

2 × 2 교차설계법에서 모집단 생물학적 동등성 검정 방법 비교

박상규¹⁾ 임남규²⁾ 이재영³⁾ 김병천⁴⁾

요약

최근 미국을 위시한 선진국에서 제제간의 생물학적 동등성을 판단하는 기준이 생체 이용률의 평균치를 비교하는 시험에서 분산까지 같이 고려하는 기준으로 바뀌고 있다. 처방성과 교차사용성을 의미하는 모집단과 개인 생물학적 동등성이 바로 그것이다. US FDA에서는 2 × 4 교차설계법을 활용해서 제제간의 생물학적 동등성을 입증하는 것을 추천하고 있다. 현재 US FDA에서 제안하고 있는 모집단 생물학적 동등성 평가 방법은 통계적으로 문제점을 가지고 있어 최근 Lee, Shao & Chow(2002), Chow, Shao & Wang(2003), 그리고 McNally, Iyer & Mathew(2002)에 의해서 수정된 평가 방법들이 제안되고 있다. 본 연구논문에서는 그동안 제제간의 생물학적 동등성 평가 설계법이었던 2 × 2 교차설계법을 이용해서 모집단 생물학적 동등성을 평가하는 방법을 논의하고 최근 제안된 방법들을 모의실험을 통해 비교하여 가장 적절한 방법론을 제안한다.

주요용어: 평균 생물학적 동등성, 개인 생물학적 동등성, 모집단 생물학적 동등성, 2 × 2 교차설계, MLS 방법.

1. 모집단 생물학적 동등성 기준

최근 미국을 위시한 선진국에서 제제간의 생물학적 동등성(이하 생동성)을 입증하는 새로운 기준이 제안되었다. 처방성(prescribability)을 의미하는 모집단 생동성(population bioequivalence)과 전환성(switchability) 혹은 교차 사용성을 의미하는 개인 생동성(individual bioequivalence)이 바로 그것들이다.

모집단 생동성은 새로운 환자에게 처방할 수 있는 오리지널 제제와 생동성 제제가 있을 때 이들 두 제제들의 생체이용률(bioavailability)의 분포가 동일하기 때문에 어떤 제제라도 자유롭게 선택할 수 있는 가에 대한 기준이다. 이에 반해 개인 생동성은 기존의 환자가 오리지널 제제를 복용해왔는데 어느 순간에 생동성 제제로 전환하고자 할 때 이것이 가능할 수 있는 가에 대한 기준이다. 기존의 생물학적 동등성 기준이 제제간의 평균 생체이용률을 기준으로 이루어졌다면 모집단 생동성이나 개인 생동성은 평균 생체이용률과 제제간의 생

1) (156-756) 서울시 동작구 흑석동 223, 중앙대학교 수확통계학부, 교수

E-mail: spark@cau.ac.kr

2) (300-716) 대전광역시 동구 용운동 96-3, 대전대학교 정보통계학과, 강의전담전임강사

E-mail: nklim@dju.ac.kr

3) (130-722) 서울시 동대문구 청량리2동 207-43, KAIST 테크노경영대학원 경영공학과, 박사과정

E-mail: rane@kaist.ac.kr

4) (130-722) 서울시 동대문구 청량리2동 207-43, KAIST 테크노경영대학원, 교수

E-mail: bckim@kgsm.kaist.ac.kr

체이용률의 분산의 차이도 함께 고려하는 것이다. 제제간의 분산은 동일하다는 가정 하에서 평균치의 차이만을 평가하는 기존의 평균 생동성 기준보다는 훨씬 사용자의 안전성을 강조한 기준이라고 볼 수 있다.

개인 생동성이 제제간의 평균 생체이용률과 기준에 고려하지 않았던 제제간의 분산이나 개인과 제제간의 교호작용을 고려하기 때문에 이를 평가하기 위해서는 기존의 2×2 교차설계법으로는 가능하지 않다. US FDA(2001)의 경우 최근 개정된 규정에서 2×4 교차설계법을 이용해서 이를 평가하도록 권장하고 있다. 최근 2×3 교차설계법을 이용해서 더 효율적인 평가 방법들이 제시되고 있고 검토하는 방법론도 활발하게 연구되고 있다(Chow, Shao & Wang, 2002).

모집단 생동성은 두 제제간의 생체이용률의 분포의 유사성을 보는 것이므로 평균 생동성을 일반화 한 것으로 볼 수 있다. 모집단 생동성은 두 제제간의 생체이용률의 분포의 유사성을 평가하기 때문에 평균과 분산을 동시에 고려하게 되지만 개인과 제제간의 교호작용을 고려하지 않기 때문에 2×2 교차설계법으로 가능하다. 그동안 우리나라를 비롯하여 미국 등 대부분의 나라에서는 제제간의 생동성 평가를 2×2 교차설계법으로 수행해 왔기 때문에 2×2 교차설계법 하에서 제제간의 생체이용률의 평균 동등성과 모집단 동등성을 동시에 평가해 보는 것은 아직 새로운 기준을 의무화하지 않은 시점에서는 가치 있는 연구가 될 수 있을 것이라 생각된다.

본 연구논문에서는 2장에서 모집단 생동성을 평가하는 통계적 모형을 고려해보고, 3장에서 현재 US FDA 기준과 그에 대응되는 통계적 방법론들을 2×2 교차설계법에 적용하여 논의한다. 4장에서는 모의실험을 통해 가장 좋은 방법론을 제안해 본다.

2. 생동성 통계 모형

US FDA에서 개인이나 모집단 생동성을 평가하기 위해 제안한 모형은 $s \times 2q$ 교차설계법 하에서 Chinchilli & Esinhart(1996)가 제안한 모형을 이용한다.

$$Y_{ijkl} = \mu_k + \gamma_{ikl} + S_{ijk} + \varepsilon_{ijkl}, \quad (2.1)$$

$$(i = 1, \dots, s; j = 1, \dots, n_i; k = T, R; l = 1, \dots, q)$$

여기서 Y_{ijkl} 은 i 번째 순서에서 k 번째 제제를 l 번째 투여한 j 번째 피험자의 반응값이고, 모수 μ_T 와 μ_R 은 시험약과 대조약의 모집단 평균을 의미하고, γ_{ikl} 은 i 번째 순서에서 k 번째 제제를 l 번 투여했을 때의 고정효과를 의미한다. j 번째 피험자 변동을 나타내는 확률 벡터 (S_{ijT}, S_{ijR}) 은 평균이 0이고 다음과 같은 공분산 행렬을 가진다.

$$\Sigma_S = \begin{pmatrix} \sigma_{BT}^2 & \sigma_{TR} \\ \sigma_{TR} & \sigma_{BR}^2 \end{pmatrix},$$

여기서 $\sigma_{TR} = \rho\sigma_{BT}\sigma_{BR}$ 이다. 제제의 변동을 의미하는 ε_{ijkl} 은 평균이 0이고 분산이 σ_{Wk}^2 , ($k = T, R$)인 서로 독립인 확률 변수이다.

만약 2×2 교차설계법이라면 $s = 2$ 이고 $q = 1$ 이 된다. 고정효과는 각 k 에 대해서 $\gamma_{1k1} + \gamma_{2k1} = 0$ 이라는 제약조건을 둘 수 있는데 이럴 경우 2×2 교차설계법에서 순서효과

는 $\gamma_{1T1} + \gamma_{1R1}$ 으로, 기간효과는 $\gamma_{1T1} - \gamma_{1R1}$ 으로 표현될 수 있다. 피험자간 변동을 의미하는 S_{ijk} 와 제제의 변동을 의미하는 ε_{ijkl} 은 2 × 2 교차설계법에서는 서로 교락된다. 즉, 2 × 2 교차설계에서 두 확률오차는 하나로 합쳐져서 다음과 같이 표현될 수 있다.

$$Y_{ijk} = \mu_k + \gamma_{ik} + \varepsilon_{ijk}^* \tag{2.2}$$

모형 (2.2)에서 오차벡터 $\varepsilon_{ij}^* = (\varepsilon_{ijT}^*, \varepsilon_{ijR}^*)'$ 의 공분산 행렬은 다음과 같다.

$$\mathbf{V} = \begin{pmatrix} \sigma_T^2 & \sigma_{TR} \\ \sigma_{TR} & \sigma_R^2 \end{pmatrix}.$$

모형 (2.2)에서 나타나는 오차항의 분산 항들은 실제 모형 (2.1)에서 구분되었던 제제의 오차 ε_{ijkl} 의 분산 σ_{Wk}^2 , ($k = R, T$)와 피험자의 오차 S_{ijk} 의 분산 σ_{Bk}^2 , ($k = R, T$)가 합쳐져 있는 형태라고 볼 수 있다.

3. 모집단 생동성 기준 및 통계적 검정 방법

3.1. JS FDA 방법

현재 US FDA에서 권하고 있는 방법은 Hyslop, Hsuan & Holder(2000)에 의한 개인 생동성을 판단하는 방법을 수정해서 제안하고 있다. 이 방법은 2 × 2 교차설계법에서도 적용될 수 있다. 이 방법은 Howe(1974)가 제안한 서로 독립인 확률 변수들의 선형 결합으로 신뢰구간을 만드는 방법을 일반화한 것이다. US FDA에서는 모집단 생동성을 다음과 같은 가설로 검정한다.

$$H_0 : \xi \geq 0, \quad H_1 : \xi < 0. \tag{3.1}$$

여기서 $\xi = \delta^2 + \sigma_T^2 - \sigma_R^2 - \theta_P \max(\sigma_0^2, \sigma_R^2)$ 이고 $\delta = \mu_T - \mu_R$ 이다. 현재 US FDA에서는 $\theta_P = 0.17448$ 과 $\sigma_0^2 = 0.04$ 를 사용하고 있다. 일반적으로 σ_R^2 은 알려져 있지 않기 때문에 실제 $\max(\sigma_0^2, \sigma_R^2)$ 의 계산은 $\hat{\sigma}_R^2$ 에 의존하게 된다. US FDA에서는 $\sigma_R^2 > \sigma_0^2$ 이면 reference-scaled 기준이라고 하고 $\sigma_R^2 \leq \sigma_0^2$ 인 경우를 constant-scaled 기준으로 구분하고 있다. 가설 (3.1)은 ξ 의 $(1 - \alpha)100\%$ 신뢰구간의 상한을 계산해서 이 구간이 0을 포함하고 있지 않으면 대립가설을 채택하는 것과 동일한 절차이다. 즉, FDA 기준에서는 5% 유의수준을 사용하고 있으므로 ξ 의 95% 신뢰구간의 상한을 계산한 후 이것을 0과 비교해서 상한이 0보다 더 작으면 제제간의 모집단 생동성을 판단하는 것이다.

US FDA에서 제시하고 있는 95% 신뢰구간의 상한은 다음과 같이 계산될 수 있다.

$$d^2 + \hat{\sigma}_T^2 - \hat{\sigma}_R^2 - \theta_P \max(\sigma_0^2, \hat{\sigma}_R^2) + \sqrt{H_d^2 + H_T^2 + (H_R + a')^2}. \tag{3.2}$$

여기서

$$\begin{aligned}
 d_{ij} &= Y_{ijT} - Y_{ijR}, \quad \bar{d}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} d_{ij}, \quad d = \frac{\bar{d}_1 + \bar{d}_2}{2}, \\
 \hat{\sigma}_T^2 &= \frac{SS_T}{\nu}, \quad \hat{\sigma}_R^2 = \frac{SS_R}{\nu}, \\
 \bar{Y}_{i.k} &= \frac{1}{n_i} \sum_{j=1}^{n_i} Y_{ijk}, \quad SS_k = \sum_{i=1}^2 \sum_{j=1}^{n_i} (Y_{ijk} - \bar{Y}_{i.k})^2, \quad k = T, R, \\
 H_d &= \left(|d| + t_{\nu, 0.95} \sqrt{cS_I^2} \right)^2 - d^2, \quad S_I^2 = \frac{1}{n_1 + n_2 - 2} \sum_{i=1}^2 \sum_{j=1}^{n_i} (d_{ij} - \bar{d}_i)^2, \\
 H_T &= SS_T \left(\frac{1}{\chi_{\nu, 0.05}^2} - \frac{1}{\nu} \right), \quad H_R = SS_R \left(\frac{1}{\chi_{\nu, 0.95}^2} - \frac{1}{\nu} \right),
 \end{aligned}$$

이고 a' 은 $\hat{\sigma}_R^2 > \sigma_0^2$ 이면 $\theta_P H_R$ 이고 그렇지 않다면 0이다. 또한 $t_{\nu, \gamma}$ 와 $\chi_{\nu, \gamma}^2$ 는 t -분포와 χ^2 분포에서 얻어진 100 γ 번째 백분위수이고, $c = \frac{1}{4} \left(\frac{1}{n_1} + \frac{1}{n_2} \right)$, $\nu = n_1 + n_2 - 2$ 이다.

US FDA에서 제안한 이 방법은 통계적으로 문제점을 가지고 있다. 즉, $\text{Cov}(\hat{\sigma}_T^2, \hat{\sigma}_R^2) = 0$ 이라는 가정 하에서 이루어지고 있지만 실제 $\text{Cov}(\hat{\sigma}_T^2, \hat{\sigma}_R^2) = 2\rho^2 \sigma_{BT}^2 \sigma_{BR}^2 / (n_1 + n_2 - 2)$ 이기 때문에 이를 무시하는 효과가 나타나 실제 유의수준은 α 보다 훨씬 낮게 나타나고 있다. 따라서 이를 개선한 방법론의 고려는 필수적일 것 같다.

3.2. Lee, Shao & Chow 방법

Lee, Shao & Chow(2002)는 US FDA에서 제시한 방법이 검정통계량의 종속성을 무시하고 계산한 오류를 지적하고 이를 점근적으로 보완하는 통계적 방법을 제안하였다. 즉, $\text{Cov}(\hat{\sigma}_T^2, \hat{\sigma}_R^2)$ 가 0이 아니므로 이에 대한 통계적 보정을 해주지 않는다면 주어진 신뢰수준을 유지하지 못한다는 것이다. $\text{Cov}(\hat{\sigma}_T^2, \hat{\sigma}_R^2)$ 을 계산해서 이를 보정한 ξ 의 $(1 - \alpha)100\%$ 신뢰구간의 상한은 다음과 같다.

$$d^2 + \frac{SS_T}{\nu} - \frac{SS_R}{\nu} - \theta_P \max(\sigma_0^2, \hat{\sigma}_R^2) + \sqrt{(H_d^2 - d^2)^2 + \hat{\lambda}_+^2 \left(\frac{\nu}{\chi_{\nu, \alpha}^2} - 1 \right)^2 + \hat{\lambda}_-^2 \left(\frac{\nu}{\chi_{\nu, 1-\alpha}^2} - 1 \right)^2} \quad (3.3)$$

여기서

$$\hat{\lambda}_{\pm} = \frac{SS_T - b' SS_R \pm \sqrt{(SS_T + b' SS_R)^2 - 4b' SS_T^2}}{2\nu}$$

이고 b' 는 $\hat{\sigma}_R^2 > \sigma_0^2$ 일 때 $1 + \theta_P$ 이고 그렇지 않다면 1이다. 이렇게 계산된 신뢰구간의 상한에 0이 포함되는지의 여부로 가설 (3.1)을 검정할 수 있다. Lee, Shao & Chow (2002)는 (3.3)에 제시한 신뢰구간 상한이 현재 US FDA에서 제시한 기준보다 항상 점근적으로 신뢰수준을 잘 유지한다고 모의실험을 통해 입증하였다.

최근 Chow, Shao & Wang (2003)은 Lee, Shao & Chow (2002)에서 사용하고 있는 방법과 비슷하지만 ξ 의 상한을 계산하는 방법을 조금 달리해서 제안하였다. 모집단 생동성을 판단하는 방법은 Lee, Shao & Chow 방법과 마찬가지로 다음과 같은 ξ 의 상한이 0보다 작으면 생동성이 있다고 판단할 수 있다.

$$d^2 + \frac{SS_T}{\nu} - \frac{SS_R}{\nu} - \theta_P \max(\sigma_0^2, \hat{\sigma}_R^2) + t_{\nu, 0.95} \sqrt{c'}. \quad (3.4)$$

여기서 만약 $\hat{\sigma}_R^2 > \sigma_0^2$ 이면 $c' = V_0$ 이고 그렇지 않다면 $c' = V$ 이다.

$$\begin{aligned} V &= (2d, 1, -(1 + \theta_p))C(2d, 1, -(1 + \theta_p))', \\ V_0 &= (2d, 1, -1)C(2d, 1, -1)', \\ C &= \begin{pmatrix} cS_I^2 & (0, 0) \\ (0, 0)' & \frac{(n_1-1)C_1 + (n_2-1)C_2}{(n_1+n_2-2)^2} \end{pmatrix}, \\ C_1 &= ((Y_{1jT} - \bar{Y}_{1.T})^2, (Y_{1jR} - \bar{Y}_{1.R})^2), j = 1, \dots, n_1 \text{의 표본 공분산 행렬}, \\ C_2 &= ((Y_{2jT} - \bar{Y}_{2.T})^2, (Y_{2jR} - \bar{Y}_{2.R})^2), j = 1, \dots, n_2 \text{의 표본 공분산 행렬}. \end{aligned}$$

이다.

현재 사용되고 있는 US FDA 방법, Lee, Shao & Chow의 방법, 그리고 Chow, Shao & Wang 방법 모두 Howe (1974), Graybill & Wang (1980)이 제시한 Modified Large Sample 방법을 일반화 한 것이다. 통계적으로 본다면 $\text{Cov}(\hat{\sigma}_T^2, \hat{\sigma}_R^2)$ 이 0이 아닌 한 FDA는 Modified Large Sample 방법을 잘못 적용한 것으로 보여지기 때문에 현재의 기준으로 모집단 생동성을 판단한다면 Lee, Shao & Chow 방법이나 Chow, Shao & Wang 방법이 FDA 방법을 대체해야 한다고 생각된다.

3.3. McNally, Iyer & Mathew 방법

모집단 생동성 검정 가설 (3.1)은 다음과 같은 가설로 바꿔 생각해 볼 수 있다.

$$H_0 : \Theta_{PBE} \geq \theta_p, \quad H_1 : \Theta_{PBE} < \theta_p. \quad (3.5)$$

여기서 $\Theta_{PBE} = \frac{\delta^2 + \sigma_T^2 - \sigma_R^2}{\max(\sigma_0^2, \sigma_R^2)}$ 이다. 2 × 2 교차설계법으로부터 모집단 생동성 가설 (3.5)의 검정에 필요한 다음과 같은 서로 독립인 추측변량들을 고려해 보자.

$$\begin{aligned} Z_d &= \frac{d - \delta}{\sqrt{c\sigma_I^2}} \sim N(0, 1), \quad U_T = \frac{SS_T}{\sigma_T^2} \sim \chi_{\nu}^2, \quad U_R = \frac{SS_R}{\sigma_R^2} \sim \chi_{\nu}^2, \\ U_{R|T} &= \frac{SS_{R|T}}{\sigma_{R|T}^2} \sim \chi_{\nu-1}^2, \quad Z_b = (b - \beta) \sqrt{\frac{SS_T}{\sigma_{R|T}^2}} \sim N(0, 1). \end{aligned}$$

여기서 $\sigma_{R|T}^2, SS_{R|T}, \beta$ 와 b 는 다음과 같이 정의 된다.

$$\begin{aligned}\sigma_{R|T}^2 &= \sigma_R^2 - \frac{\sigma_{TR}^2}{\sigma_T^2}, \quad SS_{R|T} = SS_R - \frac{SS_{TR}^2}{SS_T}, \\ SS_{TR} &= \sum_{i=1}^2 \sum_{j=1}^{n_i} (Y_{ijT} - \bar{Y}_{i,T})(Y_{ijR} - \bar{Y}_{i,R}), \\ \beta &= \frac{\sigma_{TR}}{\sigma_T^2}, \quad b = \frac{SS_{TR}}{SS_T}.\end{aligned}$$

확률벡터 $\mathbf{X} = (d, b, SS_T, SS_{R|T})'$ 를 정의하고 이에 대한 관측치 벡터를 \mathbf{x} 라고 하자. 확률벡터 \mathbf{X} 는 모수벡터 $\boldsymbol{\psi} = (\Theta_{PBE}, \zeta)'$ 를 가지고 있다고 정의하면 모수벡터 중 Θ_{PBE} 를 제외한 나머지 모수 $\zeta = (\delta, \beta, \sigma_T^2, \sigma_{R|T}^2)'$ 은 일종의 잡음모수(nuisance parameter)로 볼 수 있다. McNally, Iyer & Mathew(2002)는 모집단 생동성을 검정하는 검정함수 $T(\mathbf{X}; \mathbf{x}, \boldsymbol{\psi})$ 를 다음과 같이 제안하였다.

$$T(\mathbf{X}; \mathbf{x}, \boldsymbol{\psi}) = \frac{(d - Z_d \sqrt{cT_d})^2 + \frac{SS_T}{U_T} - T_R}{\max(T_R, \sigma_0^2)}, \quad (3.6)$$

여기서

$$\begin{aligned}T_d &= \frac{SS_T}{U_T} \left(1 - \left(b - Z_b \sqrt{\frac{SS_{R|T}}{SS_T U_{R|T}}} \right)^2 \right) + \frac{SS_{R|T}}{U_{R|T}}, \\ T_R &= \frac{SS_T}{U_T} \left(b - Z_b \sqrt{\frac{SS_{R|T}}{SS_T U_{R|T}}} \right)^2 + \frac{SS_{R|T}}{U_{R|T}}.\end{aligned}$$

검정통계량 $T(\mathbf{X}; \mathbf{x}, \boldsymbol{\psi})$ 에서 $d, b, SS_T, SS_{R|T}$ 이 주어진 값으로 간주한다면 통계량 $T(\mathbf{X}; \mathbf{x}, \boldsymbol{\psi})$ 의 분포는 잡음모수에 자유로워지며 $T(\mathbf{x}; \mathbf{x}, \boldsymbol{\psi}) = \Theta_{PBE}$ 일 것이다. 또한 검정함수 $T(\mathbf{X}; \mathbf{x}, \boldsymbol{\psi})$ 는 Θ_{PBE} 의 단조증가함수이므로 $T(\mathbf{X}; \mathbf{x}, \boldsymbol{\psi})$ 는 Tsui & Weerahandi (1989)가 제안한 일반화 검정함수 성질들을 만족하게 되어 가설 (3.5)에 대한 검정을 수행할 수 있게 된다. 검정은 검정통계량의 분포를 추론하기 어려워 귀무가설이 사실이라는 가정 하에서 귀무가설을 기각하는 p -값을 계산한 후 주어진 유의수준과 비교하는 방법으로 모집단 생동성을 판단할 수 있다. 하지만 가설 (3.5)에 대한 일반화된 p -값 $P(T(\mathbf{X}; \mathbf{x}, \boldsymbol{\psi}) \geq \theta_p | \Theta_{PBE} = \theta_p)$ 을 계산하는 데는 검정통계량 (3.6)에 포함되어 있는 일부 확률변수 값을 모르기 때문에 직접 계산할 수는 없지만 Monte Carlo 방법으로 표준 정규 확률변수 Z_d 와 Z_b , 자유도 $(n_1 + n_2 - 2)$ 와 $(n_1 + n_2 - 3)$ 의 카이제곱 확률변수 U_T 와 $U_{R|T}$ 의 값을 SAS IML과 같은 프로그램으로 난수로 발생시켜 추정할 수 있다. 즉, 최소한 10,000번의 반복을 통하여 계산된 $T(\mathbf{X}; \mathbf{x}, \boldsymbol{\psi})$ 로부터 θ_p 보다 큰 근사적인 확률을 계산할 수 있게 되고 이 근사적인 확률은 바로 p -값이 된다. McNally, Iyer & Mathew(2002) 검정은 이처럼 Monte Carlo 방법을 통한 많은 수의 반복으로 계산되기 때문에 실제적인 사용에 제약을 가져올 수 있다.

4. 모의실험과 결론

모집단 생동성을 판단하는 네 가지 방법들이 주어진 유의수준을 얼마나 잘 유지하는지를 판단하고자 모의실험을 수행하였다. 모의실험에 나타난 결과치들은 5,000번 반복을 기준으로 작성되었고 MaNally, Iyer & Mathew(2002)의 방법에서 p -값은 10,000번 반복에 기초한 Monte Carlo 방법으로 계산되었다.

모의실험은 모형 (2.2)에 의해서 이변량 정규분포로부터 확률변수를 추출하였다. 즉,

$$(y_{ijT}, y_{ijR}) \sim N \left((\mu_T, \mu_R), \begin{bmatrix} \sigma_{BT}^2 + \sigma_{WT}^2 & \sigma_{TR} \\ \sigma_{TR} & \sigma_{BR}^2 + \sigma_{WR}^2 \end{bmatrix} \right) \quad (4.1)$$

여기서 $\sigma_{TR} = \rho\sigma_{BT}\sigma_{BR}$, $\mu_T = \mu_R + \delta$ 이고 $\mu_R = 0$ 의 값을 주었다.

먼저 순서 1인 경우에 SAS IML의 NORMAL 함수를 사용해서 표준 정규 분포로부터 두 가지 독립적인 확률 표본 $(x_{11R}, \dots, x_{1n_1R})$ 와 $(x_{11T}, \dots, x_{1n_1T})$ 를 추출한다. 다음에

$$y_{1jR} = \mu_R + \sigma_R x_{1jR}, \quad y_{1jT} = \mu_T + \sigma_T \{ \rho_a x_{1jR} + \sqrt{1 - \rho_a^2} x_{1jT} \}$$

의 변환을 사용해서 식 (4.1)과 같은 공분산 구조를 가지는 n_1 개의 독립 확률 표본 (y_{1jR}, y_{1jT}) 를 생성한다. 위의 식에서 $\rho_a = \rho\sigma_{BR}\sigma_{BT}/\sigma_R\sigma_T$, $\sigma_R^2 = \sigma_{BR}^2 + \sigma_{WR}^2$, 그리고 $\sigma_T^2 = \sigma_{BT}^2 + \sigma_{WT}^2$ 가 된다. 순서 2인 경우에도 마찬가지로의 방식으로 n_2 개의 (y_{2jT}, y_{2jR}) 을 생성해서 네 가지 방법을 적용한다.

표 4.1: 모집단 생동성 검정에 대한 모의실험 결과 요약표 (평균, 표준오차)

		FDA	EMLS	CSW	GPV
높은 상관 관계	생동성 아님 ($\delta \neq 0$)	.025 (.0067)	.045 (.0034)	.035 (.0091)	.061 (.0082)
	생동성 ($\delta = 0$)	.016 (.0091)	.043 (.0041)	.064 (.015)	.054 (.011)
		.75 (.085)	.84 (.085)	.84 (.084)	.88 (.068)
낮은 상관 관계	생동성 아님 ($\delta \neq 0$)	.038 (.0030)	.041 (.0027)	.040 (.0088)	.066 (.0055)
	생동성 ($\delta = 0$)	.030 (.0032)	.040 (.0033)	.057 (.016)	.047 (.0060)
		.68 (.079)	.71 (.075)	.71 (.080)	.75 (.066)

[표 4.1]은 다양한 ($\delta, \sigma_{WR}, \sigma_{BR}, \sigma_{WT}, \sigma_{BT}$) 조건에서 수행된 모의실험 결과를 수록한 [표 4.2], [표 4.3]과 [표 4.4]를 요약한 표이다. 각각의 표에는 3장에서 설명한 현재 US FDA에서 채택하고 있는 FDA 방법, Lee, Shao & Chow (2002)에 의해서 제안된 EMLS(Extended

표 4.2: 모집단 생동성 검정 방법에 대한 근사적인 유의수준 비교 ($\delta \neq 0$)

		$\delta = .2642$				$\delta = .2954$				$\delta = .3736$			
		$\sigma_{BT} = .1, \sigma_{BR} = .1$ $\sigma_{WT} = .1, \sigma_{WR} = .1$				$\sigma_{BT} = .2, \sigma_{BR} = .2$ $\sigma_{WT} = .1, \sigma_{WR} = .1$				$\sigma_{BT} = .2, \sigma_{BR} = .2$ $\sigma_{WT} = .2, \sigma_{WR} = .2$			
n	ρ	FDA	EMLS	CSW	GPV	FDA	EMLS	CSW	GPV	FDA	EMLS	CSW	GPV
10	.2	.044	.045	.070	.056	.043	.045	.043	.079	.035	.039	.037	.069
	.3	.042	.043	.071	.055	.036	.041	.040	.069	.036	.039	.037	.064
	.8	.038	.043	.067	.065	.022	.049	.032	.062	.033	.042	.033	.070
12	.2	.037	.043	.063	.053	.015	.046	.027	.072	.032	.044	.035	.055
	.3	.040	.042	.064	.053	.041	.046	.041	.072	.043	.044	.042	.065
	.8	.044	.046	.068	.054	.038	.042	.037	.074	.037	.040	.035	.060
14	.2	.042	.046	.064	.052	.024	.050	.034	.062	.031	.041	.030	.062
	.3	.039	.045	.064	.056	.017	.044	.028	.060	.030	.044	.035	.062
	.8	.048	.049	.072	.059	.034	.037	.036	.065	.040	.042	.038	.065
	.2	.039	.040	.062	.050	.042	.044	.041	.067	.036	.038	.039	.061
	.3	.034	.040	.058	.060	.023	.045	.033	.053	.035	.042	.036	.070
	.8	.038	.043	.060	.037	.020	.049	.034	.073	.029	.039	.032	.064
		$\delta = .5907$				$\delta = .5604$				$\delta = .8861$			
		$\sigma_{BT} = .4, \sigma_{BR} = .4$ $\sigma_{WT} = .2, \sigma_{WR} = .2$				$\sigma_{BT} = .3, \sigma_{BR} = .3$ $\sigma_{WT} = .3, \sigma_{WR} = .3$				$\sigma_{BT} = .6, \sigma_{BR} = .6$ $\sigma_{WT} = .3, \sigma_{WR} = .3$			
n	ρ	FDA	EMLS	CSW	GPV	FDA	EMLS	CSW	GPV	FDA	EMLS	CSW	GPV
10	.2	.036	.038	.038	.061	.039	.041	.037	.069	.038	.040	.040	.064
	.3	.033	.037	.034	.065	.034	.038	.038	.065	.039	.045	.039	.070
	.8	.018	.041	.028	.059	.031	.040	.037	.072	.020	.047	.028	.055
12	.2	.018	.043	.027	.053	.032	.043	.037	.067	.012	.042	.029	.058
	.3	.038	.041	.037	.064	.039	.041	.041	.070	.039	.042	.037	.065
	.8	.035	.040	.030	.061	.034	.038	.036	.062	.040	.044	.040	.068
14	.2	.021	.047	.032	.073	.031	.041	.034	.064	.024	.050	.039	.050
	.3	.015	.048	.030	.051	.030	.041	.034	.044	.017	.050	.029	.066
	.8	.036	.040	.033	.061	.039	.041	.039	.064	.041	.042	.037	.065
	.2	.041	.043	.039	.064	.041	.043	.039	.064	.034	.037	.035	.061
	.3	.026	.047	.032	.057	.038	.047	.039	.052	.024	.048	.031	.057
	.8	.016	.044	.026	.057	.035	.047	.033	.066	.017	.051	.034	.072
		$\delta = .2954$				$\delta = .2954$				$\delta = .5907$			
		$\sigma_{BT} = .2, \sigma_{BR} = .1732$ $\sigma_{WT} = .1, \sigma_{WR} = .1414$				$\sigma_{BT} = .1732, \sigma_{BR} = .2$ $\sigma_{WT} = .1414, \sigma_{WR} = .1$				$\sigma_{BT} = .3873, \sigma_{BR} = .4$ $\sigma_{WT} = .2236, \sigma_{WR} = .2$			
n	ρ	FDA	EMLS	CSW	GPV	FDA	EMLS	CSW	GPV	FDA	EMLS	CSW	GPV
10	.2	.039	.040	.044	.074	.040	.044	.042	.073	.036	.040	.037	.066
	.3	.037	.041	.041	.076	.034	.040	.039	.073	.037	.041	.035	.066
	.8	.028	.046	.034	.056	.025	.043	.034	.076	.020	.040	.028	.064
12	.2	.020	.041	.031	.074	.024	.045	.034	.062	.019	.056	.035	.066
	.3	.037	.039	.040	.070	.041	.042	.043	.073	.041	.043	.037	.067
	.8	.039	.042	.039	.070	.039	.042	.039	.073	.041	.045	.040	.068
14	.2	.023	.037	.028	.064	.029	.047	.033	.068	.022	.046	.031	.069
	.3	.024	.047	.034	.073	.026	.048	.032	.073	.017	.046	.029	.052
	.8	.040	.042	.039	.072	.037	.037	.036	.062	.036	.037	.034	.062
	.2	.038	.041	.037	.070	.038	.041	.036	.068	.040	.042	.037	.061
	.3	.030	.048	.037	.068	.030	.047	.034	.065	.023	.042	.029	.059
	.8	.023	.044	.028	.058	.026	.050	.033	.070	.019	.042	.030	.059
		$\delta = .5907$				$\delta = .8861$				$\delta = .8861$			
		$\sigma_{BT} = .4, \sigma_{BR} = .3873$ $\sigma_{WT} = .2, \sigma_{WR} = .2236$				$\sigma_{BT} = .6, \sigma_{BR} = .5$ $\sigma_{WT} = .3, \sigma_{WR} = .4472$				$\sigma_{BT} = .5, \sigma_{BR} = .6$ $\sigma_{WT} = .4472, \sigma_{WR} = .3$			
n	ρ	FDA	EMLS	CSW	GPV	FDA	EMLS	CSW	GPV	FDA	EMLS	CSW	GPV
10	.2	.031	.033	.035	.062	.039	.041	.043	.072	.031	.034	.030	.061
	.3	.036	.040	.039	.067	.036	.040	.034	.068	.035	.038	.036	.064
	.8	.023	.045	.035	.050	.024	.042	.031	.068	.023	.042	.031	.063
12	.2	.014	.043	.030	.066	.023	.044	.030	.068	.029	.049	.034	.053
	.3	.041	.043	.041	.068	.038	.042	.038	.067	.038	.040	.039	.070
	.8	.038	.044	.039	.068	.038	.041	.039	.065	.038	.042	.040	.068
14	.2	.023	.046	.030	.077	.030	.044	.034	.060	.029	.047	.036	.056
	.3	.018	.044	.028	.049	.028	.045	.033	.051	.028	.049	.035	.057
	.8	.039	.042	.038	.066	.038	.041	.037	.062	.040	.043	.039	.065
	.2	.037	.041	.037	.063	.036	.040	.037	.063	.036	.039	.036	.062
	.3	.025	.044	.031	.073	.028	.043	.032	.060	.028	.044	.032	.061
	.8	.018	.049	.031	.062	.020	.038	.023	.049	.027	.042	.031	.061

Modified Large Sample) 방법, Chow, Shao & Wang(2003)에 의해 제안된 CSW 방법 그리고 McNally, Iyer & Mathew(2002)에 의해서 제안된 GPV 방법을 비교하고 있다. [표 4.2]와 [표 4.3]은 각 순서 당 표본의 크기가 10, 12, 14이고 상관관계가 낮은(0.2, 0.3) 경우와 높은(0.8, 0.9) 경우에 이들 네 가지 방법들에 대한 귀무가설 하에서 추정된 유의수준을 비교한 표이고 [표 4.4]는 실제 두 제제간의 모집단 생동성이 있는 상황 하에서 이들 방법들에 대한 추정된 검정력을 비교한 표이다.

표 4.3: 모집단 생동성 검정 방법에 대한 근사적인 유의수준 비교 ($\delta = 0$)

n	ρ	$\sigma_{BT} = .2825, \sigma_{BR} = .1$ $\sigma_{WT} = .1, \sigma_{WR} = .1$				$\sigma_{BT} = .3567, \sigma_{BR} = .2$ $\sigma_{WT} = .1, \sigma_{WR} = .1$				$\sigma_{BT} = .4238, \sigma_{BR} = .2$ $\sigma_{WT} = .2, \sigma_{WR} = .2$			
		FDA	EMLS	CSW	GPV	FDA	EMLS	CSW	GPV	FDA	EMLS	CSW	GPV
10	.2	.028	.029	.088	.035	.032	.035	.057	.058	.024	.028	.044	.041
	.3	.034	.035	.088	.039	.025	.031	.050	.047	.028	.034	.052	.051
	.8	.027	.033	.089	.040	.013	.047	.077	.084	.016	.034	.046	.051
	.9	.031	.041	.090	.034	.0092	.054	.086	.070	.015	.039	.058	.052
12	.2	.031	.033	.090	.036	.029	.033	.050	.053	.030	.033	.046	.048
	.3	.032	.035	.089	.038	.027	.034	.052	.050	.032	.036	.049	.051
	.8	.035	.043	.089	.045	.013	.045	.063	.067	.020	.042	.054	.053
	.9	.029	.041	.090	.048	.0064	.048	.078	.076	.016	.043	.054	.067
14	.2	.034	.035	.088	.039	.025	.031	.050	.047	.028	.034	.052	.051
	.3	.031	.033	.089	.035	.035	.042	.057	.058	.033	.037	.045	.049
	.8	.034	.040	.090	.045	.012	.046	.067	.073	.024	.041	.054	.054
	.9	.031	.040	.090	.047	.0078	.052	.069	.064	.017	.041	.051	.051
n	ρ	$\sigma_{BT} = .7134, \sigma_{BR} = .4$ $\sigma_{WT} = .2, \sigma_{WR} = .2$				$\sigma_{BT} = .6357, \sigma_{BR} = .3$ $\sigma_{WT} = .3, \sigma_{WR} = .3$				$\sigma_{BT} = 1.0701, \sigma_{BR} = .6$ $\sigma_{WT} = .3, \sigma_{WR} = .3$			
		FDA	EMLS	CSW	GPV	FDA	EMLS	CSW	GPV	FDA	EMLS	CSW	GPV
10	.2	.026	.031	.049	.046	.029	.033	.049	.046	.029	.034	.053	.049
	.3	.029	.037	.054	.053	.028	.033	.049	.050	.030	.036	.053	.051
	.8	.0082	.044	.053	.059	.021	.041	.051	.059	.010	.044	.059	.053
	.9	.0048	.041	.060	.053	.016	.043	.052	.038	.0066	.040	.058	.044
12	.2	.032	.038	.056	.053	.025	.030	.042	.041	.027	.031	.042	.043
	.3	.034	.040	.050	.051	.029	.033	.045	.046	.033	.040	.054	.052
	.8	.011	.044	.059	.051	.022	.042	.055	.053	.010	.038	.053	.052
	.9	.004	.044	.057	.045	.016	.041	.054	.049	.0042	.043	.055	.056
14	.2	.032	.037	.049	.047	.038	.041	.050	.052	.034	.037	.050	.049
	.3	.032	.038	.049	.048	.032	.037	.052	.052	.030	.035	.049	.048
	.8	.010	.044	.056	.057	.024	.043	.054	.064	.011	.048	.058	.051
	.9	.0042	.044	.054	.050	.019	.048	.062	.054	.0046	.040	.052	.040

[표 4.1]은 [표 4.2], [표 4.3] 그리고 [표 4.4]를 표본의 크기와 모수의 특성들을 상관관계 차이와 생동성 여부라는 특성으로 간단히 요약하고 있지만 비교적 [표 4.2], [표 4.3]과 [표 4.4]에서 나타난 특성들을 잘 요약하고 있음을 볼 수 있다. 즉, [표 4.2]와 [표 4.3]에서 나타났듯이 네 가지 방법 중에서 현재 US FDA에서 채택하고 있는 방법은 예상했듯이 상관계수가 높은 수준에서 주어진 유의수준을 전혀 맞추지 못하는 것으로 나타나고 있다. Chow, Shao & Wang의 CSW 방법은 상관관계가 높고 낮음에 상관없이 $\delta \neq 0$ 인 경우 주어진 유의수준을 과소 추정하는 경향이 나타나고 $\delta = 0$ 인 경우는 주어진 유의수준을 약간 과대 추정하는 현상을 보이고 있다. McNally, Iyer & Mathew의 GPV 방법 역시 상관계수의 값에 크게 영향을 받지 않고 있으며 $\delta \neq 0$ 인 경우에 유의수준을 과대 추정하는 경향이 나타나고 있지만 $\delta = 0$ 인 경우는 비교적 유의수준을 맞추고 있다고 볼 수 있다. Lee, Shao & Chow의 EMLS 방법 역시 상관계수의 높고 낮음에 크게 영향을 받지 않지만 $\delta \neq 0$ 인 경우나 $\delta = 0$ 인 경우 모두 주어진 유의수준을 과소 추정하고 있다. 각 평균치의 표준오차를 비교하여 보면 다른 어떤 방법보다 Lee, Shao & Chow의 EMLS 방법이 가장 작게 나타나고 있어 모의실험 결과를 가장 믿을 만한 방법임을 보여주고 있다.

표 4.4: 모집단 생동성 검정 방법에 대한 근사적 검정력 비교 ($\delta = .05$)

		$\sigma_{WT} = .15, \sigma_{WR} = .15$							
		$\sigma_{BT} = .15, \sigma_{BR} = .15, (\xi = -.076)$				$\sigma_{BT} = .3, \sigma_{BR} = .3, (\xi = -.19)$			
n	ρ	FDA	EMLS	CSW	GPV	FDA	EMLS	CSW	GPV
10	.2	.665	.684	.769	.774	.556	.580	.598	.657
	.3	.667	.686	.785	.782	.582	.619	.632	.690
	.8	.724	.784	.844	.865	.656	.810	.790	.861
	.9	.732	.799	.864	.868	.688	.873	.855	.889
12	.2	.754	.770	.832	.847	.661	.678	.691	.734
	.3	.768	.785	.848	.858	.674	.697	.704	.751
	.8	.816	.859	.899	.909	.758	.875	.864	.901
	.9	.831	.878	.911	.931	.784	.923	.908	.947
14	.2	.847	.854	.894	.903	.739	.752	.753	.796
	.3	.830	.840	.893	.896	.758	.779	.776	.816
	.8	.882	.914	.931	.943	.842	.925	.914	.939
	.9	.896	.931	.949	.966	.874	.961	.955	.961
		$\sigma_{WT} = .23, \sigma_{WR} = .23$							
		$\sigma_{BT} = .23, \sigma_{BR} = .23, (\xi = -.18)$				$\sigma_{BT} = .46, \sigma_{BR} = .46, (\xi = -.46)$			
n	ρ	FDA	EMLS	CSW	GPV	FDA	EMLS	CSW	GPV
10	.2	.560	.578	.596	.657	.582	.605	.609	.672
	.3	.566	.590	.604	.663	.585	.620	.625	.686
	.8	.608	.673	.663	.738	.659	.807	.796	.846
	.9	.605	.683	.683	.745	.697	.878	.865	.902
12	.2	.647	.660	.671	.721	.667	.686	.694	.737
	.3	.670	.684	.693	.741	.683	.707	.710	.758
	.8	.697	.751	.746	.791	.783	.890	.881	.896
	.9	.718	.781	.771	.817	.803	.926	.913	.937
14	.2	.743	.751	.751	.794	.750	.763	.770	.805
	.3	.730	.742	.739	.789	.760	.783	.781	.824
	.8	.786	.828	.821	.865	.845	.925	.916	.934
	.9	.785	.838	.829	.880	.872	.961	.954	.972
		$\sigma_{WT} = .3, \sigma_{WR} = .3$							
		$\sigma_{BT} = .3, \sigma_{BR} = .3, (\xi = -.31)$				$\sigma_{BT} = .6, \sigma_{BR} = .6, (\xi = -.78)$			
n	ρ	FDA	EMLS	CSW	GPV	FDA	EMLS	CSW	GPV
10	.2	.561	.579	.602	.659	.580	.603	.619	.673
	.3	.564	.587	.610	.662	.594	.627	.637	.694
	.8	.612	.674	.671	.729	.665	.810	.798	.856
	.9	.618	.691	.690	.750	.698	.876	.863	.899
12	.2	.671	.683	.690	.737	.680	.693	.701	.748
	.3	.677	.694	.699	.748	.692	.716	.719	.772
	.8	.707	.756	.753	.806	.774	.878	.868	.891
	.9	.711	.777	.768	.839	.810	.937	.929	.935
14	.2	.736	.747	.754	.793	.746	.759	.764	.799
	.3	.739	.754	.761	.804	.771	.791	.791	.830
	.8	.787	.831	.824	.863	.847	.929	.914	.940
	.9	.791	.844	.842	.874	.874	.962	.955	.970
		$\sigma_{WT} = .5, \sigma_{WR} = .5$							
		$\sigma_{BT} = .5, \sigma_{BR} = .5, (\xi = -.87)$				$\sigma_{BT} = 1.0, \sigma_{BR} = 1.0, (\xi = -2.1)$			
n	ρ	FDA	EMLS	CSW	GPV	FDA	EMLS	CSW	GPV
10	.2	.561	.578	.599	.655	.575	.601	.609	.668
	.3	.586	.604	.621	.677	.592	.621	.625	.685
	.8	.608	.673	.679	.755	.673	.818	.801	.857
	.9	.632	.709	.716	.759	.708	.881	.869	.895
12	.2	.653	.669	.683	.730	.671	.687	.693	.745
	.3	.679	.695	.708	.745	.683	.710	.710	.760
	.8	.714	.765	.761	.820	.780	.890	.881	.906
	.9	.729	.792	.785	.816	.809	.930	.920	.947
14	.2	.740	.751	.759	.798	.759	.774	.782	.816
	.3	.760	.774	.778	.813	.766	.789	.786	.829
	.8	.799	.839	.829	.869	.862	.933	.922	.933
	.9	.804	.852	.843	.880	.880	.963	.955	.974

검정력을 비교한 [표 4.4]의 경우는 McNally, Iyer & Mathew의 GPV 방법이 가장 검정력이 우수하게 나타나고 있는데 이것은 유의수준부터 조금씩 과대 추정하는 경향에서 기인된 것으로 보여지기 때문에 Chow, Shao & Wang의 CSW 방법이나 Lee, Shao & Chow의 EMLS 방법보다 검정력이 우수하다고 주장하는 것에는 다소 무리가 따른다고 생각된다. Chow, Shao & Wang의 CSW 방법과 Lee, Shao & Chow의 EMLS 방법은 거의 비슷한 검정

력을 보여주고 있음을 볼 수 있지만 주어진 유의수준을 유지하는 상황을 같이 고려한다면 Lee, Shao & Chow의 EMLS 방법이 현재까지는 가장 추천할 만한 방법으로 생각된다.

모집단 생동성이나 개인 생동성은 현재 미국을 비롯한 선진국에서 새롭게 채택되고 있는 기준이다. 현재 제제간의 평균치를 이용해서 생동성을 판단하는 우리나라에서는 조만간 이런 기준을 고려해야 할 것으로 보인다. 본 연구는 크게 두 가지 목적으로 시행되었는데 그 첫째는 모집단 생동성을 검정하는데 현재 US FDA에서 사용된 기준에 문제가 있다는 것을 알리고 보다 바른 통계적 방법론을 제시하고자 하는 것이고, 두 번째는 현재 제제간의 생동성을 판단하는 기준으로 사용되고 있는 2 × 2 교차설계법에서도 모집단 생동성은 가능하기 때문에 평균 생동성을 수행하는 과정에서 모집단 생동성도 같이 평가해보는 것도 의미가 있다는 것을 알리기 위해서이다. 한편 최근 미국 FDA에서도 아직 연구가 이루어지지 않고 있는 모집단 생동성을 평가할 때 결측치가 생긴 경우 이에 대한 통계적 방법론을 개발하거나 제제간의 생동성 판단에 포함시켜야 하는 이상치 유무를 판단하는 통계적 절차는 계속 연구되어야 할 것으로 생각된다.

참고문헌

- Chinchilli, V. M. and Esinhart, J. D. (1996). Design and analysis of intra-subject variability in cross-over experiment, *Statistics in Medicine*, **15**, 1619-1634.
- Chow, S. C. and Liu, J. P. (2000). *Design and Analysis of Bioavailability and Bioequivalence Studies*, Marcel Dekker, New York. 2nd Edition.
- 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., Shao, J. and Wang, H. (2003). Statistical tests for population bioequivalence, *Statistica Sinica*, **13**, 539-554.
- Graybill, F. and Wang, C. M. (1980). Confidence intervals on nonnegative linear combinations of variances, *Journal of the American Statistical Association*, **75**, 869-873.
- Howe, W. G. (1974). Approximate confidence limits on the mean of X+Y where X and Y are two tabled independent random variables, *Journal of the American Statistical Association*. **69**, 789-794.
- Hyslop, T., Hsuan, F. and Holder, D. J. (2000). A small sample confidence interval approach to assess individual bioequivalence, *Statistics in Medicine*, **19**, 2885-2897.
- Lee, Y., Shao, J. and Chow, S. C. (2002). Confidence intervals for linear combinations of variance components when estimators of variance components are dependent: an extension of the modified large sample method with applications, Submitted.
- McNally, R. J., Iyer, H. and Mathew, T. (2002). Tests for individual and population bioequivalence based on generalized p-values, *Statistics in Medicine*, **22**, 31-53.
- Tsui, K. W. and Weerahandi, S. (1989). Generalized p-values in significance testing of hypothesis in the presence of nuisance parameters, *Journal of the American Statistical Association*. **84**, 602-607.
- US Food and Drug Administration (1992). *Guidance on Statistical Procedures for Bioequivalence Using a Standard Two-treatment Crossover Design*, Rockville, MD.

US Food and Drug Administration (2001). *Statistical Approaches to Establishing Bioequivalence*, Rockville, MD.

[2004년 6월 접수, 2004년 11월 채택]

A Comparative Study of Statistical Methods for Population Bioequivalence in 2×2 Crossover Design

Sang-Gue Park¹⁾ Nam-Kyoo Lim²⁾ Jae-Young Lee³⁾ Byung-Chun Kim⁴⁾

ABSTRACT

The US Food and Drug Administration(FDA) recommends that population bioequivalence and individual bioequivalence would be assessed to address the prescribability and switchability between a brand-name drug and its new formulation or generic copy in its 2001 guidance document. The test for population bioequivalence in the latest FDA guidance is recommended in 2×4 crossover design, but it turns out to be very conservative. Recently Lee, Shao & Chow(2002), Chow, Shao & Wang(2003) and McNally, Iyer & Mathew(2002) proposed new statistical methods for assessing population bioequivalence between drugs to correct the biasness of current FDA method. Since 2×2 crossover experiment is most welcomed design in bioequivalence testing, we adopt their methods to 2×2 crossover designs and compare their methodologies with FDA one through the simulation study.

Keywords: Average bioequivalence; Individual bioequivalence; Population bioequivalence; 2×2 crossover design; Modified large sample method.

1) Professor, Division of Mathematics and Statistics, Chung-Ang University, Seoul, 156-756, Korea.

E-mail: spark@cau.ac.kr

2) Full-time Lecturer, Department of Information and Statistics, Daejeon University, Daejeon, 300-716, Korea.

E-mail: nklim@dju.ac.kr

3) PhD Candidate at Division of Management Engineering, KAIST Graduate School of Management, Seoul, 130-722, Korea.

E-mail: rane@kaist.ac.kr

4) Professor, KAIST Graduate School of Management, Seoul, 130-722, Korea.

E-mail: bckim@kgs.m.kaist.ac.kr