 분산-공분산 행렬  $\Sigma$ 는 다음과 같다.

$$\Sigma = \begin{pmatrix} \sigma_e^2 + \sigma_S^2 & \sigma_S^2 & \sigma_S^2 \\ \sigma_S^2 & \sigma_e^2 + \sigma_S^2 & \sigma_S^2 \\ \sigma_S^2 & \sigma_S^2 & \sigma_e^2 + \sigma_S^2 \end{pmatrix}. \quad (2.5)$$

식 (2.1)의 통계 모형에서 각 효과들의 유의성 검정을 위한 분산분석표는 다음과 같이 유도되는데, 우선 총제곱합을 분해하면 다음과 같다.

$$\sum_{k=1}^2 \sum_{j=1}^3 \sum_{i=1}^{n_k} (y_{ijk} - \bar{y}_{...})^2 = 3 \sum_{k=1}^2 \sum_{i=1}^{n_k} (\bar{y}_{i..k} - \bar{y}_{...})^2 + \sum_{k=1}^2 \sum_{j=1}^3 \sum_{i=1}^{n_k} (y_{ijk} - \bar{y}_{i.k})^2,$$

$$SS_{Total} = SS_{Between} + SS_{Within},$$

여기서  $SS_{Total}$ 은 총변동이고, 총변동은 개체간 변동  $SS_{Between}$ 과 개체내 변동  $SS_{Within}$ 으로 분해된다. 여기에서 다시 개체간 변동  $SS_{Between}$ 은 순서효과  $SS_{Sequence}$ , 개체간 잔차  $SS_{Inter}$ 로 분해되고, 개체내 변동  $SS_{Within}$ 은 기간효과  $SS_{Period}$ , 제제효과  $SS_{Drug}$ , 개체내 잔차  $SS_{Intra}$ 로 분해된다.

$$SS_{Between} = SS_{Sequence} + SS_{Inter},$$

$$SS_{Within} = SS_{Period} + SS_{Drug} + SS_{Intra}.$$

각 효과들에 대한 제곱합을 얻기 위한 직교대비를 Table 2.2와 Table 2.3과 같이 정의하고 직교대비를 이용해 각 효과들의 제곱합과 개체간 잔차  $SS_{Inter}$ , 개체내 잔차  $SS_{Intra}$ 를 구하면 다음과 같다.

$$SS_{Sequence} = \frac{n_1 n_2}{3(n_1 + n_2)} (\bar{y}_{.11} + \bar{y}_{.21} + \bar{y}_{.31} - \bar{y}_{.12} - \bar{y}_{.22} - \bar{y}_{.32})^2,$$

$$SS_{Inter} = SS_{Between} - SS_{Sequence},$$

**Table 2.2** Orthogonal contrasts for  $SS_{Sequence}$  and  $SS_{Drug}$

Sequence	$SS_{Sequence}$			$SS_{Drug}$		
	Period			Period		
	1	2	3	1	2	3
1	1	1	1	2	-1	-1
2	-1	-1	-1	-2	1	1

**Table 2.3** Orthogonal contrasts for  $SS_{Period}$

Sequence	$SS_{P1}$			$SS_{P2}$		
	Period			Period		
	1	2	3	1	2	3
1	1	-1	0	1	1	-2
2	1	-1	0	1	1	-2

$$\begin{aligned}
 SS_{Period} &= \frac{n_1 n_2}{2(n_1 + n_2)} (\bar{y}_{.11} - \bar{y}_{.21} + \bar{y}_{.12} - \bar{y}_{.22})^2, \\
 &+ \frac{n_1 n_2}{6(n_1 + n_2)} (\bar{y}_{.11} + \bar{y}_{.21} - 2\bar{y}_{.31} + \bar{y}_{.12} + \bar{y}_{.22} - 2\bar{y}_{.32})^2, \\
 SS_{Drug} &= \frac{n_1 n_2}{6(n_1 + n_2)} (2\bar{y}_{.11} - \bar{y}_{.21} - \bar{y}_{.31} - 2\bar{y}_{.12} + \bar{y}_{.22} + \bar{y}_{.32})^2, \\
 SS_{Intra} &= SS_{Within} - SS_{Period} - SS_{Drug}.
 \end{aligned}$$

Table 2.4는 이렇게 정리된 제곱합을 정리한 분산분석표이다.

**Table 2.4** Analysis of variance table for 2x3 dual design

sources	df	mean SS	E(mean SS)
Between subjects			
Sequence effect	1	$MS_{Sequence}$	$\frac{3n_1 n_2}{(n_1 + n_2)} + \left\{ (g_1 - g_2) - \frac{(\pi_T - \pi_R)}{3} \right\}^2 + 3\sigma_S^2 + \sigma_e^2$
Inter-subjects residuals	$(n_1 + n_2) - 2$	$MS_{Inter}$	$3\sigma_S^2 + \sigma_e^2$
Within subjects			
Period effect	2	$MS_{Period}$	$\frac{n_1 n_2}{(n_1 + n_2)} (p_1 - p_2)^2 + \frac{n_1 n_2}{3(n_1 + n_2)} (p_1 + p_2 - 2p_3)^2 + \sigma_e^2$
Drug effect	1	$MS_{Drug}$	$\frac{8n_1 n_2}{3(n_1 + n_2)} (\pi_T - \pi_R)^2 + \sigma_e^2$
Intra-subjects residuals	$2(n_1 + n_2) - 3$	$MS_{Inter}$	$\sigma_e^2$
Total	$3(n_1 + n_2) - 1$		

Table 2.4처럼 분산분석표를 계산하는 이유는 MFDS 의약품동등성시험기준에서 제제간의 생동성을 평가하기 이전에 유의수준 5%에서 순서효과에 대한 검정을 의무화하고 있기 때문이기도 하지만, 순서효과의 평균제곱합의 기댓값을 보게 되면 순서효과와 제제효과가 교락되어 있음을 볼 수 있기 때문이기도 하다. 즉, 순서효과가 존재하지 않게 되어야 제제효과에 대한 추론이 이어질 수 있어, MFDS에서는 분산분석표에서 순서효과의 검정을 요구하고 있다. 순서효과가 존재하지 않는다는 전제하에 두 제제의

생동성 평가는 두 제제의 로그변환한 평균치 차의 90% 신뢰구간을 이용한다. 2x3 이중 설계된 생동성 시험에서 90% 신뢰구간은 다음과 같이 구할 수 있다.

두 제제의 평균치 차 즉, 두 제제의 효과의 차  $\pi_T - \pi_R$ 의 추정량은 Table 2.1에서 각 셀의 표본 평균을  $\bar{y}_{.jk}$ 라 할 때 다음과 같이 나타낼 수 있다.

$$\hat{\pi}_T - \hat{\pi}_R = \frac{1}{4} \{ (2\bar{y}_{.11} - \bar{y}_{.21} - \bar{y}_{.31}) - (2\bar{y}_{.12} - \bar{y}_{.22} - \bar{y}_{.32}) \}, \quad (2.6)$$

여기서 식 (2.6)의  $\hat{\pi}_T - \hat{\pi}_R$ 은 두 제제의 효과의 차  $\pi_T - \pi_R$ 의 최량 선형 불편추정량 (best linear unbiased estimator, BLUE)으로 알려져 있고 (Kershner와 Federer, 1981),  $\hat{\pi}_T - \hat{\pi}_R$ 의 평균과 분산은 식 (2.3), 식 (2.4), 식 (2.5)에 근거하여 다음과 같다.

$$E(\hat{\pi}_T - \hat{\pi}_R) = \pi_T - \pi_R, \\ \text{Var}(\hat{\pi}_T - \hat{\pi}_R) = \frac{3}{8} \left( \frac{1}{n_1} - \frac{1}{n_2} \right) \sigma_e^2.$$

두 제제의 효과의 차  $\pi_T - \pi_R$ 의 90% 신뢰구간은 개체내 변동  $\sigma_e^2$ 을 이용하여 다음과 같이 표현할 수 있다 (Chow와 Liu, 2008).

$$\left( \hat{\pi} - t_{(0.05, 2(n_1+n_2)-3)} \hat{\sigma}_e \sqrt{\frac{3}{8} \left( \frac{1}{n_1} - \frac{1}{n_2} \right)}, \hat{\pi} + t_{(0.05, 2(n_1+n_2)-3)} \hat{\sigma}_e \sqrt{\frac{3}{8} \left( \frac{1}{n_1} - \frac{1}{n_2} \right)} \right), \quad (2.7)$$

여기서  $\hat{\pi} = \hat{\pi}_T - \hat{\pi}_R$ 이고,  $t_{(0.05, 2(n_1+n_2)-3)}$ 는 자유도가  $2(n_1 + n_2) - 3$ 인 t-분포의 상위 5% 백분위수이다.  $\hat{\sigma}_e$ 는 Table 2.4 분산분석표에서 개체내 잔차 변동인  $MS_{Intra}$ 의 제곱근이다.

두 제제간의 생동성은 시험에서 얻어진 생체이용률 AUCt와 Cmax 자료를 통해 식 (2.7)에서 주어진 90% 신뢰구간을 추정된 후 그 구간이 생동성 허용한계 안에 포함되는지 여부로 평가한다. MFDS에서는 의약품동등성시험기준 제17조를 통해 생동성 허용한계를 생체이용률 AUCt와 Cmax에 대해 각각 다르게 규정하고 있다. AUCt의 허용한계는 기존의 규정과 동일하게 ( $\log 0.8, \log 1.25$ )로 규정하고 있는 반면, Cmax의 허용한계는 대조약의 변동계수 값이 30% 이상인 경우 변동계수 값에 따라 개체내 표준편차에 비례하여 최대 ( $\log 0.6984, \log 1.4319$ )까지 확장된다는 내용으로 허용한계를 다음과 같이 규정하고 있다.

$$[\text{상한치}, \text{하한치}] = \exp[\pm 0.76 \times (\text{로그변환된 대조제제 Cmax값의 개체내 표준편차})]. \quad (2.8)$$

따라서 이 규정을 적용하기 위해 Cmax의 대조제제의 개체내 표준편차 값이 필요하고, 대조제제의 개체내 표준편차 추정값을 얻으려면 대조제제가 2회 이상 투여되는 고차 교차설계가 필요하다는 것을 알 수 있다.

의약품동등성시험기준 제17조 (평가)

④ 제2항에도 불구하고 3기 또는 4기 반복교차시험 결과 대조약 Cmax의 피험자 개체 내 변동계수 [주1]가 30% 이상인 경우, 다음 각 호를 모두 만족하는 경우에 동등으로 판정한다. 단, 이 고시 제2조 제10호에 해당하는 성분은 제외한다.

1. AUCt는 로그변환하여 통계처리 하였을 때, 로그변환한 평균치 차의 90% 신뢰구간이 log0.8 에서 log1.25이내이어야 한다.
2. Cmax는 로그변환하여 통계처리 하였을 때, 로그변환한 평균치 차가 log0.8 에서 log1.25 이내 이어야 하고, 로그변환한 평균치 차의 90% 신뢰구간은 다음의 식에 따라 계산되는 범위를 만족해야 한다. 다만, 변동계수 값이 50% 이상일 때에는 log0.8에서 log1.25이내이어야 한다.

<변동계수가 30%이상일 때 Cmax의 로그변환한 평균치 차의 90% 신뢰구간 범위>  
 [상한치, 하한치] = exp [±0.76 × (로그변환한 대조제제 Cmax값의 개체내 표준편차)]

예 :

변동계수	90% 신뢰구간 (실측치의 비)
30	0.8000 ~ 1.2500
35	0.7723 ~ 1.2948
40	0.7462 ~ 1.3402
45	0.7215 ~ 1.3859
50	0.6984 ~ 1.4319

주 1. 개체 내 변동계수(%) = 100 √{exp [(로그변환한 대조약 Cmax의 개체내 표준편차)]<sup>2</sup> - 1}

⑤ 분산분석에 의한 검정은 원칙적으로 α (유의수준)=0.05에서 실시한다.

2×3 이중설계 시험의 통계 모형인 식 (2.1)에서는 제제에 상관없이 오차항의 분산 ( $\sigma_{WT}^2 = \sigma_{WR}^2 = \sigma_e^2$ )은 동일하다고 가정했지만, 시험제제와 대조제제가 순서 2와 1에서 각각 2회씩 반복 투여되었으므로 시험제제와 대조제제의 개체내 분산 (표준편차) 추정이 가능하다. 시험제제와 대조제제의 개체내 분산을 각각  $\sigma_{WT}^2$ ,  $\sigma_{WR}^2$ 라 하면 이는 다음과 같이 추정할 수 있다.

Table 2.1의 설계에서  $d_{ikl}$ 을  $k$ 번째 순서에  $i$ 번째 개체에게서 얻은 동일 제제  $l$ 의 차라고 하면 다음과 같이 나타낼 수 있고

$$d_{i1R} = y_{i21} - y_{i31},$$

$$d_{i2T} = y_{i22} - y_{i32}.$$

$\sigma_{WT}^2$ 와  $\sigma_{WR}^2$ 에 대한 불편추정량은 다음과 같이 구할 수 있다.

$$\hat{\sigma}_{WT}^2 = \frac{1}{2(n_2 - 1)} \sum_{i=1}^{n_2} (d_{i2T} - \bar{d}_{.2T})^2,$$

$$\hat{\sigma}_{WR}^2 = \frac{1}{2(n_1 - 1)} \sum_{i=1}^{n_1} (d_{i1R} - \bar{d}_{.1R})^2, \tag{2.9}$$

여기서  $\bar{d}_{.kl}$ 은

$$\bar{d}_{.kl} = \sum_{i=1}^{n_k} \frac{d_{ikl}}{n_k}, \quad k = 1, 2, \quad l = T, R$$

이다.

따라서 생체이용률 Cmax의 허용한계는 식 (2.8)에서 로그변환된 대조제 Cmax값의 개체내 표준편차를 식 (2.9)의  $\hat{\sigma}_{WR}^2$ 을 이용하여 구할 수 있다. 새롭게 개정된 MFDS의 의약품동등성시험기준 제17조에 따르면 2x3 이중 설계를 활용했을 때 두 생체이용률 AUCt와 Cmax의 90% 신뢰구간을 식 (2.7)을 활용하여 계산한 후, AUCt는 생동성 허용한계 (log0.8, log1.25), Cmax는 제17조에 근거하여 변동계수 값에 따라 결정되는 허용한계 안에 그 신뢰구간이 포함되면 생동성을 인정한다.

**2.2. 2x3 추가 대조군 설계에서 통계적 추론**

2x3 추가 대조군 설계는 각 순서에 대조제제가 1회씩 더 추가로 투여되는 설계로 각 기간에 배치되는 순서에 따라 (TRR, RTR), (TRR, RRT) 또는 (RTR, RRT) 등으로 설계될 수 있다. 어떤 순서로 배치 되더라도 제제간의 동등성 평가는 동일하게 이루어지기 때문에 Table 2.5와 같이 (TRR, RTR)의 경우로 평균 생동성을 추론해 본다.

두 군 모두에서 대조제제를 반복 투여하는 2x3 추가 대조군 설계가 제제간의 생동성 평가에 통계적으로 관심을 끄는 이유는 의약품동등성시험규정 제17조에서도 보듯이 대조제제의 정확한 추정이 중요하기 때문이다.

**Table 2.5** 2x3 extra reference design

		Period				
		1	2	3		
Sequence	1	T	washout	R	washout	R
	2	R		T		R

2x3 추가 대조군 설계에 대한 통계 모형은 식 (2.1)과 동일하다. 다만 Table 2.5와 같이 설계되었으므로 식 (2.2)에서  $\pi_{(j,k)}$ 가 다음으로 대체된다.

$$\pi_{(j,k)} = \begin{cases} \pi_T, & (j,k) = (1, 1), (2, 2), \\ \pi_R, & (j,k) = (1, 2), (2, 1), (3, 1), (3, 2). \end{cases}$$

식 (2.1)의 통계 모형과 Table 2.5의 설계에 근거하여 3기간에 걸친 k번째 순서, i번째 개체의 생체이용률 관측값  $y_{ik} = (y_{i1k}, y_{i2k}, y_{i3k})'$  ( $i = 1, 2, \dots, n_k$ )는 식 (2.3), 식 (2.4)와 유사하게 다음과 같이 표현할 수 있고, 분산-공분산 행렬은 식 (2.5)와 동일하다.

$$y_{i1} \sim N((\mu + g_1 + p_1 + \pi_T, \mu + g_1 + p_2 + \pi_R, \mu + g_1 + p_3 + \pi_R)', \Sigma), \quad (2.10)$$

$$y_{i2} \sim N((\mu + g_2 + p_1 + \pi_R, \mu + g_2 + p_2 + \pi_T, \mu + g_2 + p_3 + \pi_R)', \Sigma). \quad (2.11)$$

2x3 추가 대조군 설계는 2x3 이중 설계와 비교했을 때 대조제제 4회, 시험제제 2회의 투여로 제제의 측면에서 균형을 이루지 못하고, 2x2 또는 2x4 교차설계와 비교해도 3기간으로 기간과 제제 모두에서 균형을 이루지 못하는 불균형 설계이다. 그렇기 때문에 순서효과와 기간효과, 제제효과 모두를 고려한 균형 있는 직교대비를 유도하는 것이 불가능하다.

또한 생동성 입증을 위한 두 제제의 평균치 차에 대한 90% 신뢰구간을 구하는데 있어서  $2 \times 3$  이중 설계와 유사하게 Table 2.5에서 각 셀의 표본 평균을  $\bar{y}_{.jk}$ 라 하면 순서1, 기간1과 순서2, 기간2에서만 시험제제가 투여되었으므로 각 순서에 1회씩 투여된 시험제제에 2배를 하여 두 제제의 효과의 차  $\pi_T - \pi_R$ 의 추정량  $\hat{\pi}_T - \hat{\pi}_R$ 을 다음과 같이 나타낼 수 있다.

$$\hat{\pi}_T - \hat{\pi}_R = \frac{1}{4} \{ (2\bar{y}_{.11} - \bar{y}_{.21} - \bar{y}_{.31}) - (\bar{y}_{.12} - 2\bar{y}_{.22} - \bar{y}_{.32}) \}. \quad (2.12)$$

식 (2.12)  $\hat{\pi}_T - \hat{\pi}_R$ 의 기댓값은 식 (2.10), 식 (2.11)에 근거하여 식 (2.13)와 같이 얻어지는데 이는 불편성을 만족하지 못한다.

$$E(\hat{\pi}_T - \hat{\pi}_R) = (\pi_T - \pi_R) + \frac{1}{4}(p_1 + p_2 - 2p_3). \quad (2.13)$$

$2 \times 3$  이중 설계에서도 언급했듯이 MFDS 의약품동등성시험기준에서는 두 제제의 평균치 차의 90% 신뢰구간을 이용하여 두 제제의 동등성에 대해 평가할 것을 규정하고 있다. 그러나  $2 \times 3$  추가 대조군 설계는 각 효과에 대한 직교대비를 유도할 수 없고 또한 식 (2.13)에서처럼 두 제제의 효과의 차이가 불편성을 만족하지 못한다. 이는 두 제제의 효과의 차이를 이용하는 생동성 시험에서 두 제제의 효과 이외의 값에 영향을 받고 있음을 의미하고 결과적으로 신뢰구간에도 영향을 주게 된다. 따라서  $2 \times 3$  추가 대조군 설계는 생동성 평가에 적절하다고 볼 수 없고, 추가 처리군 설계 역시 같은 이유로 의미가 없다고 할 수 있다.

### 3. $2 \times 3$ 이중 설계에서 피험자 크기 계산

두 제제간의 생동성을 입증하기 위해서는 생체이용률 AUCt와 Cmax 모두가 규정에서 정하고 있는 허용한계 안에 포함되어야 하는데, 최근 생동성 시험에서 생동성 실패로 판정된 경우를 살펴보면 비교적 변동성이 크게 나타나는 Cmax가 허용구간 안에 포함되지 않아서 결국 실패로 이어지는 경우가 대부분이다. 이와 같은 이유 때문에 MFDS에서도 의약품동등성시험기준 제17조를 추가하여 고변동성 제제의 생동성 평가 기준 중 Cmax에 한하여 허용한계를 기준보다 완화하고 있다. 하지만 아직 완화된 규정에 근거한 개체 크기를 계산하여 제시한 연구 자료는 없는 실정이다. 이에 본 장에서는 Chow와 Wang (2001), Park (2007)에 근거하여  $2 \times 3$  이중 설계된 평균 생동성 시험에서 고변동성 제제를 고려한 피험자 크기를 계산하여 제시하고자 한다. 또한 이 논의는  $2 \times 3$  이중설계시 피험자 크기 계산에 반영되는 변동성 크기를 잘못 계산하여 제제간의 생동성 입증에 실패했을 때  $2 \times 3$  추가시험 (Woo와 Park, 2014)을 계획할 때 함께 고려될 수 있다.

생동성 시험에서 순서1과 순서2의 개체 크기를 각각  $n_1, n_2$ 라 할 때,  $n_1 = n_2 = n$ 이라 하면 군당 피험자 크기는 식 (3.1)과 같이 유도되는데 유도과정은 Park (2007)을 참고하길 바란다.

### 4. $2 \times 3$ 이중 설계 분석 사례

Table 4.1은  $2 \times 3$  이중 설계된 생동성 시험에서 얻어진 가상의 Cmax 생체이용률 원시 자료이다.

Table 4.2는 Table 4.1에 제시된 Cmax 자료의 로그변환한 후 계산된 분산분석 결과이다. Table 4.2의 분산분석표에서 보면 유의수준 5%에서 순서효과는 존재하지 않음을 확인할 수 있어, 제제간의 생동성 평가를 수행할 수 있다. 두 제제의 평균치 차의 90% 신뢰구간은  $(-0.0489, 0.2575)$ 으로 계산된다. 만약 이 결과가  $2 \times 2$  교차설계로 이루어졌다면 신뢰구간의 상한이 기존의 허용한계인 0.2231을 넘



**Table 3.1** Minimum sample size per group

CV	$\theta$	$\Delta=\log 1.25$		$\Delta=\log 1.25 \sim \log 1.4319$	
		1- $\beta$		1- $\beta$	
		0.8	0.9	0.8	0.9
30	0	12	15	12	15
	5	14	19	14	19
	10	27	37	27	37
	15	76	104	76	104
	20	747	1034	747	1034
35	0	16	20	12	15
	5	19	26	13	18
	10	36	50	22	31
	15	101	140	47	64
	20	1001	1387	158	219
40	0	20	25	12	15
	5	24	33	13	17
	10	46	64	19	27
	15	130	179	35	48
	20	1286	1781	81	112
45	0	25	31	12	15
	5	29	40	12	16
	10	57	79	18	24
	15	161	222	28	39
	20	1597	2212	54	75
50	0	30	37	12	15
	5	35	49	12	16
	10	69	95	16	22
	15	194	269	25	34
	20	1933	2677	42	58

어서기 때문에 두 제제간의 생동성을 인정할 수 없다. 하지만 고차 교차설계인 2x3 이중 설계를 이용했기 때문에 의약품동등성시험기준 제 17조를 활용할 수 있다. 대조약 변동계수 값이 약 38%로 고변동성 제제라 판단하여 규정에 따라 확장된 허용한계 (-0.2791, 0.2791)을 적용하면 신뢰구간은 허용구간 안에 포함되어 두 제제간의 생동성을 인정할 수 있다.

### 5. 결론

최근 제네릭 의약품의 사회적 가치의 중요성이 강조되면서 제네릭 의약품을 승인하는 생동성 시험이 재조명되고 있다. 특히 오리지널 의약품의 변동성이 커서 2x2 표준 교차설계를 통해 생동성 입증에 실패하는 사례가 많아지면서, 또한 최근 MFDS 의약품동등성시험기준도 고차 교차설계에 대한 규정이 삽입되면서 2xk 고차 교차설계에 대한 관심이 높아지고 있다.

본 연구에서는 고차 교차설계로 가장 많이 활용되고 있는 2x2 교차설계를 두 번 반복한 2x4 교차설계를 효율성과 윤리성 측면에서 대체할 수 있는 2x3 교차설계의 통계적 특성과 추론에 대해 논의하였다. 2x3 이중설계는 기간이나 제제의 측면에서는 불균형 설계이기는 하나, 고차 교차설계로 인정받아 새롭게 개정된 의약품동등성시험기준 제17조를 적용할 수 있다. 또한 2x3 이중설계와 함께 고려될 수 있는 2x3 추가 대조군 설계를 함께 고찰하여 두 설계의 장단점도 비교하였고, 2x3 이중설계를 활용할 때 새로운 의약품동등성시험기준을 반영한 피험자 크기도 제시하였다.

본 연구는 대표적 고차 설계인 2x4 교차설계가 한 피험자들에게 4번의 투여과정을 실행하기가 어려

**Table 4.1** Cmax raw data in 2×3 dual design

Sequence	Subjects	Period		
		1	2	3
TRR	1	368	381	256
	2	396	140	373
	3	441	323	437
	4	501	382	495
	5	378	347	379
	6	459	316	140
	7	361	457	570
	8	479	438	356
	9	381	464	422
	10	537	365	638
	11	438	421	461
	12	396	476	219
RTT	1	386	454	217
	2	284	536	159
	3	486	416	127
	4	443	478	383
	5	485	410	364
	6	214	338	317
	7	148	375	468
	8	363	231	407
	9	506	458	228
	10	247	424	469
	11	248	532	275
	12	499	470	510

**Table 4.2** Analysis of variance table for 2×3 dual design

sources	degrees of freedom	sums of squares	mean squares	F statistics	Pr(> F)
Between-subjects					
Sequence effect	1	0.1668	0.1668	1.576	0.223
Inter-subjects residuals	22	2.3288	0.1058		
Within-subjects					
Period effect	2	0.327	0.1636	1.229	0.302
Drug effect	1	0.174	0.1740	1.307	0.259
Intra-subjects residuals	45	5.991	0.1331		
Total	71				

운 윤리적인 이슈가 발생하였을 때, 그 대안으로 활용할 수 있는 2×3 이중설계를 제시하고, 생동성 평가의 통계적 절차에 대한 전반적인 내용을 모두 제시하였다는 점에서 의의가 있다 할 것이다.

## References

- Chow, S. C. and Liu, L. (2008). *Design and analysis of bioavailability and bioequivalence studies*, Chapman & Hall.
- Chow, S. C., Shao, J. and Wang, H. (2002). Individual bioequivalence testing under 2×3 designs. *Statistics in Medicine*, **21**, 629-648.
- Chow, S. C. and Wang, H. (2001). On sample size calculation in bioequivalence trials. *Journal of Pharmacokinetics and Pharmacodynamics*, **28**, 155-169.
- EMA (2010). *MEA guideline on the investigation of bioequivalence*, CPMP/EWP/QWP/1401/98 Rev.1/Corr, London.

- Jeong, G. and Park, S. (2011). On evaluation of bioequivalence for highly variable drugs. *Korean Journal of Applied Statistics*, **24**, 1055-1076.
- Kershner R. P. and Federer, W. T. (1981). Two-treatment crossover designs for estimating a variety of effects. *Journal of the American Statistical Association*, **76**, 612-618.
- Noh, S. and Park, S. (2013). Some statistical consideration on 2×k crossover designs for bioequivalence trial. *Korean Journal of Applied Statistics*, **26**, 675-686.
- Park, J. and Park, S. (2016). Assessing bioequivalence for highly variable drugs based on 3×3 crossover designs. *Korean Journal of Applied Statistics*, **29**, 279-289.
- Park, S. (2007). On sample size determination of bioequivalence trials. *Journal of Korean Data & Information Science Society*, **18**, 365-373.
- Woo, H. and Park, S. (2014). Statistical procedures of add-on trials for bioequivalence in 2×k crossover designs. *Journal of Korean Data & Information Science Society*, **25**, 1181-1193.

## Assessing bioequivalence in $2 \times 3$ dual designs

Hwa Hyoung Woo<sup>1</sup> · Gyu Jin Jeong<sup>2</sup> · Sang-Gue Park<sup>3</sup>

<sup>1</sup>Department of Applied Statistics, Chung-Ang University

<sup>2</sup>Department of Business Statistics, Han Nam University

Received 8 June 2017, revised 3 July 2017, accepted 8 July 2017

### Abstract

Assessing bioequivalence between original drug and generic drug is traditionally based on  $2 \times 2$  crossover design. As bioequivalence trials for highly variable drugs are getting popular, the required sample size based on  $2 \times 2$  crossover design would be very large, which might cause the ethical concerns. Regulatory agencies like EMA and MFDS recommended higher order crossover designs such as  $2 \times 4$ ,  $4 \times 2$  and  $4 \times 4$  crossover designs. Alternatively, a  $2 \times 3$  dual design may be recommended in terms of economical and ethical points of view in comparison with the  $2 \times 4$  crossover design for highly variable drug. In this study, we consider some statistical characteristics of  $2 \times 3$  dual design and propose statistical procedures for calculating sample size and assessing bioequivalence based on  $2 \times 3$  dual design. We also discuss the proposed procedures from the perspective of newly revised bioequivalence guidance issued by MFDS.

*Keywords:*  $2 \times 3$  dual design,  $2 \times 3$  extra reference design, bioequivalence, highly variable drug.

---

<sup>1</sup> Lecturer, Department of Applied Statistics, Chung-Ang University, Seoul 06974, Korea.

<sup>2</sup> Professor, Department of Business Statistics, Han Nam University, Daejeon, 4430, Korea.

<sup>3</sup> Corresponding author : Professor, Department of Applied Statistics, Chung-Ang University, Seoul 06974, Korea. E-mail: spark@cau.ac.kr