

개정된 생물학적 동등성 시험 기준에 대한 통계적 고찰

고승곤¹⁾ 양지선²⁾ 김정일³⁾ 임남규⁴⁾

요약

2000년도 의약품법으로 심화된 대체 조제에 대한 논란은 의사의 사전 동의 없이 대체조제 가능 품목을 식품의약품안전청장이 생물학적 동등성이 있다고 인정한 품목으로 제한함으로써 마무리되었다. 이에 발맞추어 식품의약품안전청(KFDA)은 후발 의약품에 대한 생물학적 동등성 시험 기준을 개정하였으며, 이 중에서 눈에 띄는 부분은 통계적인 고려 사항이다. 특히, 대조약과 시험약간의 동등성 판단을 위한 비교항목의 로그 변환을 명시한 것은 통계적 검증의 타당성 확보뿐만 아니라 국제적인 의약품 허가 추세에 부응한 바람직한 조치라 할 수 있다. 본 논문에서는 개정된 생물학적 동등성 시험 기준에서 고려된 AUC 와 C_{max} 의 로그 변환(log transformation)을 중심으로 그 배경과 통계적 의미를 고찰해 보고, 기존의 생물학적 동등성 판단 기준과 비교 검토해 보고자 한다. 또한 모의실험을 통해 개정된 기준의 통계적 타당성을 확인하고, 올바른 생물학적 동등성 판단을 위한 최소 표본수를 제안해 보기로 한다.

주요용어: 생체이용율, 생물학적 동등성 시험기준, 혈중농도-시간곡선하 면적(AUC), 로그변환

1. 개요

생물학적 동등성 시험 기준은 1988년 10월 28일 국립보건안전연구원 고시 제 1988-1호로 처음 제정되어, 1990년 1차개정, 1994년 2차 개정을 거쳐 최근 1998년 8월 26일에 공포된 제 7차 개정고시인 식품의약품안전청고시 제 1998-86호에 따라 생물학적 동등성 시험을 실시해 왔고, 제 8차 개정안이 2001년 2월 22일에 입안 예고되어 각계의 의견 수렴을 거친 후, 2001년 8월 식품의약품안전청 고시 제 2001-21호에 의해 개정 고시되었다. 개정사유는 약사법 제 26조 제 6항 및 동일 법규 시행규칙 제 23조 제1항 제1호 다목의 규정에 따라 제출되는 생물학적 동등성 시험에 대한 기준 및 평가방법 등을 국제화함으로써 동 시험의 신뢰성 확보를 통한 의약품 허가관리 등에 적정을 도모하고자 하는 것이다.

1) (461-701) 경기도 성남시 수정구 복정동 산65번지, 경원대학교, 응용통계학과, 조교수

E-mail: sgk@mail.kyungwon.ac.kr

2) (156-756) 서울시 은평구 녹번동 5, 식품의약품안전청, 국립독성연구소 일반 독성부 위해도평가과, 과장

E-mail: jisun@kfda.go.kr

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

E-mail: jikim@djju.ac.kr

4) (156-756) 서울시 동작구 흑석동, 중앙대학교 통계연구소, 선임연구원

E-mail: ala88@nate.com

개정의 주요 골자는 원료약품의 분량이 동일하고 첨가제가 유효성분의 흡수에 영향을 미치지 않는 주사제, 점안제, 점이제 등의 경우 생물학적동등성 시험에서 제외하는 것을 원칙으로 하고, 동일 제조업자가 생물학적 동등성을 승인 받은 품목과 제형은 동일하나 유효성분의 함량이 다른 품목으로 첨가제들간의 함량비율이 동일하다면 기승인 품목의 생체이용률 입증자료 및 비교용출시험자료로 생물학적 동등성을 판단할 수 있도록 하였으며, 피험자의 인권보호를 위하여 생물학적 동등성 시험계획서에 피험자 보상 규약내용을 추가하였고, 심사위원회의 구성을 강화하였다. 또한 생물학적 동등성 시험 실시기관을 확대하고, 피험자의 선정기준과 건강진단을 현실화 하고자 하며, 국제 기준과의 조화를 위하여 세계적으로 추천하는 미국 FDA(1992), 일본후생성(1997), 유럽연합의 CPMP(1999) 평가방법을 도입하여 생동성 모수 중 하나인 혈중농도-시간곡선하면적(AUC)과 최고혈중농도(C_{max})의 평균치 표기방식을 기존의 기준치의 $\pm 20\%$ 이내에서 평균비가 $[\log 0.8, \log 1.25]$ 이내로 개정하고, 시험자 및 시험의뢰자의 생물학적 동등성 시험 기준 미준수시 보고 의무 조항을 신설하고자 하였다. 본 논문에서는 이 중에서 AUC 와 C_{max} 의 평가 방식의 통계적 타당성을 논의해 보고자 한다.

제 2 절에서는 중요한 생동성 모수들과 이들의 계산 방법을 알아보고, 제 3 절에서는 개정된 규정하에서의 모형과 기존모형을 비교하여 설명해 보기로 한다. 제 4 절에서는 제 3 절에서 논의한 내용을 중심으로 통계적 타당성에 대한 모의실험 결과를 제시하고, 제 5 절에서는 결론 및 보언을 논의한다.

2. 생체이용률 모수

생물학적 동등성 시험이란 제제학적으로 동등한 두 제제의 생체이용률이 동등한 것인가를 입증하기 위한 생체내 시험을 말하며, 이는 다음의 기본적인 생물학적 동등성 가정하에서 진행된다.(FDA 21 CFR, 320.50(1989))

”두 제제가 동일한 투여 경로로, 활성 성분의 흡수되는 양과 속도에서 동일한 경우 이들은 치료학적으로 동등할 것이라고 가정한다.”

즉, 생물학적 동등성 시험이란 약제학적으로 동등한 두 제제의 치료학적 동등성을 확인하기 위해 생물학적 동등성을 통계적으로 확인하는 시험이라 할 수 있으며, 두 제제 중 신약으로 허가되어 안전성과 유효성이 확인된 제제를 대조약(reference drug) 그리고 생물학적 동등성 시험의 대상이 되는 제제를 시험약(test drug)이라 한다.

의약품의 약물동태학적 성질을 나타내는 중요지표로는 흡수율과 흡수량 외에 배설반감기, 배설율 및 대사율이 있으나 실제 생체이용률(bioavailability; BA)의 중요지표는 혈중농도-시간곡선하면적(AUC) 및 최고혈중농도(C_{max}) 그리고 최고혈중농도 도달시간(t_{max})을 들 수 있다. 생체이용률은 동일성분을 대상으로 실험하기 때문에 동일한 성분의 약물이 일단 흡수되면 동일한 배설반감기 및 대사율을 나타낼 것으로 추정할 수 있으므로, 흡수율 및 흡수량을 나타내는 지표인 AUC , C_{max} 및 t_{max} 등이 생체이용률 정도에 중요 모수(parameter)가 된다. 특히, AUC 는 약물의 생체이용율을 반영하며 전신순환에 도달하는 미변화 약물의 총량을 측정하는 유용한 모수로서, 투여시점부터 무한대 시간까지

약물혈중농도와 시간곡선의 총 면적으로 혈액에 도달하는 미변화 약물의 총량을 클리어런스(Clearance)로 나눈 것으로 (2.1)과 같이 정의되고, 이는 $[AUC]_0^{t_k}$ 를 근거로 계산된다.(Rowland and Tozer(1980))

$$[AUC]_0^\infty = \frac{FD_0}{Clearance} = \frac{FD_0}{KV_d} \quad (2.1)$$

여기서, F =약물흡수비율, D_0 =투여용량, K =소실속도정수, V_d =분포면적.

AUC 는 투여경로와 약물의 소실과정과는 무관하며, 적분법, 사다리꼴 공식법, 직접적인 면적측정기 사용 등에 의해 구할 수 있다.(Chow and Liu(1992)) AUC 의 단위는 농도·시간인 $\mu\text{g}/\text{ml}$ 을 주로 사용하며, 많은 약물에서 AUC 는 투여용량에 직접 비례한다. 예를 들어, 약물의 단위 투여량이 250mg에서 1000mg으로 증가하면 AUC 는 4배 증가한다고 할 수 있다.

생물학적 동등성 시험시 혈액의 약물분석을 위한 검체의 채취는 생체이용율의 평가에 필요한 시험치를 모두 산출할 수 있는 충분한 빈도를 갖는 적절한 시점이어야 한다. 또한, 시험약과 대조약의 검체채취 빈도 및 시간은 동일해야 한다. 채혈은 반감기의 3배 이상 또는 일정시간까지의 혈중농도-시간곡선하면적(AUC_t)이 무한시간까지의 혈중농도-시간곡선하면적(AUC_∞)의 최소 80%이상에 해당되는 충분한 시간동안 실시하여야 하며, 채혈횟수는 원칙적으로 12회 이상으로 하며, 최고혈중농도 도달전 적어도 2회 이상 채혈하여야 한다. 채혈 후 의약품의 유효성분을 약물의 실제농도를 정확하게 측정할 수 있는 충분한 감도를 가진 방법으로 분석하여 시간에 따른 약물농도 곡선을 그려 투여시점부터 무한대 시간까지 약물혈중농도와 시간곡선의 총면적을 계산하여 AUC 를 구하고, 실제로 측정된 관측치로부터 C_{max} 를 정한다. 여기서 C_{max} 란 약물을 경구 투여한 후 최고치에 도달한 혈중 약물 농도를 나타내며, t_{max} 는 C_{max} 에 도달한 시간을 말한다. 많은 약물에 있어서 혈중 약물 농도와 약효간에 상관관계가 있는데, C_{max} 는 약물이 치료효과를 나타낼 만큼 충분히 전신적으로 흡수되었는지를 나타내는 지표이다. 또한 C_{max} 는 약물의 부작용 발현 농도에 대한 간접적 지표가 되기도 하며, 그 단위는 농도단위로 $\mu\text{g}/\text{ml}$ 또는 mg/ml 로 나타낸다.

생물학적 동등성 시험 기준에 의하면, 시험 방법으로 충분한 휴약 기간(washout period)을 포함하여 표2.1과 같은 표준적인 2×2 교차 설계를 원칙으로 하고, 이를 통해 구해진 AUC 및 C_{max} 에 대한 생물학적 동등성 평가 기준을 제시하고 있으며, t_{max} 및 비교용출시험 자료를 참고 파라미터로 명시하고 있다. 따라서 본 논문에서는 생체이용률 모수 중에서 AUC 를 중심으로 설명해 보도록 한다.

표 2.1: 표준 2×2 교차설계

		기 간	
		1 휴약기간	2
순 서	1	R	T
	2	T	R

분석 모형으로는 각 대상(subject)간/내에서 정규성(normality)을 가정하고, 대상내 변동은 대상과 대상 그리고 제제와 제제에서 서로 동일함을 가정한다. 즉, j 번째 순서와 k 번째 기간, 그리고 i 번째 실험 대상자들로부터 측정되는 생체이용률 모수의 측정치를 y_{ijk} 라고 할 때, 두 제제 또는 제제 형식간의 생물학적 동등성 검증을 위한 2×2 교차설계의 모형을 정의하면 (2.2)과 같다. 이에 대한 분석 및 해석 방법은 S.G Ko and H.S. Oh(1999)을 참조하기 바란다.

$$y_{ijk} = \mu + G_k + S_{ik} + P_j + F_{(j,k)} + C_{(j-1,k)} + \varepsilon_{ijk}, \quad (2.2)$$

$$S_{ik} \sim iid N(0, \sigma_s^2), \quad \varepsilon_{ijk} \sim iid N(0, \sigma_e^2), \quad i = 1, 2, \dots, n_k, \quad j = 1, 2, \quad k = 1, 2.$$

여기에서 μ 는 전체평균을 나타내고, 순서효과(G), 기간효과(P), 제제효과(F), 잔류효과(C)는 (2.3)을 만족하고, 다음과 같은 값을 갖는다.

$$\begin{aligned} \sum_k G_k &= \sum_j P_j = \sum_{j,k} F_{(j,k)} = \sum_{j,k} C_{(j-1,k)} = 0 & (2.3) \\ F_{(j,k)} &= \begin{cases} F_R, & (j,k) = (1,1), (2,2) \\ F_T, & (j,k) = (1,2), (2,1), \end{cases} \\ C_{(j-1,k)} &= \begin{cases} C_R, & (j,k) = (2,1) \\ C_T, & (j,k) = (2,2) \\ 0, & elsewhere \end{cases} \end{aligned}$$

3. 개정된 생동성 기준하에서의 AUC 비교 분석

모형 (2.2)에서는 반응치가 정규분포를 따르고 등분산성을 가정한다. 그러나 생물학적 동등성 검증을 위해 측정되는 대부분의 반응치(e.g., AUC 또는 C_{max})는 일반적으로 분포가 양의 방향으로 치우쳐 있어 정규성에 위배되고, 분산의 동일성 역시 만족하지 않는 경우가 많으며, 경우에 따라서는 이상점이 존재하기도 한다.(Metzler and Huang(1983)) [그림3.1]과 [그림3.2]은 이에 대한 예로서 국내에서 실시된 로바스타틴정에 대한 생물학적 동등성 시험에서 동일한 제제 투여량에 노출된 피험자들의 AUC와 C_{max} 에 대한 히스토그램으로 양의 치우침을 가지고 있음을 확인할 수 있다.

기본적 가정의 위배와 이상점의 존재는 결국 통계적 의사 결정의 타당성이 의심받게 되어, 이에 적합한 분석 방법이 요구된다. 이 중 가장 일반적인 통계적 교정 방법은 변수 변환이며, 특히 로그 변환이 정규성과 등분산성을 상대적으로 많이 교정한다는 사실을 근거로 세계 각국은 생물학적 동등성 관련 지침에서 로그 변환, 일반적으로, 자연로그(natural logarithm) 변환을 추천하고 있다. 생체이용률 측정치의 로그 변환의 의미는 원자료의 모형이 (3.1)과 같은 승법 모형(multiplicative model)임을 말하며, (3.2)와 같은 대수 변환 모형(log-transformed model) 또는 대수정규 선형 모형(lognormal linear model)으로도 표현된

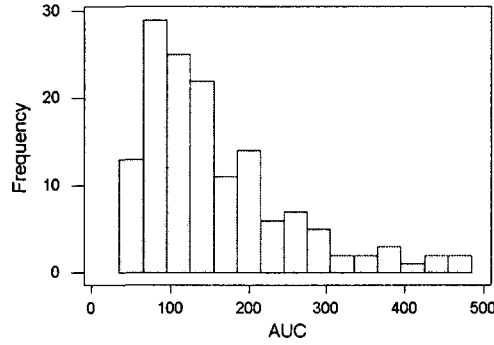


그림 3.1: 로바스타틴정을 투여받은 피험자의 AUC 히스토그램

다. 이때 y_{ijk} 는 생체이용률 측정치이며, (3.1)에서 소개된 효과들은 (2.3)에서 가정된 것과 동일하다.

$$y_{ijk} = \mu \cdot G_k \cdot S_{ik} \cdot P_j \cdot F_{(j,k)} \cdot C_{(j-1,k)} \cdot \varepsilon_{ijk}, \quad (3.1)$$

$$S_{ik} \sim iid N(0, \sigma_s^2), \varepsilon_{ijk} \sim iid N(0, \sigma_t^2), i = 1, 2, \dots, n_k, j = 1, 2, k = 1, 2, t = R, T$$

$$y_{ijk}^* = \ln(y_{ijk}) = \mu^* + G_k^* + S_{ik}^* + P_j^* + F_{(j,k)}^* + C_{(j-1,k)}^* + \varepsilon_{ijk}^*, \quad (3.2)$$

또는

$$y_{ijk} = \exp[\mu^* + G_k^* + S_{ik}^* + P_j^* + F_{(j,k)}^* + C_{(j-1,k)}^* + \varepsilon_{ijk}^*] \quad (3.3)$$

여기서,

$$\ln(S_{ik}^*) \sim iid N(0, \sigma_s^2), \ln(\varepsilon_{ijk}^*) \sim iid N(0, \sigma_t^2), i = 1, 2, \dots, n_k, j = 1, 2, k = 1, 2, t = R, T.$$

여기서 생체이용률 특성치의 모형 (3.1)과 로그변환된 모형 (3.2) 또는 (3.3)의 각 효과 모수와 변량은 (3.4)와 같은 관계에 있으며, 가법 모형화된 (3.2)에서 대상내의 랜덤 효과인 S_{ik}^* 와 대상내 오차를 나타내는 ε_{ijk}^* 는 대수정규분포를 따르고, 이들의 기대값은 (3.5)와 같다.

$$\mu = \exp(\mu^*), G = \exp(G^*), P = \exp(P^*), F = \exp(F^*), C = \exp(C^*), \quad (3.4)$$

$$E(S_{ik}^*) = \frac{1}{2}\sigma_s^2, \quad E(\varepsilon_{ijk}^*) = \begin{cases} \frac{1}{2}\sigma_T^2, & (j, k) = (1, 2), (2, 1) \\ \frac{1}{2}\sigma_R^2, & (j, k) = (1, 1), (2, 2) \end{cases} \quad (3.5)$$

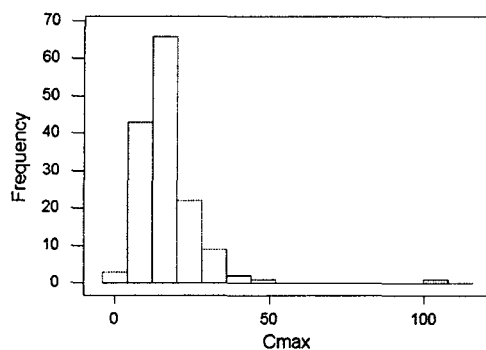


그림 3.2: 로바스타틴정을 투여받은 피험자의 C_{max} 히스토그램

즉, 로그 변환된 생체이용률 반응치의 기대값은 (3.3)을 이용하여 계산된 (3.6)식에서 알 수 있듯이, 기간효과(P)와 로그변환된 반응치의 대상내 변동(σ_s^2)과 교락될 수 있다.

$$E(y_{i11}) = \exp[\mu^* + P_1^* + F_R^* + \frac{1}{2}(\sigma_R^2 + \sigma_s^2)] \quad (3.6)$$

$$E(y_{i12}) = \exp[\mu^* + P_2^* + F_T^* + \frac{1}{2}(\sigma_T^2 + \sigma_s^2)]$$

$$E(y_{i21}) = \exp[\mu^* + P_1 + F_T^* + \frac{1}{2}(\sigma_T^2 + \sigma_s^2)]$$

$$E(y_{i22}) = \exp[\mu^* + P_2 + F_R^* + \frac{1}{2}(\sigma_R^2 + \sigma_s^2)]$$

그러나 로그 변환된 자료가 근사적으로 정규분포를 따르는 경우, 생동성 판단의 기준이 되는 제제 효과의 점추정값과 $100(1 - 2\alpha)\%$ 신뢰구간은 로그 변환된 자료, y_{ijk}^* 를 이용하여 다음과 같이 계산할 수 있다. (Chow and Liu(1992))

$\hat{F} = \bar{d}_{.1} - \bar{d}_{.2}$, 여기서, $\bar{d}_{.k}$ 는 기간의 차 $d_{ik} = \frac{1}{2}(y_{i2k}^* - y_{i1k}^*)$ 로부터 계산된 표본평균

또한 기간 비(period ratio) $r_{ik} = y_{i2k}/y_{i1k}$, $i = 1, 2, \dots, n_k$, $k = 1, 2$ 를 정의하면, $d_{ik} = \frac{1}{2} \log(r_{ik})$ 로 정의되고, 만일 잔류효과가 없다면, r_{ik} 는 (3.7)과 같이 표현할 수 있어 (3.1)의 승법 모형하에서의 대상내 변동을 없앨 수 있다. 이는 기존 생동성 모형에서 기간의 차를 이용하여 대상내 변동을 없애는 경우와 동일하게 해석할 수 있다.

$$r_{ik} = \frac{y_{i2k}}{y_{i1k}} = \begin{cases} \frac{P_2 F_{TRe_{i2k}}}{P_1 F_{Re_{i1k}}}, & k = 1 \\ \frac{P_2 F_{Re_{i2k}}}{P_1 F_{TRe_{i1k}}}, & k = 2 \end{cases} \quad (3.7)$$

또한 \hat{F} 는 $\mu_T^* - \mu_R^*$ 의 최대우도 추정량(MLE)이며, 지수 변환된 $\exp(\hat{F})$ 는 μ_T/μ_R 의 최대 우도 추정량이 된다. 여기서, 로그 변환된 자료의 평균은 원자료의 기간비와 (3.8)의 관계에 있다.

$$\exp(\hat{F}) = \frac{\hat{\mu}_T}{\hat{\mu}_R} = \left(\frac{R_1}{R_2} \right)^{1/2}, \quad R_k = \left(\prod_{i=1}^{n_k} r_{ik} \right)^{1/n_k}, \quad k = 1, 2 \quad (3.8)$$

즉, 로그 변환된 자료의 선형불편 추정량 \hat{F} 은 순서 1과 순서2를 이용한 기간 비의 기하 평균 (geometric mean)들의 비를 계산한 값의 제곱근을 구해 로그 변환한 값과 동일하다. 이를 근거로 μ_T/μ_R 의 $100(1 - 2\alpha)\%$ 신뢰구간이 $(\exp(L), \exp(U))$ 으로 주어지며, 여기서, (L, U) 는 로그 변환된 자료에서 구해진 $\mu_T^* - \mu_R^*$ 의 $100(1 - 2\alpha)\%$ 신뢰구간을 말하고, (3.9)과 같이 정의되며, 이를 동등구간과 비교하여 생물학적 동등성 여부를 확인한다. 이에 대한 구체적인 내용은 S.G. Ko and H.S. Oh (1999)에서 자세히 논의하였다.

$$(L, U) = \left(\hat{F} - t(\alpha, df) \hat{\sigma}_d \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}, \quad \hat{F} + t(\alpha, df) \hat{\sigma}_d \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \right), \quad (3.9)$$

여기서, $\hat{\sigma}_d$ 는 ANOVA의 $\frac{1}{2}MS_{intra}$.

여기서 $\exp(\hat{F})$ 는 모형 (3.1)에서 정의되는 모수 μ_T/μ_R 의 추정치로 기존의 생물학적 동등성 기준인 ± 20 규칙에 의하면, 동등 구간은 $[0.8, 1.2]$ 임이 타당하다. 그러나, \hat{F} 를 이용하여 ± 20 규칙을 적용하여 동등구간을 설정할 경우, $\ln(0.8) = -0.223$ 그리고 $\ln(1.2) = 0.182$ 가 되어, $\ln(1) = 0$ 에 대해 대칭적인 기준이 제시되지 않는다. 따라서, 동등구간의 상한을 $\ln(1.25) = 0.223$ 로 지정하여 동등구간을 $\ln(1) = 0$ 으로부터 대칭이 되도록 선정함이 보다 바람직하다.(Cartwright(1992))

지금까지 논의된 승법모형, 즉 로그 변환 모형의 생물학적 동등성 기준은 [표 3.1]의 생물학적 동등성 기준 고시(식품의약품안전청 고시 제2001-21호)에 해당한다. 여기서 유의해야 할 점은 다음의 기준은 로그 변환후 정규성을 근사적으로 만족한다는 전제하에서 제시된 것으로 로그 변환 후에도 정규성에 위배되거나, 실험의 비정밀성으로 인한 등분산성의 위배 또는 이상점이 존재한다면, 생물학적 동등성 여부를 정확히 판단할 수 없다.

4. 모의 실험

이 절에서는 좌우대칭과 중형의 형태를 갖는 확률난수와 양의 치우침을 갖는 확률난수를 이용하여 일정범위의 변동계수(Coefficient of Variation; 이하 CV)하에서 1998년 고시된 생물학적 동등성 시험 기준(OLD)과 2001년 고시된 개정된 생물학적 동등성 시험 기준(NEW)의 생물학적 동등성 채택 비율과 그 차이를 비교해 보기로 한다. 또한 개정된 생

표 3.1: 생물학적 동등성 기준 고시의 비교 : AUC 와 C_{max}

항목	식약청 고시 (1998-86호)	식약청 고시 (2001-21호)
자료변환	없음	로그변환
검 추정	시험약 \in (대조약 $\pm 20\%$)	$\bar{F} \in [\ln(0.8), \ln(1.25)]$
구간추정	대조약과 시험약의 비교항목 평균치의 차는 대조약의 20% 이내이어야 한다.	t_{max} 를 제외한 대조약과 시험약의 비교평가항목치를 로그변환하여 통계처리 하였을 때, 로그변환한 평균치 차의 90% 신뢰구간이 $\log 0.8$ 에서 $\log 1.25$ 이내 이어야 한다.
분산분석과 검출력	분산분석에 의한 검정은 원칙적으로 α (유의수준) = 0.05 ~ 0.1로 하고, 그때의 정도는 $1 - \beta$ (검출력) ≥ 0.8 및 Δ (최소 검출차) ≤ 0.2 로 함이 바람직하고 의약품의 종류에 따라 최소검출차의 의미를 고찰할 필요가 있다.	분산분석에 의한 검정은 원칙적으로 α (유의수준) = 0.05에서 실시한다. 검출력 및 최소검출차 삭제
기 타	또한, 두 제제의 생체이용률차의 신뢰한계를 구하여 분산분석의 결과와 합쳐 평가한다.	삭제

물학적 동등성 규정에서 삭제된 검출력 기준이 생물학적 동등성 판정비율에 미치는 영향을 확인해 보고, 특정 CV값을 갖는 생물학적 동등성 시험에서 적절한 표본 크기에 대한 기준을 제시해 보기로 한다.

이를 위해 본 모의실험에서는 지정된 생물학적 모수를 갖는 정규분포와 대수정규분포로부터 확률 난수를 3,000번 반복 추출하여 각각 OLD와 NEW 규정의 생물학적 동등성 채택비율을 비교하였다. 이때, 표본수는 각 제제군당 6명과 12명을 우선 고려하였고, 적절한 표본 크기에 대한 고찰에서는 각 제제군당 6명에서 18명까지 변화해 가며 그 결과를 고찰하였다. 또한 대조약의 평균은 100으로 표준화하였고, 시험약의 평균은 실제 생물학적 동등성을 만족하는 범위 내에서 10 단위로 가감하고, CV(%)를 10에서 40까지 10씩 증가시키며 그 결과를 비교해 보았다.

[표4.1]은 정규분포를 가정한 경우의 모의 실험 결과이며, [표4.2]는 대수정규분포를 가정한 경우의 모의실험 결과이다. 두 종류의 모의 실험에서 NEW의 생물학적 동등성 채택비율이 OLD에 비해서 상대적으로 높게 나타남을 확인할 수 있고, 제제간 평균 차이가 있는 경우에는 OLD에서 검출력 조항이 생물학적 동등성 판정에 큰 영향을 주었음을 확인할 수 있다.

두 경우 모두 대조약과 시험약의 표본평균의 차에 대한 90% 신뢰구간이 0을 포함할 비율은 두 규정 모두 거의 동일하며, 실제 두 제제의 평균 차가 클수록 그 비율은 작아지지만, 정해진 두 제제의 평균 하에서 CV(%)가 커질수록 그 비율은 커지는 것으로 관측되었다. 이는 통계적 구간 추정만으로 생물학적 동등성을 판정할 수 없다는 사실을 예증하는 것으로, 동등 구간을 이용한 생물학적 동등성의 타당성을 입증한다고 할 수 있다. 즉, 만일 통계

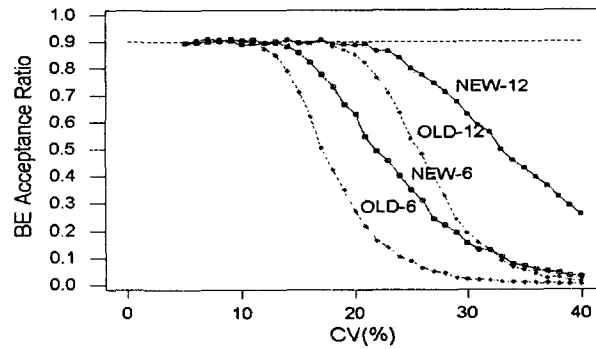


그림 4.1: CV값에 따른 OLD와 NEW의 생물학적 동등성 채택비율
:제제군당 표본수=6,12

적 구간추정 방법만으로 생물학적 동등성을 인정한다면, CV(%)가 커질수록 생물학적 동등성에 대한 채택율이 커진다는 오류를 범할 수 있게 된다. 동등구간을 이용한 생물학적 동등성 채택 비율을 보면, NEW가 OLD에 비해 높은 채택율을 보이고 있으며, 표본의 크기가 작아도 CV(%)가 작고 실제 두 제제간 평균차가 크지 않다면, 상당히 높은 채택율을 보이는 것을 알 수 있다. 또한 두 규정 모두 동등구간을 이용한 OLD의 채택 비율과 Power Approach에서의 채택율을 비교해 보면, 두 제제간 차이가 상대적으로 큰 경우 검출력 기준이 실제 채택 비율에 많은 영향을 주고 있으며, 이는 실제 두 제제의 평균의 차이를 적절하게 반영하고 있지 않은 것으로 나타났다. 이를 통해, NEW에서 검출력 기준의 생략은 타당했던 조치로 해석된다.

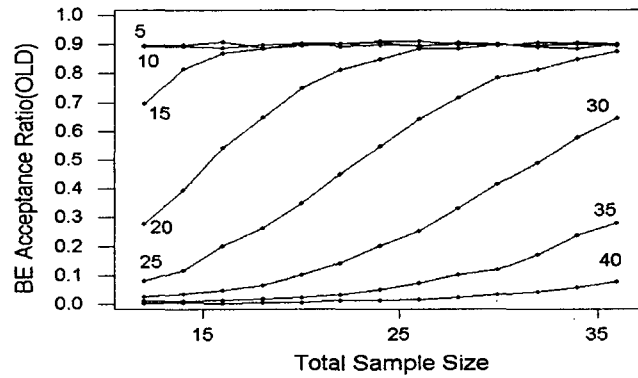
NEW의 경우, 고려된 두 제제 간의 차이가 ± 10 의 범위내에 있고 CV(%)가 10인 경우, 약 90%이상의 채택율을 보이고 있는 반면, OLD는 두 제제간 차이가 없고, CV(%)가 10인 경우에 한해 90%정도의 채택률을 보이고 있음을 확인할 수 있다. 그리고, OLD에서 검출력의 조건을 고려하지 않고 ± 20 동등구간만을 고려하는 경우, 각 제제군당 표본의 수가 큰 경우에는 OLD가 NEW보다 더 높은 채택률을 보이는 경우도 있으나 그 차이는 크지 않다.

[표4.1]에서는 NEW의 채택율에서 시험약의 평균 생체이용률이 대조약의 평균 생체이용율보다 큰 경우가 그렇지 않은 경우보다 훨씬 더 채택률이 높음을 확인할 수 있다. 예를 들어, 각 제제군당 표본수가 6명이고 CV(%)가 10인 경우는 $(\mu_R, \mu_T) = (100, 120)$ 과 $(\mu_R, \mu_T) = (100, 80)$ 의 생물학적 동등성 채택 비율은 각각 24.7%와 4.7%로 약 20%정도 차이가 있으며, 동일한 표본수에서 CV(%)가 40인 경우는 각각 1.5%와 0.2%이다. 이는 각 제제군당 표본수가 12명인 경우에도 비슷한 결과를 보이고 있다. 이는 AUC의 분포로 정규 분포를 가정되었고, NEW에서 로그 변환된 제제 별 평균비의 동등구간 한계를 $\ln(1.2)$ 에서 $\ln(1.25)$ 로 확장한 것에 기인하는 것으로 해석할 수 있으며, 이러한 현상은 AUC의 분포로

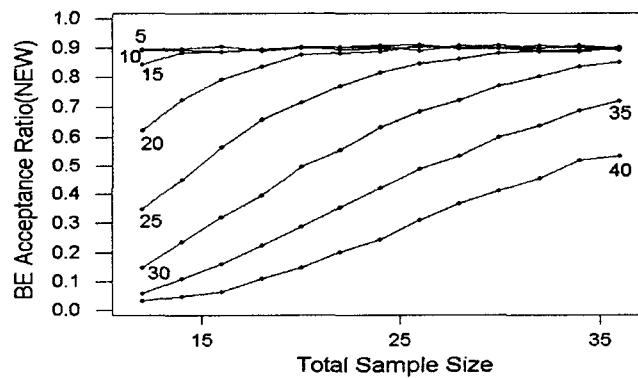
대수 정규분포를 가정한 경우 현격히 줄어듦을 [표4.2]를 통해 확인할 수 있다.

[그림4.1]은 대조약과 시험약이 동일한 모수를 갖는 대수정규분포를 가정한 경우, $CV(\%)$ 값의 변화에 따른 OLD와 NEW의 생물학적 동등성 채택 비율을 비교한 것이다. 제제군당 표본수가 6인 경우는 $CV(\%)$ 가 약 10정도까지는 두 규정 모두 대략 90%의 채택비율로서 비슷하지만, $CV(\%)$ 가 10이상 커지는 경우에는 채택 비율이 급격히 작아지고, 그 정도는 NEW에 비해 OLD가 큰 것으로 나타났다. 또한 제제군 당 표본수가 12인 경우 역시 비슷한 결론을 내릴 수 있으나, 채택 비율이 떨어지는 지점이 $CV(\%)$ 가 약 20 정도임을 확인할 수 있다. 따라서, $CV(\%)$ 가 큰 경우에는 좀 더 많은 표본 수를 고려해야 함을 확인할 수 있고, 동일한 표본수의 경우에는 OLD가 NEW에 비해 채택 비율이 더 많이 떨어짐을 확인할 수 있다.

그림 4.2: $CV(\%)$ 값과 전체 표본수에 따른 생물학적 동등성 채택 비율



(a) OLD의 생물학적 동등성 채택비율: $CV(\%) = 5(5)40$



(b) NEW의 생물학적 동등성 채택비율: $CV(\%) = 5(5)40$

[그림4.2]의 (a)와 (b)는 서로 다른 CV 값과 표본 크기에 따른 생물학적 동등성 채택 비율을 보여준다. 이 경우 역시 대조약과 시험약이 동일한 모수를 갖는 대수정규분포를 가정하

표 4.1: 기존 기준(OLD)과 개정된 기준(NEW)의 생물학적 동등성 채택비율: AUC의 분포가 좌우대칭, 중형인 경우(정규분포)

sample size	R	T	CV (%)	OLD			NEW		
				C.I(0) ^{a)}	C.I(±20) ^{b)}	PA ^{c)}	C.I(0) ^{d)}	C.I*(±20) ^{e)}	PA ^{f)}
6	100	80	10	0.002	0.038	0.002	0.002	0.047	0.002
			20	0.260	0.033	0.041	0.268	0.042	0.047
			30	0.581	0.017	0.007	0.598	0.018	0.010
			40	0.710	0.001	0.000	0.726	0.002	0.000
6	100	90	10	0.283	0.757	0.282	0.286	0.799	0.286
			20	0.692	0.251	0.209	0.701	0.293	0.248
			30	0.798	0.046	0.023	0.810	0.058	0.026
			40	0.842	0.005	0.002	0.855	0.006	0.001
6	100	100	10	0.898	0.993	0.898	0.899	0.999	0.899
			20	0.901	0.483	0.332	0.899	0.565	0.433
			30	0.896	0.089	0.035	0.899	0.101	0.048
			40	0.899	0.012	0.002	0.902	0.013	0.004
6	100	110	10	0.242	0.711	0.241	0.241	0.901	0.241
			20	0.701	0.294	0.249	0.700	0.397	0.359
			30	0.793	0.074	0.027	0.801	0.101	0.049
			40	0.841	0.015	0.005	0.856	0.022	0.006
6	100	120	10	0.002	0.067	0.002	0.001	0.247	0.001
			20	0.273	0.066	0.080	0.279	0.113	0.133
			30	0.561	0.044	0.022	0.577	0.059	0.040
			40	0.680	0.013	0.002	0.706	0.015	0.005
12	100	80	10	0.000	0.038	0.000	0.000	0.044	0.000
			20	0.046	0.038	0.042	0.048	0.043	0.038
			30	0.274	0.034	0.033	0.309	0.026	0.024
			40	0.530	0.020	0.005	0.567	0.012	0.002
12	100	90	10	0.042	0.968	0.042	0.045	0.976	0.045
			20	0.483	0.512	0.474	0.497	0.516	0.481
			30	0.713	0.229	0.145	0.713	0.203	0.118
			40	0.791	0.067	0.009	0.808	0.038	0.007
12	100	100	10	0.900	1.000	0.900	0.900	1.000	0.900
			20	0.893	0.899	0.879	0.895	0.927	0.884
			30	0.900	0.448	0.236	0.903	0.411	0.256
			40	0.904	0.121	0.016	0.902	0.077	0.009
12	100	110	10	0.040	0.938	0.040	0.040	0.996	0.040
			20	0.493	0.515	0.486	0.495	0.672	0.492
			30	0.709	0.284	0.190	0.720	0.304	0.259
			40	0.778	0.105	0.020	0.796	0.076	0.016
12	100	120	10	0.000	0.071	0.000	0.000	0.422	0.000
			20	0.042	0.068	0.042	0.045	0.167	0.045
			30	0.272	0.067	0.078	0.295	0.092	0.119
			40	0.506	0.050	0.010	0.535	0.041	0.015

a) confidence interval including 0
 b) confidence interval within (-20, 20)
 c) Power Approach
 d) confidence interval including 0
 e) confidence interval within (ln 0.8, ln 1.25)
 f) Power Approach

표 4.2: 기존 기준(OLD)과 개정된 기준(NEW)의 생물학적 동등성 채택비율: AUC의 분포가 양의 치우침을 갖는 경우(대수정규분포)

sample size	R	T	CV (%)	OLD			NEW		
				C.I.(0) ^{a)}	C.I.(±20) ^{b)}	PA ^{c)}	C.I.(0) ^{d)}	C.I*(±20) ^{e)}	PA ^{f)}
6	100	80	10	0.000	0.035	0.000	0.000	0.048	0.000
			20	0.177	0.037	0.044	0.172	0.049	0.061
			30	0.465	0.020	0.013	0.458	0.031	0.020
			40	0.645	0.005	0.002	0.625	0.010	0.002
6	100	90	10	0.225	0.803	0.225	0.222	0.852	0.222
			20	0.685	0.294	0.265	0.679	0.383	0.365
			30	0.795	0.064	0.024	0.788	0.111	0.054
			40	0.835	0.011	0.006	0.826	0.025	0.009
6	100	100	10	0.900	0.994	0.900	0.904	0.999	0.904
			20	0.900	0.485	0.341	0.904	0.635	0.519
			30	0.906	0.080	0.030	0.908	0.158	0.069
			40	0.906	0.017	0.004	0.911	0.032	0.008
6	100	110	10	0.292	0.672	0.292	0.286	0.888	0.286
			20	0.702	0.285	0.218	0.702	0.417	0.382
			30	0.824	0.063	0.022	0.817	0.115	0.053
			40	0.842	0.014	0.005	0.837	0.030	0.008
6	100	120	10	0.006	0.064	0.006	0.006	0.230	0.006
			20	0.334	0.060	0.063	0.335	0.120	0.145
			30	0.583	0.032	0.012	0.579	0.060	0.033
			40	0.724	0.005	0.001	0.705	0.013	0.005
12	100	80	10	0.000	0.034	0.000	0.000	0.048	0.000
			20	0.021	0.038	0.020	0.016	0.054	0.016
			30	0.200	0.038	0.057	0.188	0.052	0.083
			40	0.403	0.021	0.011	0.381	0.038	0.021
12	100	90	10	0.029	0.978	0.029	0.028	0.985	0.028
			20	0.447	0.544	0.442	0.436	0.625	0.436
			30	0.680	0.274	0.219	0.673	0.365	0.371
			40	0.779	0.080	0.016	0.774	0.171	0.045
12	100	100	10	0.900	1.000	0.900	0.901	1.000	0.901
			20	0.907	0.917	0.894	0.906	0.971	0.905
			30	0.907	0.462	0.266	0.904	0.638	0.530
			40	0.902	0.109	0.015	0.903	0.245	0.062
12	100	110	10	0.063	0.927	0.063	0.060	0.996	0.060
			20	0.526	0.493	0.502	0.516	0.709	0.516
			30	0.718	0.262	0.155	0.711	0.411	0.394
			40	0.782	0.074	0.015	0.780	0.170	0.045
12	100	120	10	0.000	0.057	0.000	0.000	0.395	0.000
			20	0.080	0.065	0.072	0.080	0.185	0.080
			30	0.343	0.063	0.052	0.326	0.119	0.162
			40	0.551	0.035	0.006	0.532	0.074	0.029

- a) confidence interval including 0
b) confidence interval within (-20, 20)
c) Power Approach
d) confidence interval including 0
e) confidence interval within (ln 0.8, ln 1.25)
f) Power Approach

였으며, 3,000번의 모의 실험을 통해 채택 비율을 계산하였다. CV값은 5에서 40까지를 고려하였으며, 전체 표본수는 현실적인 기준인 12명(제제군당 6명)에서 36명(제제군당 18명)까지를 고려하였다.

동일한 표본수를 가정한다면, CV(%)가 작은 경우에는 두 규정이 거의 동일한 채택비율을 보이지만, CV(%)가 커짐에 따라 두 규정의 채택비율에서 큰 차이를 확인할 수 있다. 다시 말해 동일한 CV(%)하에서 동일한 채택비율을 얻기 위해 필요되는 표본수가 다를 수 있다. 이를 근거로 생물학적 동등성 채택 비율에 따른 각 제제군당 최소 표본수를 정리해보면 [표4.3]과 같다. 이는 고려하고 있는 생동성 시험 자료의 CV값을 대략 예상할 수 있는 경우, 표본 크기에 따른 가능한 채택 비율의 기준을 제시한다. 전체적으로 동일한 채택비율을 가정하는 경우, CV(%)값이 커지면서 제제 군당 표본수는 커지고, NEW가 OLD에 비해 요구되는 표본의 수가 적음을 확인할 수 있다. 또한 CV(%)가 5인 경우, 채택비율을 90%이상 달성하는 제제군 당 최소 표본수는 두 경우 모두 6명으로, 이는 미국, 캐나다 등의 생물학적 동등성 시험 기준에서의 최소 피험자수와 일치한다. 그러나 CV(%)가 35인 경우에는 최대 표본수인 제제군당 18명을 고려하더라도 OLD의 채택비율은 50% 미만이며, NEW의 경우에도 약 70%정도의 채택 비율만을 달성하며, CV(%)가 40인 경우에는 두 규정 모두 주어진 표본 수에서 60%의 채택비율을 달성하지 못한다. 즉, 피험자의 수는 통계적 또는 생물학적 동등성 기준에 의해 제시되는 것이 아니라, 생물학적 동등성 시험의 설계에서 올바른 피험자의 선택 그리고 생체 모수의 측정 정도에 따라 결정되는 CV(%)에 따라 결정됨을 확인할 수 있다.

표 4.3: 생물학적 동등성 채택 비율(%)과 CV(%)에 따른 제제군당 최소 표본 수

CV (%)	BE 채택 비율(%)	50	60	70	80	90
5	OLD	* ^{a)}	*	*	*	6
	NEW	*	*	*	*	6
10	OLD	*	*	*	*	9
	NEW	*	*	*	*	9
15	OLD	*	*	6	7	10
	NEW	*	*	*	6	9
20	OLD	8	9	10	11	15
	NEW	*	6	7	9	13
25	OLD	12	13	14	16	*
	NEW	*	9	10	12	18
30	OLD	17	18	*	*	*
	NEW	10	12	14	16	*
35	OLD	*	*	*	*	*
	NEW	14	15	18	*	*
40	OLD	*	*	*	*	*
	NEW	17	*	*	*	*

a) 고려한 제제군당 표본수(6 ~ 18명)에서 달성 불가능한 경우

5. 결론 및 보언

생물학적 동등성 시험은 인체를 대상으로 실시된다. 따라서 무엇보다 윤리적 측면이 강조되어야 하며, 제제학적으로 동등한 의약품은 생물학적으로 동등한지를 평가하기 위해 여러 생체 검체 분석기술, 약동력학적 기초, 그리고 실제 데이터를 통한 올바른 판단 기준이 되는 통계적 실험설계와 분석기법들이 잘 적용되도록 하여야 한다. 또한 생물학적 동등성을 만족하는 제제에 대한 올바른 판단이 매우 중요시되어야 하지만, 생물학적 동등성을 만족하지 않는 제제에 대한 올바른 기각 역시 우선 고려될 수 있어야 하며, 그 개정의 목표가 국내적으로는 의약품 산업에 따른 대체조제 요구에 부응하고 국제적으로는 각국의 약효 동등성 관리제도의 합리적 운영과 조화를 이룰 수 있어야 한다. 이러한 점에서 2001년 개정된 생물학적 동등성 규정은 세계적 추세에 부응하여, 각국에서 반영하고 있는 복제의약품의 생산 및 유통에 대한 생물학적 동등성 시험 기준을 잘 보완하여 개정되었다고 평가되며, 본 논문에서 모의실험을 통해 확인해 본 바와 같이 경험적 그리고 통계적 기준을 잘 반영하였다고 보여진다. 그러나 현재 고려하고 있는 생체이용률 모수의 평균뿐 아니라 분산의 영향을 반영할 수 있는 방법 등이 반영되지 않아 아쉬움을 남긴다. 앞으로도 의약품과 국민건강보험 재정 안정을 위해 생물학적 동등성 시험 수요가 급증하고 있는 만큼, 생물학적 동등성 시험의 설계와 데이터 수집 그리고 분석에 이르는 다양한 통계적 고찰과 보완이 꾸준히 이루어져야 하겠다.

참고문헌

- [1] 식품의약품안전청(1998), 고시 제 1998-51 호, 생물학적 동등성 시험 기준, 제 5 장, 제 13조.
- [2] 식품의약품안전청(2001), 고시 제 2001-21 호, 생물학적 동등성 시험 기준, 제 5 장, 제 19조.
- [3] Cartwright. (1992). International harmonization and consensus DIA meetings on bioavailability and bioequivalence requirements and standards, *Drug Information Journal*, 25, 471-482.
- [4] Chow, S.C. and Liu, J.P. (1992). *Design and Analysis of Bioavailability and Bioequivalence Studies*, Marcel Dekker, New York.
- [5] Commission of the European Communities. (1999). CPMP(committee for Proprietary Medicinal Products) Working Party in the Efficacy of Medicinal Products Note for Guidance, Investigation of Bioavailability and Bioequivalence, Brussels, Belgium.
- [6] FDA. (1989). Federal Register, Vol. 42, No. 5, Sects. 320.05.
- [7] Food and Drug Administration(FDA). (1992). Guidance on Statistical Procedures for

Bioequivalence Studies using Standard Two-treatment Crossover Design. Division of Bioequivalence, Office of Generic Drugs, Rockville, MD.

- [8] Liu, J.P and Chow, S.C. (1992). On the assessment of variability in bioavailability/bioequivalence studies, *Comm. Stat. - Theory and Meth.* 21(9), 2591-2607.
- [9] Metzler, C.M. and Huang, D.C. (1983). Statistical methods for bioavailability and bioequivalence, *Clin. Res. Practice and Drug Reg. Affairs*, 1, 109 - 132.
- [10] Rowland, M. and Tozer, T.N. (1980). *Clinical Pharmacokinetics Concepts and Applications*. Lea & Febiger, Philadelphia, PA.
- [11] S.G. Ko, H.S. Oh. (1999). Assessment and Interpretation of Bioequivalence for Two Drug Formulations using Crossover Design, *The Korean Communication in Statistics*, Vol, 241-251.
- [12] The National Institute of Health Sciences. (1997). Guideline for Bioequivalence Studies of Generic products, Tokyo, Japan.

[2001년 12월 접수, 2002년 3월 채택]

Statistical consideration of assessing bioequivalence in the new KFDA regulation

S. G. Ko ¹⁾ J. S. Yang ²⁾ J. I. Kim ³⁾ N. K. Lim ⁴⁾

ABSTRACT

On the new medical system separating the prescription and dispensing of the drug, the qualification of pharmacist in substitution of prescribed medicine was restricted, except bioequivalence-certified drugs. Also, Korean Food and Drug Association(KFDA) revised the bioequivalence regulation on August, 2001. Among many changes from old guideline, impressive one is the statistical consideration. Specially, to estimate and analyze bioequivalence measures, AUC and C_{max} , the log-transformed model is recommended and the equivalence interval is modified from ± 20 rule to $[\ln(0.8), \ln(1.25)]$ one. This meaningful act is very hope-for because it is statistically reasonable and is agreed with worldwide bioequivalence guideline, including USA, EU, Japan and Canada. In this paper, we introduce the new regulation of assessing bioequivalence, announced at August, 2001, mainly on statistical view points. Key points for the new regulation are discussed and the minimum sample size based on simulation studies are proposed.

Keywords: Bioavailability, Bioequivalence Regulation, Area under the plasma or blood concentration-time curve(AUC), log transformation

1) Assistant Professor, Dept. of Applied Statistics, Kyungwon Univ. E-mail:sgk@mail.kyungwon.ac.kr

2) Chief, Risk Assessment Division, General Toxicology Department, National Institute of Toxicological Research, Korea Food and Drug Administration E-mail:jisun@kfda.go.kr

3) Professor, Dept. of Information Statistics, Daejeon Univ. E-mail:jikim@dju.ac.kr

4) Research Assistant, Statistical Consulting Center, Chungang Univ. E-mail:ala88@nate.com