The Korean Journal of Applied Statistics (2011) DOI: http://dx.doi.org/10.5351/KJAS.2011.24.6.1095
24*(6), 1095–1102*

Bayesian Estimation in Bioequivalence Study

Seung-Chun Lee[1]

[1]Department of Statistics, Hanshin University

(Received October 2011; accepted December 2011)

Abstract

The classical two-period, two-sequence crossover design is no longer sufficient to assess various demands in a bioequivalence study. For instance, to estimate the within-subject and between-subject variances of test and reference formulations separately, it is necessary to use a replicate design in which each subject receives at least the reference formulation in two periods. Several designs were studied to satisfy the demands. It is provided a unified Bayesian approach applicable to those study designs. The benefit of the method in the bioequivalence study is discussed.

Keywords: Gibbs sampling, crossover design, bioequivalence, linear mixed model.

1. Introduction

Bioequivalence studies are generally conducted through a comparison of the *in vivo* rate and extent of drug absorption of a test drug and an innovative drug. In a standard *in vivo* bioequivalence study, a conventional two-period, two-sequence randomized crossover design has been widely used. In this 2×2 crossover design, study subjects receive a single dose of both test and innovative drugs on separate occasions through random assignment to the two possible sequences of drug administration. The 2×2 crossover design enables to assess various statistical objectives as well as the average bioequivalence(ABE). See Hauschke *et al.* (2007) for further detail.

In 1997, the US Food and Drug Administration(FDA) published a draft guidance to propose two new concepts of bioequivalence: population bioequivalence(PBE) introduced by Anderson and Hauck (1990) and individual bioequivalence(IBE) introduced by Anderson and Hauck (1990). This draft guidance recommended that the PBE or IBE should be used to assess bioequivalence instead of the ABE. This recommendation caused a series of controversial debate, which ended up by omitting the PBE and IBE concepts from the FDA guidance in 2003. However, many statistical methods for PBE and IBE were proposed during the debate (Hyslop *et al.*, 2000; Carrasco and Jover, 2003; Chow *et al.*, 2003; Hsuan and Reeve, 2003; Oh *et al.*, 2003).

These two criteria compare the population means of the reference and test formulation as well as the variances. In particular, the IBE criterion requires the estimation of the variance of subject-by-formulation interaction. Since we cannot estimate the within-subject and between-subject variances

This work was supported by Hanshin University research grant.
[1]Professor, Department of Statistics, Hanshin University, 411 Yangsan-Dong, Osan, Kyunggi-Do 447-791, Korea. E-mail: seung@hs.ac.kr

of each test and innovative reference formulation separately under the 2×2 design, it is not sufficient to assess the IBE based on the criterion recommended by the Food and Drug Administration (1997). For the assessment, it is necessary to use a replicate design where each subject receives the reference formulation at least two periods. The FDA recommended a four-period, two-sequence, randomized crossover design. However, some authors studied other designs instead of the recommended design. For example, Shao *et al.* (2002) considered the 2×3 design, while Jung *et al.* (2010) investigated the 3×2 design to assess the IBE.

The current FDA bioequivalence guidance does not require a demonstration of the PBE or IBE. However, even if we assume the normality of pharmacokinetic characteristic, it is not appropriate that the decision on bioequivalence of the two formulations is solely based on a comparison of the means. We should somehow investigate the between-subject variance or subject-by-formulation interactions. For this purpose, the classical two-period, two-sequence crossover design could no longer be a standard study design. In addition, there are situations in which more than two formulations are needed. For instance, dose linearity study typically needs to test more than two formulations. Thus, to achieve various objectives in bioequivalence studies, we should consider a general crossover design. Note that the design can be expressed as a linear mixed model and the Bayesian theory is well-established in the linear mixed model with significant literature published on the Bayesian treatment for the model. For example, Lee and Lee (2002) provided Bayesian analysis in a quite general multivariate linear mixed model using Gibbs sampling. Note that the Bayesian approach has several advantages in that the posterior distribution of model parameters can be estimated reliably by a Markov Chain Monte Carlo method regardless of the sample size. Therefore, all inferences in bioequivalence study can be easily established. However, the Frequentist's method, proposed by Hyslop *et al.* (2000) and recommended by the FDA, tries to find the distribution of estimates by using the bootstrap method of which the efficiency may depend on sample size.

There were also researches in regards to the Bayesian approach in bioequivalence study. See, Best *et al.* (1995), Lunn *et al.* (2002) and Oh *et al.* (2003). However, most studies provided a conceptual algorithm that may be dependent on the study design. The method described in this paper is somewhat independent of the study design.

2. Gibbs Sampler in a Mixed Effect Linear Model

In this section, we will briefly review a noninformative Bayesian approach for a mixed effect linear model which can be written as:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \sum_{i=1}^{L} \mathbf{Z}_i \boldsymbol{\theta}_i + \boldsymbol{\epsilon}, \tag{2.1}$$

where $\mathbf{y}$ is a $N \times 1$ vector of observations; $\mathbf{X}$ is a full column rank design matrix of independent variables; $\boldsymbol{\beta}$ is the $p \times 1$ vector of parameters for the fixed effects; $\mathbf{Z}_i$ is the design matrix of predictor variables for random effects $\boldsymbol{\theta}_i$; $\boldsymbol{\theta}_i$ is the $m \times 1$ vector of random effects; $\boldsymbol{\epsilon}$ is the error vector. Further, $\boldsymbol{\epsilon} \sim N(\mathbf{0}, \sum_{h=1}^{m} \sigma_h^2 \mathbf{D}_h)$ where $\mathbf{D}_h$'s are known diagonal matrices of 0's and 1's such that $\sum_{h=1}^{m} \mathbf{D}_h = \mathbf{I}$, a $N \times N$ identity matrix, and $\boldsymbol{\theta} = (\boldsymbol{\theta}_1', \ldots, \boldsymbol{\theta}_L')'$ and $\boldsymbol{\epsilon}$ are independent.

For (2.1), a hierarchical Bayesian model can be set up as:

(I) $\mathbf{y} | \boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{\sigma} \sim N(\mathbf{X}\boldsymbol{\beta} + \sum_{i=1}^{L} \mathbf{Z}_i \boldsymbol{\theta}_i, \boldsymbol{\Sigma})$, where $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_m)'$ and $\boldsymbol{\Sigma} = \sum_{h=1}^{m} \sigma_h^2 \mathbf{D}_h$.

(II) $\boldsymbol{\beta}, \boldsymbol{\theta}$ and $\boldsymbol{\sigma}$ have a certain joint prior distribution, proper or improper.

In the absence of prior information, the uniform prior is customarily chosen for the fixed effects $\boldsymbol{\beta}$ (e.g., Zeger and Karim, 1991; Hobert and Casella, 1996; Natarajan and Kass, 2000; Kass and Natarajan, 2006). For the random effects, $N(\mathbf{0}, \boldsymbol{\Omega})$ is chosen by default, because frequentist thought that $\boldsymbol{\theta}_i$'s are independently and identically distributed normal random vectors with a mean $\mathbf{0}$ and variance-covariance matrix $\boldsymbol{\Omega}$. There is no reason to consider other priors: however, this prior requires a second stage prior for $\boldsymbol{\Omega}$.

The two frequently used noninformative priors are well-known Jeffrey's prior, $\pi(\boldsymbol{\Omega}) \propto |\boldsymbol{\Omega}|^{-(m+1)/2}$ (Tiao and Tan, 1965; Wang *et al.*, 1994), and a uniform prior on $\boldsymbol{\Omega}$. The first prior has the advantage of simplifying the full conditional distribution required by the Gibbs sampler: however, it can be shown that the prior leads to an improper joint posterior distribution. Note that Gibbs chains corresponding to improper posteriors are quite ill-behaved, but the Gibbs output does not inform the user that the posterior is improper (Hobert and Casella, 1996). Thus, one must be careful in choosing the improper prior distribution. On the other hand, the uniform prior can be quite informative for a small dataset (Daniels and Kass, 1999), which is a usual case in the bioequivalence study.

An easy way to avoid improper posterior distributions is to use a conjugate prior which is an inverse-Wishart distribution. It is commonly used as a noninformative proper prior with m degrees of freedom and some fixed scale matrix Ψ. The scale matrix can be chosen as the maximum likelihood estimator (Daniels and Kass, 1999). We can also choose the scale matrix given in Kass and Natarajan (2006). Finally, we refer to Gelman (2006) for the prior of $\boldsymbol{\sigma}$. That is, we assign a uniform prior distribution. Consequently, we will employ following prior distributions:

(II*)
1. $\boldsymbol{\beta}$ and $\boldsymbol{\theta}_i|\boldsymbol{\Omega}, i = 1, \ldots, L$ are independent, and $p(\boldsymbol{\beta}) \propto 1$ and $p(\boldsymbol{\theta}_i|\boldsymbol{\Omega}) \overset{iid}{\sim} N(\mathbf{0}, \boldsymbol{\Omega})$,
2. $p(\boldsymbol{\Omega}) \propto |\boldsymbol{\Omega}|^{-(2m+1)/2} \exp\left[-1/2 \operatorname{tr}(\Psi \boldsymbol{\Omega}^{-1})\right]$,
3. $p(\boldsymbol{\sigma}) \propto 1$ or equivalently $p(\sigma_1^2, \ldots, \sigma_m^2) \propto \prod_{h=1}^{m} 1/\sigma_h$,

where $\operatorname{tr}(\mathbf{A})$ denotes the trace of a matrix $\mathbf{A}$.

The Gibbs sampler consists of a set of full conditional posterior distributions of unknown parameters. The posterior distributions can be obtained from the joint posterior distribution of parameters and it can be obtained by the Bayes theorem,

$$\begin{aligned} p(\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{\Omega}, \boldsymbol{\sigma}|\mathbf{y}) &\propto p(\mathbf{y}|\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{\Omega}, \boldsymbol{\sigma}) p(\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{\Omega}, \boldsymbol{\sigma}) \\ &= p(\mathbf{y}|\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{\sigma}) p(\boldsymbol{\beta}) p(\boldsymbol{\theta}|\boldsymbol{\Omega}) p(\boldsymbol{\Omega}) p(\sigma_1^2, \ldots, \sigma_m^2). \end{aligned} \tag{2.2}$$

Among the five terms in (2.2), the first and the second terms are the only ones that are functions of $\boldsymbol{\beta}$. Thus, the full posterior distribution of $\boldsymbol{\beta}$ is proportional to the product of these two conditional distributions,

$$\begin{aligned} p(\boldsymbol{\beta}|\boldsymbol{\theta}, \boldsymbol{\Omega}, \boldsymbol{\sigma}, \mathbf{y}) &\propto p(\mathbf{y}|\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{\sigma}) p(\boldsymbol{\beta}) \\ &\propto \exp\left[-\frac{1}{2}\left(\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \sum_{i=1}^{L} \mathbf{Z}_i \boldsymbol{\theta}_i\right)' \boldsymbol{\Sigma}^{-1} \left(\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \sum_{i=1}^{L} \mathbf{Z}_i \boldsymbol{\theta}_i\right)\right] \\ &\propto \exp\left[-\frac{1}{2}\left\{\boldsymbol{\beta}' \mathbf{X}' \boldsymbol{\Sigma}^{-1} \mathbf{X} \boldsymbol{\beta} - 2\boldsymbol{\beta}' \mathbf{X}' \boldsymbol{\Sigma}^{-1} \left(\mathbf{y} - \sum_{i=1}^{L} \mathbf{Z}_i \boldsymbol{\theta}_i\right)\right\}\right] \\ &\propto \exp\left[-\frac{1}{2}\left(\boldsymbol{\beta} - \tilde{\boldsymbol{\beta}}\right)' \mathbf{X}' \boldsymbol{\Sigma}^{-1} \mathbf{X} \left(\boldsymbol{\beta} - \tilde{\boldsymbol{\beta}}\right)\right], \end{aligned}$$

where $\tilde{\boldsymbol{\beta}} = (\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\Sigma}^{-1}(\mathbf{y} - \sum_{i=1}^{L}\mathbf{Z}_i\boldsymbol{\theta}_i)$. This shows that the full posterior distribution of $\boldsymbol{\beta}$ is a normal with mean $\tilde{\boldsymbol{\beta}}$ and variance-covariance matrix $(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}$. In particular, the full posterior distribution of a subset of $\boldsymbol{\beta}$ is a normal distribution given in Lemma 2.1

Lemma 2.1. *Suppose $\boldsymbol{\beta}' = (\boldsymbol{\beta}_1', \boldsymbol{\beta}_2')$. Then under model (I) and (II*), the full posterior distribution of $\boldsymbol{\beta}_1$ is*

$$\boldsymbol{\beta}_1|\boldsymbol{\beta}_2, \boldsymbol{\theta}, \boldsymbol{\Omega}, \boldsymbol{\sigma}, \mathbf{y} \sim N\left(\left(\mathbf{X}_1'\boldsymbol{\Sigma}^{-1}\mathbf{X}_1\right)^{-1}\mathbf{X}_1'\boldsymbol{\Sigma}^{-1}\left(\mathbf{y} - \mathbf{X}_2\boldsymbol{\beta}_2 - \sum_{i=1}^{L}\mathbf{Z}_i\boldsymbol{\theta}_i\right), \left(\mathbf{X}_1'\boldsymbol{\Sigma}^{-1}\mathbf{X}_1\right)^{-1}\right),$$

where $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2)$ partitioned according to $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$.

In what follows, we will use a minus subscript to delete an element or a vector from a matrix appropriately. For instance, $\boldsymbol{\theta}_{-j}$ will denote the vector $\boldsymbol{\theta}$ with the j^{th} element θ_j being deleted.

The full posterior distribution of $\boldsymbol{\theta}_j, j = 1, \ldots, L$ is proportional to the product of $p(\mathbf{y}|\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{\sigma})$ and $p(\boldsymbol{\theta}_j|\boldsymbol{\Omega})$,

$$\begin{aligned} p(\boldsymbol{\theta}_j|\boldsymbol{\beta}, \boldsymbol{\theta}_{-j}, \boldsymbol{\Omega}, \boldsymbol{\sigma}, \mathbf{y}) &\propto p(\mathbf{y}|\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{\sigma})p(\boldsymbol{\theta}_j|\boldsymbol{\Omega}) \\ &\propto \exp\left[-\frac{1}{2}\left\{\left(\mathbf{y} - \mathbf{X}\boldsymbol{\beta} + \sum_{i=1}^{L}\mathbf{Z}_i\boldsymbol{\theta}_i\right)'\boldsymbol{\Sigma}^{-1}\left(\mathbf{y} - \mathbf{X}\boldsymbol{\beta} + \sum_{i=1}^{L}\mathbf{Z}_i\boldsymbol{\theta}_i\right) + \boldsymbol{\theta}_j'\boldsymbol{\Omega}^{-1}\boldsymbol{\theta}_j\right\}\right] \\ &\propto \exp\left[-\frac{1}{2}\left(\boldsymbol{\theta}_j - \tilde{\boldsymbol{\theta}}_j\right)'\left(\mathbf{Z}_j'\boldsymbol{\Sigma}^{-1}\mathbf{Z}_j + \boldsymbol{\Omega}^{-1}\right)\left(\boldsymbol{\theta}_j - \tilde{\boldsymbol{\theta}}_j\right)\right], \end{aligned}$$

where $\tilde{\boldsymbol{\theta}}_j = (\mathbf{Z}_j'\boldsymbol{\Sigma}^{-1}\mathbf{Z}_j + \boldsymbol{\Omega}^{-1})^{-1}\mathbf{Z}_j'\boldsymbol{\Sigma}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \sum_{i\neq j}\mathbf{Z}_i\boldsymbol{\theta}_i)$. Again, the full posterior distribution is a normal distribution.

Similarly, the full posterior distribution of $\boldsymbol{\Omega}$ will be given by the third and the fourth terms in (2.2). Since,

$$p(\boldsymbol{\Omega}|\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{\sigma}, \mathbf{y}) \propto p(\boldsymbol{\theta}|\boldsymbol{\Omega})p(\boldsymbol{\Omega}) \propto |\boldsymbol{\Omega}|^{-\frac{2m+L+1}{2}}\exp\left[-\frac{1}{2}\mathrm{tr}\left\{\left(\Psi + \sum_{\mathrm{i}=1}^{\mathrm{L}}\boldsymbol{\theta}_\mathrm{i}\boldsymbol{\theta}_\mathrm{i}'\right)\boldsymbol{\Omega}^{-1}\right\}\right],$$

we see that the full posterior distribution of $\boldsymbol{\Omega}$ is an inverse Wishart distribution. Likewise, the full posterior distribution of $\sigma_h^2, h = 1, \ldots, m$ can be obtained from

$$\begin{aligned} p(\sigma_1^2, \ldots, \sigma_m^2|\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{\Omega}, \mathbf{y}) &\propto p(\mathbf{y}|\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{\sigma})p(\sigma_1^2, \ldots, \sigma_m^2) \\ &\propto \prod_{h=1}^{m}\frac{1}{(\sigma_h^2)^{\frac{n_h+1}{2}}}\exp\left[-\frac{1}{2\sigma_h^2}(\mathbf{y} - \mathcal{M})'\mathbf{D}_h(\mathbf{y} - \mathcal{M})'\right], \end{aligned}$$

where n_h is the number of 1's in $\mathbf{D}_h$ and $\mathcal{M} = \mathbf{X}\boldsymbol{\beta} + \sum_{i=1}^{L}\mathbf{Z}_i\boldsymbol{\theta}_i$.

Lemma 2.2. *Under the model (I) and (II*), the full posterior distributions of $\boldsymbol{\theta}_j$'s, $\boldsymbol{\Omega}$ and σ_h^2's to implement the Gibbs sampler are given as follow:*

$$\boldsymbol{\theta}_j|\boldsymbol{\beta}, \boldsymbol{\theta}_{-j}, \boldsymbol{\Omega}, \boldsymbol{\sigma}, \mathbf{y} \sim N\left(\left(\mathbf{Z}_j'\boldsymbol{\Sigma}^{-1}\mathbf{Z}_j + \boldsymbol{\Omega}^{-1}\right)^{-1}\mathbf{Z}_j'\boldsymbol{\Sigma}^{-1}\left(\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \sum_{i\neq j}\mathbf{Z}_i\boldsymbol{\theta}_i\right), \left(\mathbf{Z}_j'\boldsymbol{\Sigma}^{-1}\mathbf{Z}_j + \boldsymbol{\Omega}^{-1}\right)^{-1}\right)$$

$$\boldsymbol{\Omega}|\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{\sigma}, \mathbf{y} \sim \text{Inv-Wishart}\left(\Psi + \sum_{i=1}^{L}\boldsymbol{\theta}_i\boldsymbol{\theta}_i', m + L\right)$$

and

$$\sigma_h^2|\boldsymbol{\beta},\boldsymbol{\theta},\boldsymbol{\Omega},\mathbf{y} \sim \text{Inv-Wishart}\left((\mathbf{y}-\mathcal{M})'\mathbf{D}_h(\mathbf{y}-\mathcal{M}), n_h-1\right),$$

where Inv-Wishart$(\boldsymbol{\Psi}, m)$ *denotes an inverse Wishart distribution with* m *degrees of freedom and inverse scale matrix* $\boldsymbol{\Psi}$*.*

3. Replicate Crossover Design

Let y_{ijkh} denote the underlying pharmacokinetic characteristic (possibly taking logarithms) on the j^{th} subject in the i^{th} sequence during k^{th} period where $i = 1,\ldots,s; j = 1,\ldots,n_i; k = 1,\ldots,c; h = 1,\ldots,m$. We consider following linear mixed model (Hyslop *et al.*, 2000):

$$y_{ijkh} = \mu_h + \nu_i + \pi_k + \gamma_{ikh} + \theta_{ijh} + \epsilon_{ijkh}, \tag{3.1}$$

where μ_h is the mean response of the h^{th} formulation; ν_i is the fixed effect of the i^{th} sequence; π_k is the fixed effect of the k^{th} period; γ_{ikh} is the fixed effect of interaction between sequence, period and formulation; θ_{ijh} is the random effect of the j^{th} subject in the i^{th} sequence under formulation h. For the random components, it is assumed that the distribution of $\boldsymbol{\theta}_{ij} = (\theta_{ij1},\ldots,\theta_{ijm})'$ is a multivariate normal with mean $\mathbf{0}$ and variance-covariance matrix $\boldsymbol{\Omega} = \{\sigma_{h\ell}\}_{m\times m}$ with $\sigma_{h\ell} = \sigma_{Bh}^2$ if $h = \ell$, and $\sigma_{h\ell} = \rho_{h\ell}\sigma_{Bh}\sigma_{B\ell}$ if $h \neq \ell$, and the customary random error ϵ_{ijkh}'s are mutually independent and normally distributed with mean zero and variance σ_{Wh}^2. Here σ_{Bh}^2 and σ_{Wh}^2 are the between-subject and the within-subject variances of formulation h, respectively, and ρ's are the correlations between two formulations on the same subject.

Note that in (3.1), h is determined by i and k by the nature of the replicate crossover design. It would suffice to use three subscripts to describe the model, but we use four subscripts for notational convenience. Note also that (3.1) has more parameters than can be estimated from the data. To avoid this over-parameterization, some constraints on the fixed effects should be imposed. We need one constraint on each set of ν_i's and π_k's, and $s + c$ constraints on γ_{ikh}'s. Sum-to-zero constrains are customarily imposed, but other constraints are also possible with the different interpretation of model parameters.

Then, we can write the replicate crossover design in a matrix notation as:

$$\mathbf{y} = \mathbf{X}_1\boldsymbol{\mu} + \mathbf{X}_2\boldsymbol{\nu} + \mathbf{X}_3\boldsymbol{\pi} + \mathbf{X}_4\boldsymbol{\gamma} + \sum_{i=1}^{s}\sum_{j=1}^{n_i}\mathbf{Z}_{ij}\boldsymbol{\theta}_{ij} + \boldsymbol{\epsilon}, \tag{3.2}$$

where $\boldsymbol{\mu} = \{\mu_h\}_{m\times 1}, \boldsymbol{\nu} = \{\nu_i\}_{(s-1)\times 1}, \boldsymbol{\pi} = \{\pi_k\}_{(c-1)\times 1}, \boldsymbol{\gamma} = \{\gamma_{ikh}\}_{(sc-s-c)\times 1}$, and $\mathbf{X}_i$'s and $\mathbf{Z}_{ij}$'s are design matrices. The specific form of the design matrices depends upon the study design, but it can be given in an obvious manner. Furthermore, the design matrices of fixed effects have the full column ranks, and the column spaces are mutually disjoint so that $\mathbf{X} = (\mathbf{X}_1,\ldots,\mathbf{X}_4)$ would have the full column rank. Since $\boldsymbol{\theta}_{ij} \overset{iid}{\sim} N(\mathbf{0},\boldsymbol{\Omega})$ and $\boldsymbol{\epsilon} \sim N(\mathbf{0},\boldsymbol{\Sigma})$ where $\boldsymbol{\Sigma} = \sum_{h=1}^{m}\sigma_{Wh}^2 D_h$, the Gibbs sampler proposed in the previous section is applicable to the replicate crossover design.

Example 3.1. The FDA provided data sets (http://www.fda.gov/Drugs/ScienceResearch/ResearchAreas/Biostatistics/ucm081434.htm) which were from replicate and non-replicate *in vivo* bioequivalence studies submitted to the FDA and used by agencies in support of proposals. We reanalyze a data set referred to as 17A. The data set consists of the values of AUC of 37 subjects who

Table 3.1. Bayes and ANOVA estimate(standard error) of parameters in dataset 17A

Method	Parameter						
	μ_T	μ_R	σ^2_{BT}	σ^2_{BR}	ρ_B	σ^2_{WT}	σ^2_{WR}
Bayes	7.622(.109)	7.664(.103)	0.405(.114)	0.384(.105)	0.957(.032)	0.113(.028)	0.076(.019)
ANOVA	7.621(.110)	7.662(.104)	0.400	0.365	0.937	0.098	0.067

were administered an antihypertensive path in a four-period, two-sequence crossover trial with RTTR/TRRT manner. The model of the data is a typical example of (3.2) where $s = 2, c = 4$ and $h = \{T, R\}$. Here T and R represent the test and the reference formulation, respectively.

For the analysis, the proposed Gibbs sampling was conducted for the logarithmic scale of AUC for five times to check the convergence of Gibbs sampling. Each chain consisted of 50,000 iterations with the first 20,000 iterations discarded, and every 10^{th} sample retained. The mean and standard deviation of posterior distributions were calculated. We also obtained the ANOVA estimates and their standard errors for comparison. The results are shown in Table 3.1. It seems that the classical method underestimates the parameters slightly from a Bayesian point of view.

The FDA guidance recommended the use of an ABE criterion to compare bioavailability measures for replicate and nonreplicate bioequivalence studies of both immediate- and modified-release products. The European Medicines Agency guideline (2010) also states that "To date, most bioequivalence studies are designed to evaluate average bioequivalence". Both regulatory agencies recommend to test $H_0 : \mu_T - \mu_R \leq \log(0.80)$ or $\mu_T - \mu_R \geq \log(1.25)$ against $H_1 : \log(0.80) < \mu_T - \mu_R < \log(1.25)$ at 5% significance level to claim the ABE. The TOST (two one-side tests) is a well-known classical approach for the problem. That is, if 90% confidence interval of $\mu_T - \mu_R$ is contained in $(\log(0.8), \log(1.25))$, then the ABE can be claimed. A classical 90% confidence interval for $\mu_T - \mu_R$ is $(-0.1409, 0.0578)$, while Bayesian credible set is calculated as $(-0.1402, 0.0557)$. Thus, the ABE can be concluded by both frequentist and Bayesian. The Bayesian treats the testing problem somewhat differently. However, we used the TOST principle to follow the recommendation of regulatory agencies. The credible set is computed from the posterior distribution of $\mu_T - \mu_R$.

The FDA stated the logarithmic scale of AUC of dataset 17A is considered to have large subject-by-formulation interaction ($\mathrm{Var}(\theta_{ijT} - \theta_{ijR}) > (0.15)^2$) that would make it necessary to assess the IBE. To assess IBE by the moment-based criterion, which was recommended by the FDA, it needs to test $H_0^{\mathrm{IBE}} : \Theta_{\mathrm{IBE}} \geq \theta_{\mathrm{ibe}}$ vs. $H_1^{\mathrm{IBE}} : \Theta_{\mathrm{IBE}} < \theta_{\mathrm{ibe}}$ at 5% significance level where

$$\Theta_{\mathrm{IBE}} = \frac{(\mu_T - \mu_R)^2 + \sigma_D^2 + \sigma_{WT}^2 - \sigma_{WR}^2}{\max\{\sigma_{W0}^2, \sigma_{WR}^2\}},$$

and θ_{ibe} and σ_{W0}^2 are predefined values. Similarly, PBE can be assessed by testing $H_0^{\mathrm{PBE}} : \Theta_{\mathrm{PBE}} \geq \theta_{\mathrm{pbe}}$ vs. $H_1^{\mathrm{PBE}} : \Theta_{\mathrm{PBE}} < \theta_{\mathrm{pbe}}$ with given θ_{pbe} and σ_0^2, and

$$\Theta_{\mathrm{PBE}} = \frac{(\mu_T - \mu_R)^2 + \sigma_T^2 + \sigma_R^2}{\max\{\sigma_0^2, \sigma_R^2\}}.$$

The FDA gave $\sigma_{W0}^2 = \sigma_0^2 = 0.04, \theta_{\mathrm{ibe}} = 2.4948$ and $\theta_{\mathrm{pbe}} = 1.74483$. One may refer Hauschke *et al.* (2007) for the rational of these values.

The estimates of Θ_{IBE} and Θ_{PBE} can be obtained easily by substituting REML or ANOVA estimates for corresponding parameters; however, these estimates cannot be used directly for the testing problems, since the distributions of those estimates are unknown. The frequentist solved this problem by employing the bootstrap method or constructing approximate confidence intervals for

Table 3.2. Point estimate and 90% confidence interval for $\mu_R - \mu_T$ for assessing ABE, and Point estimates and 95% one-side confidence limits for Δ^{PBE} and Δ^{IBE} for assessing PBE and IBE

Bioequivalence concept	Point estimate		Confidence limits			
			Lower	Upper	Lower	Upper
	ANOVA	Bayes	ANOVA		Bayes	
ABE	−0.0416	−0.0420	−0.1409	0.0578	−0.1402	0.0557
PBE	−0.6869	−0.7396		−0.2708		−0.4438
IBE	−0.0853	−0.1098		0.2709		0.0061

linear functions of variance components. However, the Bayesian can enjoy the advantage of Gibbs sampling for the testing problems. The posterior distributions of Θ_{PBE} and Θ_{IBE} are obtainable during the Gibbs sampling process.

According to the FDA guidance (1997), the test should be based on a one-side upper 95% confidence interval, and the null hypothesis is rejected if the upper bound is smaller than the predefined value θ. The upper bounds for testing H^{PBE} and H^{IBE} are 0.4706 and 2.6142, respectively. Thus, PBE can be concluded. However, IBE cannot be concluded. This may be due to a high subject-by-formulation interaction.

A classical approach for assessing PBE or IBE is to test a linearized version of the hypothesis that can be written as $H_0 : \Delta \geq 0$ vs. $H_1 : \Delta < 0$ where $\Delta^{\text{IBE}} = (\mu_T - \mu_R)^2 + \sigma_D^2 + \sigma_{WT}^2 - \sigma_{WR}^2 - \theta_{\text{ibm}} \max\{\sigma_{W0}^2, \sigma_{WR}^2\}$ and $\Delta^{\text{PBE}} = (\mu_T - \mu_R)^2 + \sigma_T^2 + \sigma_R^2 - \theta_{\text{pbe}} \max\{\sigma_0^2, \sigma_R^2\}$. Table 3.2 shows the upper bounds of one-side 95% confidence intervals for the linearized version of hypotheses for comparison of a Bayesian and a classical approach. If the upper bound is smaller than zero, then the corresponding bioequivalence can be concluded. Thus, we see that both methods reach the same conclusions.

4. Remarks

In this paper, we introduce a hierarchical Bayesian method for a general lineal mixed model. The method is applicable to bioequivalence studies. The method is general since replicate and nonreplicate crossover designs employed in a bioequivalence study can be expressed as a linear mixed model. In addition, it is flexible as well because it can meet most demands in a bioequivalence study. For instance, the joint posterior distribution of parameters enables us to assess bioequivalence concepts with the different criterion such as the moment-based or probability-based. As previously mentioned, the FDA omitted the PBE and IBE concepts from their guidance. This is mainly due to some drawbacks of the moment-based and probability-based criterions. However, there may be a common sense that the decision on bioequivalence of two formulations should not be solely based on a comparison of the means. The between-subject variance or subject-by-formulation interactions should somehow be investigated. We believe that a new criterion that incorporates the common sense will require the flexibility of a Bayesian method in the future.

References

Anderson, S. and Hauck, W. W. (1990). Consideration of individual bioequivalence, *Journal of Pharmacokinetics and Biopharmaceutic*, **18**, 259–273.

Best, N. G., Tan, K. K. C., Gilks, W. R. and Spiegelhalter, D. J. (1995). Estimation of population pharmacokinetics using the Gibbs sampler, *Journal of Pharmacokinetics and Biopharmaceutics*, **23**, 407–435.

Carrasco, L. and Jover, L. (2003). Assessing individual bioequivalence using structural equation model, *Statistics in Medicine*, **15**, 1619–1634.

Chow, S. C., Shao, J. and Wang, H. (2003). Statistical test for population bioequivalence, *Statistica Sinica*, **13**, 539–554.

Daniels, M. D. and Kass, R. E. (1999). Nonconjugate Bayesian estimation of covariance matrices and its use in hierarchical models, *Journal of the American Statistical Association*, **94**, 1254–1263.

Europian Medical Agency (2010). *Guideline on the Investigation of Bioequivalence*, Committee for Medicinal Products for Human Use, CPMP/EWP/QWP/1401/98 Rev. 1., London.

Food and Drug Administration (1997). *In Vivo Bioequivalence Studies Based on Population and Individual Bioequivalence Approached: Guidance for Industry*, Center for Drug Evaluation and Research, Rockville, MD 20857.

Food and Drug Administration (2003). *Bioavailability and Bioequivalence Studies for Orally Administered Drug Products: General Considerations*, Center for Drug Evaluation and Research, Rockville, MD.

Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models, *Bayesian Analysis*, **1**, 515–533.

Hauck, W. W. and Anderson, S. (1992). Types of bioequivalence and related statistical considerations, *International Journal of Clinical Pharmacology, Therarp, and Toxicology*, **30**, 181–187.

Hauschke, D., Steinijas, V. and Pigeot, I. (2007). *Bioequivalence Studies in Drug Development: Methods and Applications*, John Wiley & Sons, New York.

Hobert, J. P. and Casella, G. (1996). The effect of improper priors on Gibbs sampling in hierarchical linear mixed models, *Journal of the American Statistical Association*, **91**, 1461–1473.

Hsuan, F. C. and Reeve, R. (2003). Assessing individual bioequivalence with high-order cross-over designs: a unified procedure, *Statistic in Medicine*, **22**, 2847–2860.

Hylsop, T., Hsuan, F. and Holder, D. J. (2000). A small sample confidence interval approach to access individual bioequivalence, *Statistic in Medicine*, **19**, 2885–2897.

Jung, G.-J., Lim, N.-K. and Park, S.-G. (2010). Individual bioequivalence tests under 3×2 design, *The Korean Journal of Applied Statistics*, **23**, 139–150.

Kass, R. E. and Natarajan, R. (2006). A default conjugate prior for variance components in generalized linear mixed models (Comment on article by Browne and Draper), *Bayesian Analysis*, **1**, 535–542.

Lee, S.-C. and Lee, D. (2002). Bayesian analysis of multivariate threshold animal models using Gibbs sampling. *Journal of the Korean Statistical Society* **31**, 177–198.

Lunn, D. J., Best, N., Thomas, A. and Wakefield, J. (2002). Bayesian analysis of population PK/PD models: General concepts and software, *Journal of Pharmacokinetics and Pharmacodynamics*, **29**, 271–307.

Natarajan, R. and Kass, R. E. (2000). Reference Bayesian methods for general linear mixed models, *Journal of American Statistical Association*, **95**, 227–237.

Oh, H. S., Ko, S.-G. and Oh, M.-S. (2003). A Bayesian approach to accessing population bioequivalence in a 2×2 crossover design, *Journal of Applied Statistics*, **30**, 881–891.

Shao, J., Chow, S. C. and Wang, H. (2002). Individual bioequivalence testing under 2×3 designs, *Statistics in Medicine*, **21**, 629–648.

Tiao, G. C. and Tan, W. Y. (1965). Bayesian analysis of random-effects models in the analysis of variance. I. Posterior distribution of variance components, *Biometrika*, **52**, 37–53.

Wang, C. S., Rutledge, J. J. and Gianola, D. (1994). Bayesian analysis of mixed linear models via Gibbs sampling with an application to litter size of Iberian pigs, *Genetique, Selection, Evolution*, **26**, 91–115.

Zeger, S. L. and Karim, M. R. (1991). Generalized linear models with random effects; A Gibbs sampling approach, *Journal of the American Statistical Association*, **86**, 79–86.