

공변량을 고려한 2×2 교차설계법에서 평균 생물학적 동등성 평가

정규진¹ · 박상규² · 김관엽³

¹한남대학교 비즈니스통계학과, ²중앙대학교 통계학과, ³한국스코어링 컨설팅사업부

(2010년 10월 접수, 2010년 11월 채택)

요약

소수의 피험자로 이루어지는 생물학적 동등성 시험에서 제제의 특성으로 인해 측정된 생체이용률이 크게 영향을 받는 경우 교차설계법을 사용하더라도 적절한 통제가 어려워지는 경우가 많다. 이 때 적절한 공변량을 통해 이러한 영향력을 통제할 수 있다면 공변량을 고려한 교차설계법 모형이 고려될 필요가 있다. 본 연구논문에서는 공변량을 갖는 2×2 교차설계법을 고려하고 제제간의 생물학적 동등성 평가를 위한 통계적 추론을 제안한다. 제안된 방법은 실제 사례를 통해 논의한다.

주요어: 2×2 교차설계, 공변량, 생물학적 동등성 시험, 생체이용률.

1. 서론

특정 물질로 개발하여 처음 특허를 받은 의약품을 신약이라 하고 신약의 특허기간이 만료된 후 그 신약과 주성분이 같은 물질로 만든 의약품을 복제 의약품이라고 한다. 이러한 복제 의약품은 생물학적 동등성 시험으로 신약과의 동등성을 평가 받는다.

생물학적 동등성(이하 생동성) 시험에 대한 정의는 식품의약품안전청고시로 관리되는 생물학적 동등성 시험기준 제2조에 명시되어 있다. 식품의약품안전청 (2010)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.

일반적으로 생동성 시험은 소수의 피험자들로 실험을 하기 때문에 제제의 특성이나 피험자의 상태에 따라 제제효과가 영향을 받을 수 있다. 특히 생체이용률의 평균을 비교하게 되는 생동성 시험에서 특정 피험자의 상태가 생동성 통계량에 크게 영향을 미치게 되어 생동성 결과를 잘못 판단하는 경우가 생길 수 있다.

보통 2×2 교차설계법으로 진행되는 생동성 시험에서는 동일한 피험자에 대해 실험제제와 대조제제의 생체이용률을 얻기 때문에 두 특성값의 차이를 고려함으로써 각 피험자의 특성을 나타내는 공변량을 고려할 필요가 없어진다. 하지만 제제 혹은 측정되는 생체이용률의 특성 때문에 제제 투여전의 피험자의 상태가 생체이용률의 값에 영향을 준다면 역시 공변량의 고려도 필요할 수 있다. 최근 당뇨병 치료제

본 연구는 2010년도 한남대학교 학술연구조성비 지원에 의하여 연구되었음(2010A364).

²교신저자: (156-756) 서울시 동작구 흑석동 221, 중앙대학교 통계학과, 교수. E-mail: spark@cau.ac.kr

중 탄수화물 흡수를 지연시키는 보글리보스(Voglibose) 제제가 많이 사용되고 있는데 이 제제는 평소 피험자의 식습관이나 식사량과 관련이 많이 되어 있어 이 제제의 생동성 시험에서 잘 관리된 피험자들이라 하더라도 변동성이 크다고 알려져 있다. 이런 경우 피험자의 생체이용률을 직접 비교하는 것 보다 실험 실시 직전의 혈당량 등을 공변량으로 활용하는 것이 제제의 효과를 올바르게 평가할 수 있다. Patel (1983)은 기저치(baseline measurements)가 존재하는 2×2 교차설계법을 고려했지만 이 방법은 동일한 반응값을 반복 측정할 경우로 볼 수 있어 공변량의 특성을 사용했다라고 보기 어렵다. 따라서 공변량 특성을 반영하는 보다 일반적인 2×2 교차설계법의 논의가 필요한 상황이다. 아직까지는 식약청에서 공변량이 있는 2×2 교차설계법을 통한 생동성 평가를 공식적으로 인정하고 있지 않지만 이론적 체계가 갖추어지면 이런 설계법에 의한 분석은 올바른 생동성 평가에 도움을 줄 수 있을 것으로 생각된다.

2장에서는 공변량을 고려한 2×2 교차설계모형에 대한 설명이 주어지고, 3장에서는 이 모형을 이용하여 얻어지는 평균 생동성 평가 방법을 제안한다. 4장에서는 이 모형이 적용되는 사례를 통해 제안된 모형의 효율성을 논의한다.

2. 공변량이 있는 2×2 교차설계모형

교차설계법으로 진행되는 생동성 시험의 특성상 대부분의 경우가 두 개 이상의 공변량이 존재하는 상황은 흔치 않을 것으로 생각되어 하나의 공변량을 가진 모형을 위주로 설명한다. 공변량이 두 개 이상인 경우로의 확장은 선형모형에서 보통 사용하는 방법을 이용하면 가능하다.

공변량이 하나이고 적절한 휴약기간(washout period)이 주어진 2×2 교차설계법으로 이루어진 생동성 시험 모형은 다음과 같이 표현할 수 있다.

$$v_{ijk} = \mu + G_k + S_{ik} + P_j + F_{(j,k)} + \beta u_{ijk} + \epsilon_{ijk},$$

$$S_{ik} \stackrel{iid}{\sim} N(0, \sigma_S^2), \quad \epsilon_{ijk} \stackrel{iid}{\sim} N(0, \sigma_\epsilon^2), \quad i = 1, 2, \dots, n_k, \quad k = 1, 2. \quad (2.1)$$

여기서 v 와 u 는 각각 로그 변환된 생체이용률값(혈중농도-시간곡선하면적 AUC, 최고혈중농도 C_{max} 등)과 공변량을 의미하며, μ 는 생체이용률값의 전체평균을, 그리고 β 는 공변량의 기울기를 나타낸다. 순서효과(G), 기간효과(P), 제제효과(F)는 다음의 조건을 만족하는 생동성 모수들이다.

$$\sum_{k=1}^2 G_k = 0, \quad \sum_{j=1}^2 P_j = 0, \quad \sum_{j=1}^2 \sum_{k=1}^2 F_{j,k} = 0, \quad F_{j,k} = \begin{cases} F_T, & (j, k) = (1, 1), (2, 2), \\ F_R, & (j, k) = (1, 2), (2, 1). \end{cases}$$

S_{ik} 는 k -번째 순서집단에 투입된 i -번째 피험자에게 관측되는 피험자의 변동성을 의미하는 확률변수이고, ϵ_{ijk} 는 k -번째 순서의 j -번째 기간에 i -번째 피험자에게 관측되는 제제의 변동성을 의미하는 확률변수이다. V 와 ϵ_{ijk} 는 통계적으로 서로 독립을 가정하고, S_{ik} 의 분산 σ_S^2 는 피험자간 변동(inter-subject variability)을, ϵ_{ijk} 의 분산 σ_ϵ^2 는 피험자내 변동(intra-subject variability)을 의미한다.

생동성 시험 모형 (2.1)에서 βu_{ijk} 가 없으면 Chow와 Liu (2008)에서 제시하고 있는 일반적인 생동성 시험 모형과 같다. 2×2 교차설계법으로 진행되는 생동성 시험의 통계적 모형에 관한 논의는 Ko와 Oh (1999)를 참고하면 된다. 주어진 순서에서 시험제제 T 와 대조제제 R 의 반응값 v 의 차이를 y 로, 공변량 u 의 차이를 x 로 나타내면, 즉, 첫 번째 순서에서 i -번째 피험자에 대해

$$y_{i1} = v_{i11} - v_{i21}, \quad x_{i1} = u_{i11} - u_{i21}$$

이라 하면 첫 번째 순서에서 y 에 대한 모형으로 다음을 얻는다.

$$y_{i1} = (F_T - F_R) + (P_1 - P_2) + \beta x_{i1} + (e_{i11} - e_{i21}).$$

이와 비슷하게 두 번째 순서에서 i 번째 피험자에 대해

$$y_{i2} = v_{i22} - v_{i12}, \quad x_{i2} = u_{i22} - u_{i12}$$

이라 하면 두 번째 순서에서 y 에 대한 모형으로 다음을 얻는다.

$$y_{i2} = (F_T - F_R) + (P_2 - P_1) + \beta x_{i2} + (e_{i22} - e_{i12}).$$

이를 결합하면 다음과 같은 모형을 얻는다: ($i = 1, 2, \dots, n_k, k = 1, 2$)

$$y_{ik} = \delta + \tau_k + \beta x_{ik} + e_{ik}^*, \quad e_{ik}^* \stackrel{iid}{\sim} N(0, 2\sigma_e^2). \quad (2.2)$$

여기서 $\delta = F_T - F_R$ 이고, $\tau_1 = P_1 - P_2, \tau_2 = P_2 - P_1 = -\tau_1$ 이다.

3. 평균 생동성 평가

공변량이 있는 2×2 교차설계법 모형 (2.2)는 공변량이 있는 일원배치법 모형과 유사한 형태를 가지고 있다. 따라서 평균 생동성의 통계적 추론은 유사하게 진행할 수 있다. 반응변수 y 와 공변량 x 의 제곱합과 곱의 합을 각각 S_{yy}, S_{xx}, S_{xy} 라 하자.

$$\begin{aligned} S_{yy} &= \sum_{k=1}^2 \sum_{i=1}^{n_k} (y_{ik} - \bar{y}_{.k})^2, \\ S_{xx} &= \sum_{k=1}^2 \sum_{i=1}^{n_k} (x_{ik} - \bar{x}_{.k})^2, \\ S_{xy} &= \sum_{k=1}^2 \sum_{i=1}^{n_k} (x_{ik} - \bar{x}_{.k})(y_{ik} - \bar{y}_{.k}). \end{aligned}$$

모형 (2.2)에 대해 β 와 δ 에 대한 최소제곱 추정량과 그 분포는 다음과 같다.

$$\hat{\beta} = \frac{S_{xy}}{S_{xx}} \sim N\left(\beta, \frac{2\sigma_e^2}{S_{xx}}\right), \quad (3.1)$$

$$\hat{\delta} = \frac{\bar{y}_{.1} + \bar{y}_{.2}}{2} - \hat{\beta} \frac{\bar{x}_{.1} + \bar{x}_{.2}}{2} \sim N\left(\delta, \text{Var}(\hat{\delta})\right). \quad (3.2)$$

모형 (2.2)에서 모수 δ 는 $\delta + \tau_1$ 과 $\delta + \tau_2$ 의 평균이므로 최소제곱 추정량 $\hat{\delta}$ 은 $\delta + \tau_1$ 과 $\delta + \tau_2$ 의 최소제곱 추정량 $\widehat{\delta + \tau_k} = \bar{y}_{.k} - \hat{\beta} \bar{x}_{.k}$ ($k = 1, 2$)의 평균을 구하여 얻는다. 추정량 $\hat{\delta}$ 의 분산은 다음과 같다.

$$\text{Var}(\hat{\delta}) = 2\sigma_e^2 \left\{ \frac{1}{4n_1} + \frac{1}{4n_2} + \left(\frac{\bar{x}_{.1} + \bar{x}_{.2}}{2}\right)^2 / S_{xx} \right\} \quad (3.3)$$

위 분산식 (3.3)의 오른쪽에서 마지막 항을 제거하면 공변량이 없는 경우 $\hat{\delta}$ 의 분산이 된다. β 와 δ 에 대한 추론을 하려면 오차항의 분산 $2\sigma_e^2$ 에 대한 추정량이 필요한데 이는 보통의 분산분석에서처럼 오차 제곱합 SSE를 계산하면 된다. 오차제곱합과 $2\sigma_e^2$ 에 대한 추정량, 그리고 그들의 분포는 다음과 같다.

$$\begin{aligned} \text{SSE} &= S_{yy} - \hat{\beta}^2 S_{xx} \sim 2\sigma_e^2 \chi_{n_1+n_2-3}^2, \\ 2\hat{\sigma}_e^2 &= \frac{\text{SSE}}{n_1 + n_2 - 3} \sim 2\sigma_e^2 \frac{\chi_{n_1+n_2-3}^2}{n_1 + n_2 - 3}, \end{aligned} \quad (3.4)$$

표 4.1. 로그변환된 생체이용률 값

순서 집단	피험자 번호	반응값		공변량값	
		기간		기간	
1	1	1.10	0.75	1.31	0.85
	2	1.28	1.33	1.09	1.24
	3	1.84	2.11	1.38	1.90
	4	2.46	2.43	2.27	2.19
	5	1.61	1.81	1.34	1.47
	6	1.20	0.85	1.31	0.85
	7	1.12	1.20	0.96	1.12
	8	0.90	0.88	0.66	0.78
	9	2.48	2.77	1.65	1.90
	10	1.28	1.33	1.09	1.24
	11	1.55	1.78	1.39	1.47
	12	0.94	0.90	0.66	0.78
2	13	2.06	1.88	1.74	1.54
	14	2.48	2.10	2.41	2.13
	15	2.50	2.32	3.05	2.18
	16	1.48	1.30	1.20	1.41
	17	2.28	2.34	1.70	2.21
	18	2.72	2.48	1.89	2.05
	19	1.94	1.11	0.89	0.72
	20	3.35	3.23	2.41	2.83
	21	1.16	1.25	0.96	1.01
	22	2.06	1.38	1.74	1.54
	23	1.48	1.30	1.20	1.41

여기서 χ_n^2 은 자유도 n 인 카이제곱 분포를 나타낸다. 식 (3.1)–(3.4)로부터 $\hat{\delta}$ 의 분포는 다음과 같다.

$$\frac{\hat{\delta} - \delta}{\widehat{\text{Var}}(\hat{\delta})} \sim t_{n_1+n_2-3}, \quad (3.5)$$

여기서 $\widehat{\text{Var}}(\hat{\delta})$ 은 식 (3.3)에 식 (3.4)를 대입하여 얻어지고, t_n 은 자유도 n 인 t -분포를 나타낸다.

두 제제간의 차이 δ 에 대한 평균 생동성 평가는 식 (3.5)를 이용하여 수행할 수 있다. Schuirmann (1987)의 TOST(two one-sided test)도 가능하지만 현재 식품의약품안전청에서 제제간의 생동성 평가에 사용하고 있는 신뢰구간 방법을 제안한다. δ 에 대한 $100(1 - 2\alpha)\%$ 신뢰구간은 다음과 같다.

$$\hat{\delta} \pm t_{\alpha;n_1+n_2-3} \sqrt{\widehat{\text{Var}}(\hat{\delta})} \quad (3.6)$$

만약 식 (3.6)이 $\pm \ln 1.25$ 에 포함되면 두 제제의 평균 생동성을 인정한다. $t_{\alpha;n}$ 은 자유도 n 인 t -분포의 상위 α 백분위수를 나타낸다. 현재 우리나라뿐만 아니라 국제적으로 통용되는 신뢰수준은 90%이다.

4. 사례 및 논의

공변량이 있는 2×2 교차설계법 분석 사례는 Patel (1983)이 사용했던 예제를 본 논문에 맞게 일부 수정하여 사용하였다. 이 예제는 기존의 기관지 천식 치료를 위한 대조제제와 이를 복제한 처리제제의 동등

표 4.2. 공변량을 고려하지 않았을 때 분산분석표

Source	<i>d.f.</i>	SS	MS	<i>F</i>	<i>P</i> -값
Between Subject					
Sequence	1	3.023757	3.023757	3.99	0.059
Subject(Sequence)	21	15.928865	0.758517		
Within Subject					
Drug	1	0.238063	0.238063	7.79	0.0109
Period	1	0.144881	0.144881	4.74	0.0410
Residuals	21	0.641611	0.030553		
Total	45	19.961722			

성을 평가하는 실험 자료이다. 총 23명의 피험자를 대상으로 일정 시간 간격으로 처리제와 대조제를 복용시킨 후 두 제제간의 생동성을 평가하였다. 다만 제제를 투여하기 전 공변량으로 사용할 수 있는 값을 측정한 후 제제를 투여하고 반응값을 측정하였다. 이렇게 설계한 실험에서 투여 전 측정값인 공변량의 유무에 따라서 생동성 평가에 차이가 있는지 확인해 보기로 한다. 표 4.1은 하나의 공변량을 고려한 2×2 교차설계법 자료이고, 측정된 생체이용률값은 생동성 평가를 위해 로그변환한 값이다.

표 4.2는 공변량이 없을 경우 2×2 교차설계법으로 측정된 생체이용률만 가지고 두 제제간의 생체이용률을 비교했을 경우 SAS GLM으로 구현된 분산분석표이다. $\beta_{u_{ijk}}$ 가 없는 생동성 모형 (2.1)에서 두 제제간의 생동성을 평가하는데 모형 순서효과가 잔류효과(carryover effect)와 교락되어 있고, 잔류효과가 존재할 때 처리효과와 교락되어 있다고 알려져 있다 (Ko와 Oh, 1999). 따라서 적절한 처리효과의 통계적 추론을 위해 순서효과의 존재여부를 5% 유의수준에서 검정하는 것이 바람직하다. 분산분석표를 보면 Sequence 항목의 *p*-값이 0.059로 5% 유의수준에서 순서효과가 존재하지 않고 있어 두 제제간의 생동성 평가가 통계적으로 가능함을 알 수 있다. 현재 제제간의 생동성 평가 방법으로 식품의약품안전청에서 요구하고 있는 두 제제간의 90% 신뢰구간을 계산해보면 (0.05524, 0.2328)으로 얻어진다. 이 신뢰구간은 우리나라뿐만 아니라 국제적인 생동성 평가 기준 $\pm \ln 1.25$ 인 [-0.2231, 0.2231]을 벗어나고 있어 두 제제간의 생동성은 통계적으로 유의하지 않은 것으로 결론이 난다.

한편, 공변량을 함께 고려했을 경우 식 (3.1)–(3.4)를 이용하면 다음과 같은 값을 얻게 된다.

$$\hat{\beta} = 0.42136, \quad \text{MSE} = 0.02169,$$

$$\hat{\delta} = 0.12901, \quad \text{Var}(\hat{\delta}) = 0.001913.$$

이를 이용한 δ 의 90% 신뢰구간은 (0.0536, 0.2044)이다. 이 신뢰구간은 생물학적 동등성 기준값인 [-0.2231, 0.2231]에 포함되므로 두 제제가 생물학적으로 동등하다고 할 수 있다. 공변량이 있는 2×2 교차설계법에 대한 SAS Code는 교신저자에 연락하면 제공된다.

일반적으로 생동성 시험은 소수의 피험자로 이루어지기 때문에 교차설계법을 사용하더라도 제제의 특성이 피험자에 영향을 주어 제제효과의 평가가 왜곡되어 시험 결과에 영향을 줄 수 있다. 이런 경우 제제 투여 전에 적절한 공변량을 이용하면 보다 정확한 평가가 내려질 수 있다. 본 연구 논문은 공변량이 존재할 때 두 제제간의 생동성 평가에 관련된 통계적 추론과정을 제안하였다. 현재 우리나라뿐만 아니라 국제적으로도 공변량을 사용하는 생동성 평가는 아직 공식적으로 인정되고 있지 않지만 최근 생동성 평가가 제제의 특성을 충분히 고려하는 방향으로 가고 있기 때문에 본 연구논문의 의의는 있다 하겠다.

참고문헌

- 식품의약품안전청 (2010). 생물학적동등성시험기준(고시 제2010-43호).
- Chow, S. C. and Liu, J. P. (2008). *Design and Analysis of Bioavailability and Bioequivalence Studies*, 3rd edition, Chapman & Hall/CRC.
- Ko, S. and Oh, H. (1999). Assessment and Interpretation for Two Drug Formulations using Crossover Design, *The Korean Communications in Statistics*, **6**, 601-609.
- Patel, H. I. (1983). Use of baseline measurements in the two-period crossover design, *Communications in Statistics, Theory & Methods*, **12**, 2693-2712.
- Schirmann, D. J. (1987). A comparison of the two one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability, *Journal of Pharmacokinetics and Pharmacodynamics*, **15**, 657-680.

Assessing Average Bioequivalence for 2×2 Crossover Design with Covariates

Gyu-Jin Jeong¹ · Sang-Gue Park² · Kwan-Yup Kim³

¹Department of Business Statistics, Hannam University

²Department of Statistics, Chung-Ang University; ³Consulting Section, Korea Scoring Company

(Received October 2010; accepted November 2010)

Abstract

The primary variables are often systematically related to other influences apart from drug effect. For instance, there may be relationships to covariates such as health conditions or prognostic factors. When a 2×2 crossover experiment for bioequivalence is designed, the statistical adjustment for the influence of covariates should be considered if some covariates influence the drug effect. Statistical inference for assessing average bioequivalence for a 2×2 crossover design with covariates is given and an illustrated example is presented with discussion.

Keywords: 2×2 crossover design, bioavailability, average bioequivalence, covariates.

This research was supported by the 2010 Hannam University Research Fund(2010A364).

²Corresponding author: Professor, Department of Statistics, Chung-Ang University, Heukseok-dong, Dongjak-gu, Seoul 156-756, Korea. E-mail: spark@cau.ac.kr