Testing Population and Individual Bioequivalence: A Hierarchical Bayesian Approach

Pulak Ghosh* and Ioannis Ntzoufras†

October 20, 2005

Abstract

The US Food and Drug Administration has recommended statistical guidelines that introduce two new bioequivalence criteria: population bioequivalence (PBE) and individual bioequivalence (IBE). In this paper we propose a hierarchical Bayesian methodology for evaluating these recently introduced criteria. We derive the joint posterior distribution of the parameters involved in a crossover design and propose two Bayesian testing procedures for bioequivalence. The method is also extended to incorporate the t distribution in order to facilitate a more robust approach. All methods are illustrated using a popular dataset.

*Department of Mathematics and Statistics, Georgia State University, Atlanta, GA, 30303-3083, USA, Email: pghosh@mathstat.gsu.edu

†Department of Statistics, Athens University of Economics and Business, Greece, e-mail: ntzoufras@aueb.gr.

Key words: Crossover design; Gibbs sampling; Individual bioequivalence; Markov Chain Monte Carlo; Population bioequivalence; t-distribution;

1 Introduction

Bioequivalence studies are performed in order to demonstrate that a test (T) and a reference (R) drug products are equivalent in terms of efficacy and safety (Patterson, 2001). Thus, bioequivalence plays an important role in the drug development process. Until recently, bioequivalence has been assessed based on average bioequivalence (ABE). However, ABE has limitations since it focuses only on the comparison of population means between test and reference drugs. Therefore, neither the intra-subject variance of the formulation under study, nor the subject-by-formulation interaction is taken into account. Hence, under average bioequivalence, it is of major concern to know whether approved generic drug products can be used safely and interchangeably. Drug interchangeability is usually classified as drug prescribability or drug switchability. Drug prescribability is defined as the physician's available choice to prescribe either a reference or a test formulation when a patient starts receiving a treatment. On the other hand, drug switchability is related to the possibility of changing one drug product to an alternative one during the patient's treatment without observing any side effects or noticeable variations concerning the treatment's efficacy and the patient's safety. While prescribability requires that the test and reference formulations are population bioequivalent (PBE), switchability refers to individual bioequivalence (IBE) between formulations.

Anderson and Hauck (1990) first demonstrated that ABE may be insufficient to guarantee switchability between formulations. Later on, in 1997 the US Food and Drug Administration (FDA) put forward a draft "Guidance" and has been updated regularly by FDA (1999, 2002), proposing means of addressing these limitations of ABE. A considerable amount of work has been done since then on IBE and PBE (Chow and Liu, 2000; Hsuan and Reeve, 2003; Carrasco and Jover, 2003; McNally, *et al.* , 2003; also see the special issue of Statistics in Medicine on individual bioequivalence, 2000, vol 20, issue 19). These two criteria (IBE and PBE) allow the assessment of subject-by-formulation interaction and compare the population means and

variances of the test and the reference drugs. IBE and PBE can provide flexible equivalence criteria for different types of drugs based on their therapeutic windows. Moreover, they encourage pharmaceutical companies to produce less variable formulations.

FDA (2001) recommends a statistical test procedure for IBE and PBE based on the methodology proposed by Hyslop *et al.* (2000) under a crossover design. The method of Hyslop *et al.* (2001) is based on finding a 95% upper confidence bound using an approximate methods. However, their approach does not account for the uncertainty of the parameter estimates. In this paper we focus on assessing IBE and PBE using the Bayesian approach which overcomes such problems. A fully Bayesian approach has important advantages. For example, it accounts for various sources of parameter uncertainty, especially for variance components, while classical frequentist approach considers only sampling uncertainty (Gill, 2002; Gelman *et al.* , 2004). In addition, inference is based on the whole posterior distribution of model parameters which can be assessed accurately enough by generating samples using Markov chain Monte Carlo methodology (Bernardo and Smith, 1994, p.353). Hence, all descriptive measures including credible intervals can be obtained from the estimated posterior distribution of the parameters of interest and can be estimated accurately and reliably regardless of the size of the data (Agresti, 1996). Another advantage is the ability to incorporate "background" information thought pertinent to the clinical question being addressed (Ghosh and Khattree, 2003). Bayesian inference thus, offers substantial benefit not only in terms of model understanding in bioequivalence assessment but also in terms of data exploration (Patterson, 2001). All the above points as well as the comment of Breslow (1990) make us agree that bioequivalence is a natural field to be assessed using Bayesian methods.

The Bayesian approach has been adopted by several researchers in order to assess ABE (Rodda and Davis, 1980; Mandallaz and Mau, 1981; Grieve, 1985; Racine-Poon *et al.* , 1987; Ghosh and Khattree, 2003). However, the application and usefulness of the Bayesian methods in IBE and PBE are not fully explored. Oh *et al.* (2003) develops a Bayesian testing for PBE and

Wellek (2000) developed a Bayesian test for a probability based IBE criteria. However, in both the above cases the approaches are different than the IBE and PBE criteria proposed by FDA.

In this paper we propose Bayesian methods of assessing IBE and PBE using the criteria proposed by FDA. We facilitate posterior model odds and probabilities in order to examine the hypotheses of IBE and PBE. These quantities have a direct probability interpretation which p-values do not have (Carlin and Louis, 2000, p.38-39). Moreover, they can be directly calculated from the output of a single MCMC run. Finally, we use the t distribution to construct a robust version of the proposed model and base the IBE/PBE inference.

In section which follows, we present the underlying statistical model and the FDA based criteria for individual and population bioequivalence. In section 3 we describe the Bayesian procedure including the specification of prior distributions and the evaluation of IBE and PBE using posterior model probabilities. A detailed numerical example based on a published bioequivalence study is provided in section 4. In Section 5 we introduce the use of IBE/PBE model based on the $t-$distribution and how it can be compared to the standard Normal model. Final comments and conclusions are provided in section 6.

2 Statistical Model and Criteria

2.1 Model Formulation

The FDA draft guidance (1999), suggests the use of two-period replicate crossover designs for assessing IBE. However, standard two-period crossover designs may be used for PBE. But, for uniformity we use a two-period replicate crossover design for both IBE and PBE throughout this article. In this paper we consider the following linear mixed-effects model with the assumption of no carryover effect, (Chinchilli and Esinhart, 1996; Hyslop *et al.* , 2000).

$$y_{ijkl} = \mu_k + \gamma_{ikl} + \delta_{ijk} + e_{ijkl}, \tag{1}$$

where $i = 1, \cdots s$ indicates the sequence, $j = 1, \cdots n_i$ indicates the subject within sequence i, $k = R, T$ indicates the treatment and $l = 1, \cdots p_{ik}$ indicates the replicate number, and p_{ik} represents the number of replicates for the kth treatment in the i th sequence. Thus, y_{ijkl} is the logarithm of the response for replicate l on treatment k for subject j in sequence i, μ_k is the population average response for the kth formulation, γ_{ikl} represents the fixed effect for replicate l on treatment k in sequence i. This fixed effect includes the period effect and sequence effects. δ_{ijk} is the random subject effect for subject j in sequence i on treatment k and e_{ijkl} is the random error for subject j within sequence i on replicate l of treatment k. In a typical four period design for two treatments one have two replications for each treatment. Thus we assume in our model $p_{iT} = p_{iR} = 2$. The minimum number of sequence is 2 ($s \geq 2$).

To avoid model over parametrization we need to impose constraints on the location parameters γ_{ikl}. For model formulation (1) parameters γ_{ikl} represent the sequence by period interaction. Usually, sum-to-zero constraints $\sum_{i=1}^{s} \sum_{l=1}^{p_{ik}} \gamma_{ikl} = 0$ are implemented.

Furthermore, the 2×1 vectors of random subject effects $\boldsymbol{\delta}_{ij}$=$(\delta_{ijR}, \delta_{ijT})^T$ are mutually independent bivariate normal with zero means and variance-covariance matrix $\boldsymbol{\Omega}$:

$$\boldsymbol{\delta}_{ij} = (\delta_{ijR}, \delta_{ijT})^T \sim N_2(\mathbf{0}, \boldsymbol{\Omega}) \text{ with } \boldsymbol{\Omega} = \begin{bmatrix} \sigma^2_{BR} & \rho\sigma_{BT}\sigma_{BR} \\ & \sigma^2_{BT} \end{bmatrix}. \tag{2}$$

Here σ^2_{BT} and σ^2_{BR} are the between subject variance components for test and reference formulations, respectively, and ρ is the correlation between the responses on the same subject corresponding to the two formulations.

To complete the model formulation, errors e_{ijkl} are assumed to be mutually independent and distributed normally with mean 0 and intra-individual variance $\sigma^2_{W_k}$ ($k = R, T$), hence

$$e_{ijkl} \sim N(0, \sigma^2_{W_k}). \tag{3}$$

Furthermore, δ_{ij} and e_{ijkl} are assumed to be mutually independent. Therefore y_{ijkl} are normally distributed random variables.

2.2 FDA Criteria

Based on the above model let us define the following parameters:

$$
\begin{aligned}
\mu_T &= \text{mean for test formulation} \\
\mu_R &= \text{mean for reference formulation} \\
\sigma^2_{WT} &= \text{within-subject variance for test drug} \\
\sigma^2_{WR} &= \text{within-subject variance for reference drug} \\
\sigma^2_{BT} &= \text{between-subject variance for test formulation} \\
\sigma^2_{BR} &= \text{between-subject variance for reference formulation} \\
\rho &= \text{correlation coefficient between individual average test and reference formulation} \\
\sigma^2_{TT} &= \text{total variance for test formulation} \\
\sigma^2_{TR} &= \text{total variance for reference formulation} \\
\sigma^2_D &= (\sigma_{BT} - \sigma_{BR})^2 + 2(1-\rho)\sigma_{BT}\sigma_{BR}
\end{aligned}
$$

Let us now denote by y_R, $y_{R'}$ the responses on two randomly selected subjects receiving the reference drug and by y_T the response of a third independently selected subject receiving the test drug. Then drug prescribability can be measured by

$$\frac{E(y_T - y_R)^2 - E(y_R - y_R^{'})^2}{\max\{\frac{1}{2}E(y_R - y_R')^2, \sigma^2_{T0}\}}. \tag{4}$$

Under model (1) and the prescribability criteria (4) the parameter of interest for PBE is given by

$$\Theta_{PBE} = \begin{cases} \frac{(\mu_T-\mu_R)^2+\sigma^2_{TT}-\sigma^2_{TR}}{\sigma^2_{TR}} & \text{when } \sigma^2_{TR} > \sigma^2_{T0} \text{ (Reference scaled criterion)} \\ \\ \frac{(\mu_T-\mu_R)^2+\sigma^2_{TT}-\sigma^2_{TR}}{\sigma^2_{T0}} & \text{when } \sigma^2_{TR} \leq \sigma^2_{T0} \text{ (Constant scaled criterion)} \end{cases}, \tag{5}$$

where σ^2_{T0} is a specified constant. At present, the FDA recommended value of σ^2_{T0} is 0.04. In

order to identify PBE we test for

$$H_0^{PBE} : \Theta_{PBE} \geq \theta_P \qquad vs. \qquad H_1^{PBE} : \Theta_{PBE} < \theta_P. \tag{6}$$

PBE is concluded if H_1^{PBE} is accepted. The FDA recommended the value of 1.7448 for θ_P. Of course, the demonstration of PBE from the data requires that $\Theta_{PBE} < \theta_P$; for details, see FDA (1999) and appendix A therein.

A similar justification like (4) can be given for switchability for IBE. The parameter of interest for IBE is

$$\Theta_{IBE} = \begin{cases} \frac{(\mu_T - \mu_R)^2 + \sigma_D^2 + \sigma_{WT}^2 - \sigma_{WR}^2}{\sigma_{WR}^2} & \text{when } \sigma_{WR}^2 > \sigma_{W0}^2 \text{ (Reference scaled criterion)} \\ \\ \frac{(\mu_T - \mu_R)^2 + \sigma_D^2 + \sigma_{WT}^2 - \sigma_{WR}^2}{\sigma_{W0}^2} & \text{when } \sigma_{WR}^2 \leq \sigma_{W0}^2 \text{ (Constant scaled criterion)} \end{cases}. \tag{7}$$

Here σ_{W0}^2 is a constant. As earlier, the FDA recommended value for σ_{W0}^2 is 0.04. We identify IBE by testing

$$H_0^{IBE} : \Theta_{IBE} \geq \theta_I \qquad vs. \qquad H_1^{IBE} : \Theta_{IBE} < \theta_I. \tag{8}$$

IBE is concluded if $\Theta_{IBE} < \theta_I$ is accepted. Currently, the FDA recommended value for θ_I is 2.4948.

3 Bayesian Analysis of the Normal Bioequivalence Model

In the following, we describe in detail the Bayesian model and the procedures used to test for individual and population bioequivalence. The model likelihood, the prior distributions, the posterior simulation and the testing procedures for each hypothesis is described in detail.

3.1 Model Likelihood

The above model formulation can be summarized by the following:

$$\begin{aligned} f(y_{ijkl}|m_{ijkl}, \sigma_{W_k}) &\sim N(m_{ijkl}, \sigma^2_{W_k}) \\ m_{ijkl} &= \mu_k + \gamma_{ikl} + \delta_{ijk} \\ f(\boldsymbol{\delta}_{ij}|\boldsymbol{\Omega}) &\sim N_2(\mathbf{0}, \boldsymbol{\Omega}), \end{aligned}$$

where $\boldsymbol{\delta}_{ij}$ and $\boldsymbol{\Omega}$ are given by (2). The model likelihood of the above model is given by

$$\begin{aligned} & f(\boldsymbol{y}|\boldsymbol{\mu}, \boldsymbol{\gamma}, \boldsymbol{\delta}, \sigma^2_{WR}, \sigma^2_{WT}, \boldsymbol{\Omega}) = \\ &= \exp\left(-\frac{1}{2}\left\{ \sum_{i=1}^{s} n_i(p_{iR} + p_{iT}) \right\} \log(2\pi) - \sum_{i=1}^{s} n_i p_{iR} \log \sigma_{WR} - \sum_{i=1}^{s} n_i p_{iT} \log \sigma_{WT} \right. \\ &\quad -\frac{1}{2} \sum_{i=1}^{s} \sum_{j}^{n_i} \sum_{k\in\{R,T\}} \sum_{l=1}^{p_{ik}} \left(\frac{y_{ijkl} - \mu_k - \gamma_{ikl} - \delta_{ijk}}{\sigma_{W_k}} \right)^2 - \frac{n}{2} \log(2\pi) - \frac{n}{2} \log|\boldsymbol{\Omega}| \\ &\quad \left. -\frac{1}{2} \sum_{i=1}^{s} \sum_{j=1}^{n_i} \boldsymbol{\delta}_{ij}^T \boldsymbol{\Omega}^{-1} \boldsymbol{\delta}_{ij} \right) \mathcal{I}\left(\sum_{i=1}^{s} \sum_{l=1}^{p_{iR}} \gamma_{iRl} = 0 \right) \mathcal{I}\left(\sum_{i=1}^{s} \sum_{l=1}^{p_{iT}} \gamma_{iTl} = 0 \right) \end{aligned}$$

where $n = \sum_{i=1}^{s} n_i$ is the total number of subjects , $\boldsymbol{\mu} = (\mu_T, \mu_R)^T$, $\boldsymbol{\gamma}$ is a vector containing all γ_{ikl} parameters and $\mathcal{I}(x)$ is a indicator function taking the value of one if x true and zero otherwise.

3.2 Prior Distributions

To complete the Bayesian specification of the model, we must assign priors to the unknown fixed effect parameters μ_k, γ_{ikl}, within subject variances $\sigma^2_{W_k}$ and the random effect precision parameter Ω^{-1}. For convenience, we represent all parameters by $\boldsymbol{\theta}$ and the prior distribution of $(\mu_k, \gamma_{ikl}, \sigma^2_{WT}, \sigma^2_{WR}, \boldsymbol{\Omega}^{-1})$ is broadly decomposed as

$$f(\boldsymbol{\theta}) = f(\boldsymbol{\mu}, \boldsymbol{\gamma}, \sigma^2_{WR}, \sigma^2_{WT}, \boldsymbol{\Omega}^{-1}) = f(\boldsymbol{\gamma}) f(\boldsymbol{\Omega}^{-1}) \prod_{k\in\{R,T\}} f(\mu_k) f(\sigma^2_{W_k}).$$

We will use conditionally conjugate prior distributions for each of the above parameters. This choice simplifies computations since the conditional distributions involved in the Gibbs sampler are of known form and hence easy to generate from. In the absence of any expert opinion or generally prior information, we choose the parameters of the prior distributions in such way that reflect high uncertainty about the quantities of interest. We thus, use proper prior distributions (ensuring proper posteriors) but, in order to express minimal or low information, we specify large prior variance which ensures that they are flat over a realistic range of parameter values. In particular, we assume the following prior distributions for the unknown parameters.

3.2.1 Prior for μ_k

For the treatment effects μ_k we use a two stage hierarchical prior given by

$$f(\mu_k) \sim N(\mu_{0k}, \sigma^2_{0k}). \tag{9}$$

The above prior is assumed to vary with treatment effect. We assign a second stage prior distribution for the hyperparameters μ_{0k} and σ^2_{0k}, which creates the second stage of the prior structure:

$$f(\mu_{0k}) \sim N(\mu_{00}, \sigma^2_{00}) \quad \text{and} \quad f(\sigma^2_{0k}) \sim IG(a, b)\ , \tag{10}$$

where $x \sim IG(c, d)$ denotes the inverse gamma distribution with mean and variance of x^{-1} equal to c/d and c/d^2 respectively. The parameter d of the inverse gamma distribution has an interpretation as measuring the strength of prior belief in terms of an equivalent 'sample size' (Gelman, Carlin, Stern and Rubin, 1995).

3.2.2 Prior for γ_{ikl}

Under the usual sum-to-zero constraints we use prior distributions:

$$\gamma_{ikl} \sim N(0, \sigma^2_k) \quad \text{for } (i, l) \neq (1, 1)$$

and

$$\gamma_{1k1} = -\sum_{l=2}^{p_{1k}} \gamma_{1kl} - \sum_{i=2}^{s}\sum_{l=2}^{p_{ik}} \gamma_{ikl}.$$

3.2.3 Prior distributions for $\sigma^2_{W_k}$

To stabilize the treatment specific variance parameter we assume the following prior distribution for $\sigma^2_{W_k}$:

$$f(\sigma^2_{W_k}) \sim IG(a_k, b_k).$$

3.2.4 Prior distributions for $\mathbf{\Omega}$

Concerning $\mathbf{\Omega}$, we propose to use a set of univariate prior distributions for the components of $\mathbf{\Omega}$: $\sigma^2_{BT}, \sigma^2_{BR}$ and ρ. Therefore we propose to use univariate inverse gamma distributions for σ^2_{BT} , σ^2_{BR} and a transformed beta based distribution for ρ. Hence

$$\sigma^2_{BT} \sim IG(a_{\sigma_{BT}}, b_{\sigma_{BT}}) \quad (11)$$

$$\sigma^2_{BR} \sim IG(a_{\sigma_{BR}}, b_{\sigma_{BR}}) \quad (12)$$

$$\rho = 2\,U - 1 \text{ with } U \sim Beta(a_\rho, b_\rho) \quad (13)$$

This prior setup offers great insight to the practitioner concerning the marginal prior distributions. They can be easily specified if the expert can express his prior information in terms of expected values and the degree of their accuracy. For example, if we need a-priori support for the independence between the test and reference individual random effects then we can set the prior mean for ρ equal to zero by $a_\rho = b_\rho$. Then we can control the degree that we believe this prior assumption by the variance given by $(2a_\rho + 1)^{-1}$.

When no prior information is available, relatively non-informative priors can be selected using the prior parameters $a_{\sigma_{BT}} = b_{\sigma_{BT}} = 10^{-4}$, $a_{\sigma_{BR}} = b_{\sigma_{BR}} = 10^{-4}$ and $a_\rho = b_\rho = 1$. The

latter imposes a uniform prior for ρ defined on the interval $[-1, 1]$ which can be thought non-informative in the sense that any interval of the same length which is included in $[-1, 1]$ has the same probability.

An alternative natural choice for the prior of $\boldsymbol{\Omega}$ is an inverse Wishart distribution: $f(\boldsymbol{\Omega}^{-1}) \sim W_2(\alpha, \boldsymbol{S})$ with $W_q(\alpha, S)$ representing a q dimensional Wishart distribution with α degrees of freedom and mean $\alpha \boldsymbol{S}^{-1}$. For our analysis, diffuse priors can be chosen so that the analysis is dominated by the data likelihood. Specifically, to represent the vague prior knowledge, we propose to set the degrees of freedom for the Wishart distribution to be the minimum possible namely, to the rank of Ω (here $\alpha = 2$). We choose $\boldsymbol{S} = \boldsymbol{I}_2$. This prior seems to be a priori sensible since $\boldsymbol{S}$ can be thought as a prior belief concerning the magnitude of the covariance matrix $\boldsymbol{\Omega}$ for the random effects $\boldsymbol{\delta}_{ij}$ (Lindley, 1970, Spiegelhalter *et al.* , 1996, p.56). However, for our example, using this prior distribution has resulted to quite sensitive results highly depending on the choices of α and $\boldsymbol{S}$ and hence we will not use it to the subsequent of the paper.

Note that both of the above priors are fairly natural choices of prior distributions giving some of the advantages of conjugacy. The hierarchical prior model has the advantage that it does not rely on the fixing of the prior parameters at indirectly estimated values. Instead, the specified family of the prior distributions is integrated over, with integration measures being determined by the hyperprior specification (e.g., $\mu_{0k} \sim N(\mu_{00}, \sigma^2_{00})$ and $\sigma^2_{0k} \sim IG(a, b)$ in section 2.1.1), allowing a full propagation of uncertainty concerning the values of the prior parameters. The values of all the hyper-hyperparameters are taken to be weakly informative.

3.3 Posterior Inference via Gibbs Sampling

In order to estimate model parameters we use Markov chain Monte Carlo approach to estimate each model parameter $\boldsymbol{\theta}$. Gibbs sampler is used to iteratively generate random samples of $\boldsymbol{\theta}$ from the corresponding conditional posterior distributions. Let $(\theta_i|rest)$ denote the full conditional

distribution of parameter θ_i given the current values of all other quantities in the model. Then, we generate $l = 1, 2, \ldots, T$, random numbers from the conditional distributions $f(\mu_k|rest)$, $(\gamma_{ikl}|rest)$, $(\sigma^2_{W_k}|rest)$ and $(\mathbf{\Omega}^{-1}|rest)$. In order to avoid possible effect of the initial starting point $(\mu_k^{(0)}, \gamma_{ikl}^{(0)}, \sigma_{W_k}^{(0)}, \mathbf{\Omega}^{(0)})$ we remove the initial B iterations which are called burn-in period. Note that the total number of iterations kept $(T - B)$ should be as large as possible to ensure convergence of the chain and reduce the Monte Carlo error. Computationally, the model can be implemented using `WinBUGS` software (Spiegelhalter *et al.* , 2003) which facilitates Gibbs sampling to obtain samples from the posterior distribution. Convergence of the generated samples can be assessed using standard tools within `WinBUGS` software (trace plots, ACF plots, as well as Gelman-Rubin convergence diagnostic) or use CODA software available for Splus and R (Best *et al.* , 1995).

3.4 Evaluating Population and Individual Bioequivalence

In practice, we estimate the posterior distribution of model parameters $\boldsymbol{\theta}$ and then for any functions of these parameters such as Θ_{PBE} and Θ_{IBE} which are of main interest. Usually the posterior distribution is difficult to calculate hence we use MCMC algorithms as described in section 3.3. Parameters Θ_{PBE} and Θ_{IBE} are simply calculated as functions of the generated values of the model parameters.

The most natural approach in the Bayesian set up to evaluate evidence in favor of a hypothesis H is to use the posterior probability $f(H|\boldsymbol{y})$ of the hypothesis of interest and the corresponding posterior hypothesis odds which is given by

$$PO_{H\overline{H}} = \frac{f(H|\boldsymbol{y})}{f(\overline{H}|\boldsymbol{y})} = \frac{f(H|\boldsymbol{y})}{1 - f(H|\boldsymbol{y})}$$

where $\overline{H}$ is the complementary hypothesis of H.

Using this approach we are interested in the calculation of $f(H_1^{IBE}|\boldsymbol{y})$ and the corresponding

posterior odds

$$PO_{10}^{IBE} = \frac{f(H_1^{IBE}|\boldsymbol{y})}{f(H_0^{IBE}|\boldsymbol{y})} = \frac{f(H_1^{IBE}|\boldsymbol{y})}{1 - f(H_1^{IBE}|\boldsymbol{y})}. \tag{14}$$

We can interpret the PO_{10}^{IBE} as the odds in favor of H_1^{IBE} against H_0^{IBE} given the data $\boldsymbol{y}$.

Calculation of the above quantity from the MCMC output in our case is straightforward. Since we can draw T values from the posterior distribution of Θ_{IBE} and discarding the initial B values as burn-in period, we can estimate this probability $f(H_1^{IBE}|\boldsymbol{y})$ by

$$\widehat{f}(H_1^{IBE}|\boldsymbol{y}) = \frac{1}{T-B} \sum_{l=B+1}^{T} \mathcal{I}\left(\Theta_{IBE}^{(l)} < \theta_I\right),$$

where $\Theta_{IBE}^{(l)}$ is the value of Θ_{IBE} generated in the l-th iteration of the algorithm. Therefore an MCMC estimate of PO_{10}^{IBE} will be given by

$$\widehat{PO}_{10}^{IBE} = \frac{\widehat{f}(H_1^{IBE}|\boldsymbol{y})}{\widehat{f}(H_0^{IBE}|\boldsymbol{y})} = \frac{\sum_{l=B+1}^{T} \mathcal{I}\left(\Theta_{IBE}^{(l)} < \theta_I\right)}{\sum_{l=B+1}^{T} \mathcal{I}\left(\Theta_{IBE}^{(l)} \geq \theta_I\right)}.$$

Similarly we can evaluate the evidence in favor of H_1^{PBE} by calculating $\widehat{f}(H_1^{PBE}|\boldsymbol{y})$ and $\widehat{PO}_{10}^{PBE}$.

Alternatively, we may use the following approach for testing hypotheses (6) and (8) in the spirit of a frequentist analysis. Wang and Ghosh (2004) have developed this kind of test in an autoregressive models. Under this approach we use MCMC to estimate the posterior distribution of Θ_{IBE} and Θ_{PBE} and compute α and $1-\alpha$ percentile values of the posterior distributions of Θ_{IBE} and Θ_{PBE}, respectively, noted by $(\Theta_{IBE;\alpha},\ \Theta_{IBE;1-\alpha})$ and $(\Theta_{PBE;\alpha},\ \Theta_{PBE;1-\alpha})$. Thus, when we wish to test for H_0^{PBE} versus H_1^{PBE} then we reject H_0 if $\Theta_{PBE,1-\alpha} < \theta_p$ and hence we have evidence in favor of PBE. If H_0 is not rejected then we can try for the inverse test (H_1^{PBE} versus H_0^{PBE}). If $\Theta_{PBE,\alpha} > \theta_p$ then we reject H_1^{PBE} (PBE). Therefore, when $\Theta_{PBE,\alpha} \leq \theta_p \leq \Theta_{PBE,1-\alpha}$ there is not enough evidence in favor of either of the two hypotheses tested for. Using similar arguments we can construct the corresponding comparison for IBE. Wang and Ghosh (2004) have shown that this simple rule performs reasonably well in the sense of maintaining good frequentist properties, such as high power with low total error rates.

4 Illustrative Example

4.1 MAO Inhibitor Data

In this section, we use one of FDA's data set $14c$ (published on their website, `http://www.fda.gov/cder/bioequivdata`). The data come from a bioequivalence study between two MAO inhibitor drugs. We analyze the area under the curve (AUC) as the pharmacokinetic measure in this example. This is a 2×4 crossover design (RTRT, TRTR) with 18 and 20 subjects per sequence. Here R stands for reference drugs and T stands for test drug. Thus, there are two replication for each drug.

For our analysis, diffuse priors are chosen so that the analysis is dominated by the data likelihood. Specifically, for μ_k we use the prior distribution defined by equations (9) and (10) with $\mu_{00} = 0$, $\sigma^2_{00} = 10^{-4}$ and $a = b = 10^{-4}$. The intra-subject variance components $\sigma^2_{W_k}$ is assumed to follow $IG(10^{-4}, 10^{-4})$. For $\mathbf{\Omega^{-1}}$ we have tried both prior setups presented in 3.2.3. For the multivariate prior, following the arguments of section 3.2.3 we have assumed $\mathbf{\Omega}^{-1} \sim Wishart(2, \boldsymbol{S} = \boldsymbol{I}_2)$. Unfortunately, results were quite sensitive in terms of the specification of $\boldsymbol{S}$ hence we decided to focus on the prior setup using univariate beta and inverse gamma distributions given by equations (11 – 13) with $a_{\sigma_{BT}} = b_{\sigma_{BT}} = a_{\sigma_{BR}} = b_{\sigma_{BR}}$ and $a_\rho = b_\rho = 1$. These choices result to relatively low-information priors for σ^{-2}_{BT} and σ^{-2}_{BR} with prior means equal to one and large variances equal to 10^4. Moreover, the choice of $a_\rho = b_\rho = 1$ leads to uniform prior for ρ which can be thought as non-informative in the sense that puts the same probability to any subinterval of $[-1, 1]$ of the same length.

Details concerning the posterior probabilities of IBE and PBE given in Table 1 while posterior summaries are provided in Table 2. The posterior probabilities of individual and population bioequivalence are both very high since $f(H_0^{IBE})$ and $f(H_0^{PBE}) < 10^{-4}$ ($\widehat{PO}_{10}^{IBE}$ and $\widehat{PO}_{10}^{PBE} > 10^4$) strongly supporting that the two drug formulations are IBE and PBE. We reach to the

same conclusion using the alternative procedure proposed in Section 3.4 since $\Theta_{\text{PBE};0.95} = 0.18(< \Theta_P = 1.745)$ rejecting H_0^{PBE}. The high posterior probability of PBE can be attributed to the fact that the estimates of treatment means and marginal variances are very close for the two formulations. Similarly for individual bioequivalence, we observe that $\Theta_{IBE;0.95} = 0.63 < \theta_I = 2.49$ which rejects the null hypothesis of no IBE. It should be noted here that the two formulations are found to IBE and PBE in the FDA's method also.

The estimates of the two treatment effects are quite close. The posterior mean of parameter ρ was found to be equal to 0.95 indicating strong correlation between the responses within a subject to different formulation of the same drug. The correlation ρ of the responses within subjects is a key parameter for assessing IBE. When two formulations of the same drug are truly IBE, we expect that the responses of a subject to the two formulations will be highly correlated. Also note that σ_D^2 is inversely related to ρ. The subject by formulation interaction σ_D^2 is not substantial (posterior mean equal to 0.006). The subject by formulation interaction, expresses the lack of consistency of subject's true bioavailabilities on the test and reference formulations. The estimates of the between formulation variances are higher than the within subject variances, which is a usual case in crossover design (Schumaker and Metzler, 1998). The correlation does not have any direct effect to PBE.

4.2 Sensitivity analysis

We performed sensitivity analysis for various choices of prior parameters. In all the results we have focused our attention on the posterior means of Θ_{PBE}, Θ_{IBE} and ρ. Results seem to be generally robust for all choices of prior parameters we have tried. All the results were obtained by changing only one parameter at a time and keeping all other parameters constant to their default values given above.

For σ_k^2 we have considered values $\in \{10^\xi : \xi \in \{-4, -3, -2, -1, 0, 1, 2\}\}$ while for parameters

σ^2_{0k}, and $\sigma^2_{W_k}$ we have used priors of type $IG(a, a)$ with $a = 10^{\xi}$ for $\xi \in \{-4, -3, -2, -1, 0, 1, 2\}$. For all values of σ^2_{0k} and σ^2_k results were robust, while for $\sigma^2_{W_k}$ results were robust for $a \leq 1$. We have considered similar values for σ^2_{BT} and σ^2_{BR}. Results were robust for values of $a \leq 10$. In all the above cases the hypotheses of both IBE and PBE were strongly supported with the exception of the prior $IG(100, 100)$ for σ^2_{BT} and σ^2_{BR}.

The above analysis indicates that the proposed priors are quite robust on prior values (at least for this dataset). Moreover, in our example both IBE and PBE assumptions are strongly supported for most of the prior choices used.

5 Using Student's t Distribution for Bioequivalence

The bioequivalence model given by equations (1-3), which is widely used in bioequivalence trials (Chow *et al.* , 2000), heavily depends on the assumption of normality. Although model presented in section (2.1) offers great flexibility for modeling the within-subject correlation, frequently it suffers from the lack of robustness against outliers and skewness (Chow and Tse, 2000; Bolton ,1991). Actual bioavailability data tend to have outliers and often considerable skewness and thus a log-transformation of the response may not be sufficient. Hence, it is of practical interest, to examine alternative models that are more robust than the normal distribution. An appealing alternative is to use a distribution with thicker tail areas than the Normal distribution and similar bell shaped behavior like the Student's t distribution. Therefore we propose to use a mixed model (1) with robust distributions for the errors. In this paper, we consider Student's t distribution for modeling the measurement error distribution. Note that, to our knowledge, the effect of such robust distributions has not been examined in the context of bioequivalence. In particular, we assume,

$$e_{ijkl}|\nu \sim t(0, \sigma^2_{Wk}, \nu)$$

here, $t(0, \sigma^2_{Wk}, \nu)$ is the univariate t distribution with degrees of freedom ν, with within subject variance σ^2_{Wk}.

For the parameter ν we have used several prior set-ups. We propose to restrict our choices to a range of values $2 - w$ and hence implement a uniform prior of type $U(2, w)$. In order to stay within the t-model we can argue that a choice of $w = 30$ will be good enough to leave space for the model to decide and moreover to distinguish from the Normal model. In our examples we have used various Uniform setups for choices $w = 30, 50, 100$ and 200.

5.1 Evaluation of the t-distribution Using Posterior Model Odds.

In this section , we use posterior model odds to test various hypotheses concerning the assumption of the t-distribution for the Bioequivalence model used to estimate PBE and IBE.

In the first approach we consider a prior of type $Uniform(2, w)$ and we split the interval in K subintervals of type $[c_k, c_{k+1}]$ for $k = 1, 2, \ldots, K$ with $c_1 = 2$ and $c_{K+1} = w$. Two crucial points for the values of ν can be thought the values of 30 and 50 since for $\nu > 30$ or $\nu > 50$ the t-distribution gets sufficiently close to the normal distribution. Therefore as $k \to K$ intervals should approximate the normal distribution and hence have similar posterior model probabilities. We propose two choices of hyperparameters:
a) for $w = 194$ and $c^{[1]}_{k+1} = 2 + 48k$ for $k = 1, 2, 3, 4$ such that $(c_1, c_2) = (2, 50)$ and
b) for $w = 198$ and $c^{[2]}_{k+1} = 2 + 28k$ for $k = 1, 2, 3, 4, 5, 6, 7$ such that $(c_1, c_2) = (2, 30)$.
Using this approach, the posterior probability of each hypothesis $H_k : c_j \leq \nu < c_{k+1}$ will be

given by

$$\begin{aligned} f(H_k|\boldsymbol{y}) &= \int_{c_k}^{c_{k+1}} f(\nu|\boldsymbol{y})d\nu = \int_{c_k}^{c_{k+1}} \int_{\boldsymbol{\Theta^*}} f(\nu,\boldsymbol{\theta}^*|\boldsymbol{y})d\boldsymbol{\theta}^*d\nu \\ &= \int_{c_k}^{c_{k+1}} \int_{\boldsymbol{\Theta^*}} f(\boldsymbol{y}|\nu,\boldsymbol{\theta}^*)f(\nu,\boldsymbol{\theta}^*)d\boldsymbol{\theta}^*d\nu \left(\int_2^w \int_{\boldsymbol{\Theta^*}} f(\boldsymbol{y}|\nu,\boldsymbol{\theta}^*)f(\nu,\boldsymbol{\theta}^*)d\boldsymbol{\theta}^*d\nu\right)^{-1} \\ &= \int_{c_k}^{c_{k+1}} \int_{\boldsymbol{\Theta^*}} f(\boldsymbol{y}|\nu,\boldsymbol{\theta}^*)f(\boldsymbol{\theta}^*)f(\nu)d\boldsymbol{\theta}^*d\nu \left(\int_2^w \int_{\boldsymbol{\Theta^*}} f(\boldsymbol{y}|\nu,\boldsymbol{\theta}^*)f(\boldsymbol{\theta}^*)f(\nu)d\boldsymbol{\theta}^*d\nu\right)^{-1} \\ &= \frac{(c_{k+1}-c_k)^{-1}\int_{c_k}^{c_{k+1}}\int_{\boldsymbol{\Theta^*}} f(\boldsymbol{y}|\nu,\boldsymbol{\theta}^*)f(\boldsymbol{\theta}^*)d\boldsymbol{\theta}^*d\nu}{\sum_{k=1}^{K}(c_{k+1}-c_k)^{-1}\int_{c_k}^{c_{k+1}}\int_{\boldsymbol{\Theta^*}} f(\boldsymbol{y}|\nu,\boldsymbol{\theta}^*)f(\boldsymbol{\theta}^*)d\boldsymbol{\theta}^*d\nu}. \end{aligned} \tag{15}$$

where $\boldsymbol{\theta}^*$ is the model parameter vector $\boldsymbol{\theta}$ excluding the degrees of freedom ν and $\boldsymbol{\Theta}^*$ is the corresponding parameter space. Using MCMC, the above quantity can be estimated in a straightforward manner using

$$\widehat{f}(H_k|\boldsymbol{y}) = \frac{1}{T-B}\sum_{l=B+1}^{T} \mathcal{I}\left(c_k \leq \nu^{(l)} < c_{k+1}\right)$$

where $\nu^{(l)}$ is the value of ν generated at the l-th iteration of the MCMC algorithm.

The second approach is similar to the above but it is more model oriented in the sense that we specify different prior $f(\nu|H_k)$ where H_k is the same as in the first approach. We select $f(H_k) = 1/K$ in order to give equal prior probabilities in all hypotheses considered. Notice that posterior probabilities will be sensitive to the choice of w but not the posterior model odds provided that the two hypotheses have the same interval length. Furthermore, for each model

we define $f(\nu|H_k) \sim Uniform(c_k, c_{k+1})$. Posterior model probabilities will be now given by

$$\begin{aligned} f(H_k|\boldsymbol{y}) &= = \int_{c_k}^{c_{k+1}} \int_{\boldsymbol{\Theta^*}} f(H_k, \nu, \boldsymbol{\theta}^*|\boldsymbol{y}) d\boldsymbol{\theta}^* d\nu \\ &= \frac{\int_{c_k}^{c_{k+1}} \int_{\boldsymbol{\Theta^*}} f(\boldsymbol{y}|H_k, \nu, \boldsymbol{\theta}^*) f(\boldsymbol{\theta}^*|H_k) f(\nu|H_k) f(H_k) d\boldsymbol{\theta}^* d\nu}{\sum_{k=1}^{K} \int_{c_k}^{c_{k+1}} \int_{\boldsymbol{\Theta^*}} f(\boldsymbol{y}|H_k, \nu, \boldsymbol{\theta}^*) f(\boldsymbol{\theta}^*|H_k) f(\nu|H_k) f(H_k) d\boldsymbol{\theta}^* d\nu} \\ &= \frac{K^{-1}(c_{k+1} - c_k)^{-1} \int_{c_k}^{c_{k+1}} \int_{\boldsymbol{\Theta^*}} f(\boldsymbol{y}|H_k, \nu, \boldsymbol{\theta}^*) f(\boldsymbol{\theta}^*|H_k) d\boldsymbol{\theta}^* d\nu}{K^{-1} \sum_{k=1}^{K} (c_{k+1} - c_k)^{-1} \int_{c_k}^{c_{k+1}} \int_{\boldsymbol{\Theta^*}} f(\boldsymbol{y}|H_k, \nu, \boldsymbol{\theta}^*) f(\boldsymbol{\theta}^*|H_k) d\boldsymbol{\theta}^* d\nu} \\ &= \frac{(c_{k+1} - c_k)^{-1} \int_{c_k}^{c_{k+1}} \int_{\boldsymbol{\Theta^*}} f(\boldsymbol{y}|H_k, \nu, \boldsymbol{\theta}^*) f(\boldsymbol{\theta}^*|H_k) d\boldsymbol{\theta}^* d\nu}{\sum_{k=1}^{K} (c_{k+1} - c_k)^{-1} \int_{c_k}^{c_{k+1}} \int_{\boldsymbol{\Theta^*}} f(\boldsymbol{y}|H_k, \nu, \boldsymbol{\theta}^*) f(\boldsymbol{\theta}^*|H_k) d\boldsymbol{\theta}^* d\nu}. \end{aligned}$$

Considering the same prior distributions for $\boldsymbol{\theta}^*$ under all hypotheses (that is $f(\boldsymbol{\theta}^*|H_k) = f(\boldsymbol{\theta}^*)$) and that the model likelihood does not directly depends on H_k we end up with posterior model odds equal to the posterior model odds of the first approach (see equation 15). Following the 2nd approach, estimation of the posterior model probabilities and the corresponding odds is more computationally demanding since we need to construct more advanced MCMC samplers based on the ideas of reversible jump MCMC (Green, 1995) or Gibbs based model and variable selection techniques (see for examples in Dellaportas *et al.* , 2002, Ntzoufras, 2002, Katsis and Ntzoufras, 2005).

Finally, in our last approach, we use different t-distributions with degrees of freedom ν taking integer values from 2 to w which are a priori equally probable. We use m_ν for $\nu = 2, \ldots, w$ to denote a t-distributed model with ν degrees of freedom and with m_1 we denote the normal model. We use a discrete uniform distribution for the prior probability of each model. Hence $f(m_\nu) = 1/(w-1)$ for all $\nu = 2, \ldots, w$. Using this approach, the posterior model probabilities will be given by

$$\begin{aligned} f(m_k|\boldsymbol{y}) &= \int_{\boldsymbol{\Theta^*}} f(m_k, \nu = k, \boldsymbol{\theta}^*|\boldsymbol{y}) d\boldsymbol{\theta}^* \left(\sum_{k=1}^{K} \int_{\boldsymbol{\Theta^*}} f(m_k, \nu = k, \boldsymbol{\theta}^*|\boldsymbol{y}) d\boldsymbol{\theta}^* \right)^{-1} \\ &= \int_{\boldsymbol{\Theta^*}} f(\boldsymbol{y}|m_k, \nu = k, \boldsymbol{\theta}^*) f(\boldsymbol{\theta}^*|m_k) d\boldsymbol{\theta}^* \left(\sum_{k=1}^{K} \int_{\boldsymbol{\Theta^*}} f(\boldsymbol{y}|m_k, \nu = k, \boldsymbol{\theta}^*) f(\boldsymbol{\theta}^*|m_k) d\boldsymbol{\theta}^* \right)^{-1}. \end{aligned}$$

All posterior model probabilities $f(m_k|\boldsymbol{y})$ (for $k = 2, \ldots, w-1$) are compared with the corresponding posterior probability of m_w with $\nu = w$. Posterior model odds $PO_{kw} > 3$ indicate positive evidence in favor of m_k model with $\nu = k$.

Concerning our dataset, comparisons using the first two approaches were close as expected; see Table 3 for detailed results. All chains were run 100,000 iterations in order to achieve Monte Carlo error lower than 2.5% for the a-posteriori most probable hypothesis. Using both priors specified above, the hypothesis H_1 of the t-distribution was supported. For $w = 194$ and $K = 4$, the posterior probability of $H_1 : 2 \leq \nu < 50$ (which favors the t-distribution) was found equal to 0.51 and 0.47 indicating an increase of 102% (and 87%) due to the likelihood since this interval was a-priori supported by 0.25 . The posterior odds PO_{14} was found equal to 3.30 (or 2.84 for the 2nd approach) indicating positive evidence in favor of H_1; see Kass and Raftery (1995). Hypotheses H_2, H_3 and H_4 were close in terms of posterior probabilities indicating that the data do not carry enough information in order to discriminate between these hypotheses. Results using the two different sampling approaches slightly differ but departures of posterior model probabilities are within Monte Carlo error precision. The choice of $w = 198$ and $K = 7$ gave similar results. The hypothesis $H_1 : 2 \leq \nu < 30$ of the t-distribution was also supported with posterior probabilities 0.41 and 0.44 and posterior model odds $PO_{16} = 4.34$ and 4.71 indicating positive evidence in favor of H_1. All other hypotheses had similar posterior probabilities with corresponding posterior odds close to one, indicating minor or no differences between them. Similar results were extracted when we have further divided the interval $[2, 198]$ in 14 sub-intervals of equal length. The posterior probability of the first interval $[2, 16)$ was found equal to 6.7 times the posterior probability of the last interval $[184, 198)$ (which can be thought as an approximation of the Normal distribution). All other intervals have posterior odds lower than 2.1 indicating minor differences between them.

Finally, when a discrete prior on ν with $w = 30$ was used, we found that for $\nu = 3, 4, \ldots, 10$ the posterior odds $PO_{\nu,30} \geq 3$ with the highest values (6.68 and 6.70) corresponding to $\nu = 4$

and $\nu = 5$.

To sum up, from the above comparisons we conclude that there is clear evidence in favor of small values of ν indicating that $t - distribution$ is more appropriate than the Normal distribution, at least for this example.

6 Conclusion

In this article we present Bayesian estimation and testing procedures for bioequivalence models. In particular, we have developed two testing procedures, whose results are consistent. We go beyond normality by modeling the error distribution with t distribution. Moreover, we use posterior odds to compare various hypotheses concerning the values of the degrees of freedom (ν) of the t distribution. These comparisons help us to infer in favor of the Student's t-distribution and/or against the normal assumption. However, this is still an ongoing research. We are investigating the impact of other distributions like skew normal and skew t distribution in a bioequivalence trial as well. Our proposed methods are general, flexible and easily applicable using standard Bayesian software such as WINBUGS (code is available by the authors upon request). The Bayesian approach provides a reliable alternative to the existing FDA tests using posterior probabilities with straightforward interpretation rather than using p-values which do not have a direct probability interpretation.

References

[1] Agresti, A. (1996), *An Introduction to Categorical Data Analysis.* Wiley and Sons, New York.

[2] Anderson, S. and Hauck, W. W. (1990), Consideration of individual bioequivalence. *Journal of Pharmacokinetics and Biopharmaceutics* **18**, 259-273.

[3] Bernardo, J.M. and Smith, A.F.M. (1994), *Bayesian Theory.* Willey and Sons, Chichester, UK.

[4] Best, N., Cowles, M.K. and Vines, K. (1995), *CODA: Convergence Diagnostics and Output Analysis Software for Gibbs Sampling Output.* MRC Biostatistics Unit, Institute of Public health, Cambridge, UK.

[5] Bolton, S. (1991), Outliers-examples and opinions. Presented at *Bioavailability/Bioequivalence: Pharmacokinetic and Statistical Considerations* sponsored by Drug Information Association, August, 1991, Bethesda, Maryland.

[6] Breslow, N. (1990), Biostatistics and Bayes. *Statistical Science* **5**, 269-298.

[7] Carlin, B.P. and Louis, T.A. (2000), *Bayes and Empirical Bayes Methods for Data Analysis.* Second Edition. Chapman & Hall/CRC.

[8] Carrasco, L and Jover, L. (2003), Assessing individual bioequivalence using the structural equation model. *Statistics in Medicine* **22**, 901-912.

[9] Chinchilli, V. M. and Eisenhart, J. (1996), Design and analysis of intra-subject variability in crossover experiment. *Statistics in Medicine*, **15**, 1619-1634.

[10] Chow, S. C. and Liu, J. P. (2000), Design and analysis of bioavailabilty and bioequivalence studies. 2nd ed. *New York: Marcel Dekker.*

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

[12] Chow, S. C. and Tse, S. K. (1990), Outlier detection in bioavailability/bioequivalence studies. *Statistics in Medicine* **9** , 549-558.

[13] Dellaportas, P., Forster, J.J. and Ntzoufras, I. (2002), On Bayesian Model and Variable Selection Using MCMC. *Statistics and Computing*, **12**, 27–36.

[14] FDA (1999), Draft guidance: average, population and individual approaches to establishing bioequivalence. *Office of Generic Drugs, Center for Drug Evaluation and Research, Food and Drug Administration, Public Health Service, Food and Drug Administration, Rockville, Maryland*. `http://www.fda.gov/cder/guidance`.

[15] FDA (2001), Statistical Approaches to Establishing Bioequivalence. *Office of Generic Drugs, Center for Drug Evaluation and Research, Food and Drug Administration, Public Health Service, Food and Drug Administartion, Rockville, Maryland*. `http://www.fda.gov/cder/guidance`.

[16] FDA (2002), Bioavailability and Bioequivalence studies for orally administered drug products-General considerations. *Office of Generic Drugs, Center for Drug Evaluation and Research, Food and Drug Administration, Public Health Service, Food and Drug Administartion, Rockville, Maryland*. `http://www.fda.gov/cder/guidance`.

[17] Gelman, A. and Carlin, J.B. and Stern, H.S. and Rubin, D.B. (2003), *Bayesian Data Analysis,* 2nd ed, London: CRC Press.

[18] Ghosh, P. and Khattree, R. (2003), Bayesian Approach to average Bioequivalence Using Bayes Factor. *Journal of Biopharmaceutical Statistics* **13**, 719-734.

[19] Gill, J. (2002), *Bayesian Methods: A Social and Behavioral Sciences Approach.* Chapman & Hall/CRC. New York.

[20] Green (1995), Reversible Jump Markov Chain Monte Carlo Computation and Bayesian Model Determination. *Biometrika*, **82**, 711–732.

[21] Grieve, A.P. (1985), A Bayesian analysis of the two-period crossover design for clinical trials. *Biometrics* **41**, 979–990.

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

[23] 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.

[24] Kass, R.E. and Raftery, A.E. (1995), Bayes Factors. *Journal of the American Statistical Association*, **90**, 773–795.

[25] Katsis, A. and Ntzoufras, I. (2005), Testing Hypotheses for the Distribution of Insurance Claim Counts Using the Gibbs Sampler. *Journal of Computational Methods in Sciences and Engineering*, (to appear).

[26] Lindley, D.V. (1970), A non-frequentist view of Probability and Statistics. *TchgProbSt*, 209–222.

[27] Mandallaz, D. and Mau, J. (1981), Comparison of different methods for decision making in bioequivalence assessment. *Biometrics* **37**, 213–222.

[28] McNally, R.J., Iyer, H. and Mathew, T. (2003), Test for individual and population bioequivalence based on generalized p values. *Statistics in Medicine* 22, 31-53.

[29] Ntzoufras, I. (2002), Gibbs Variable Selection Using BUGS. *Journal of Statistical Software*, Volume **7**, Issue 7, 1 - 19 .

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

[31] Patterson, S. D. (2001), A Review of the development of Biostatistical Design and Analysis Techniques for Assessing In Vivo Bioequivalence. *Indian journal of Pharamaceuttical Sciences* 63, 81-100.

[32] Racine-Poon, A., Grieve, A. P., Fluhler, H., and Smith, A. F. M. (1987), A two-stage procedure for bioequivalence studies. *Biometrics* **43**, 847–856.

[33] Rodda, B. E., and Davis, R. L. (1980), Determining the probability of an important difference in bioavailability. *Clinical Pharmacology and Therapy* **28**, 247–252.

[34] Shumaker, R. C. and Metzler, C. M. (1998), The phenytoin trial is a case study of individual bioequivalence. *Drug information Journal*, **32**, 1063–1072.

[35] Spiegelhalter, D., Thomas, A., Best, N. and Gilks, W.(1996), *BUGS 0.5: Bayesian Inference Using Gibbs Sampling Manual*. MRC Biostatistics Unit, Institute of Public Health and Department of Epidemiology & Public Health, Imperial College School of Medicine, available at `http://www.mrc-bsu.cam.ac.uk/bugs`.

[36] Spiegelhalter, D., Thomas, A., Best, N. and Lunn, D. (2003), *WinBUGS User Manual, Version 1.4* . MRC Biostatistics Unit, Institute of Public Health and Department of Epidemiology & Public Health, Imperial College School of Medicine, available at `http://www.mrc-bsu.cam.ac.uk/bugs`.

[37] Wang, D. and Ghosh, S. K. (2004), Bayesian analysis of Random Coefficient autoregressive models. *Institute of Statistics Mimeo Series* 2566.

[38] Wellek, S. (2000), Bayesian Construction of an Improved Parametric Test for Probability-Based Individual Bioequivalence. *Biometrical Journal* **42**, 1039-1052.

			percentiles			
Parameter	mean	St.Dev.	5%	95%	θ	$f(H_0\|\mathbf{y})$
PBE	-0.113	0.16	-0.35	0.18	1.74	0.0000
IBE	0.327	0.18	0.06	0.63	2.49	0.0000

Table 1: Bioequivalence Testing Results for MAO Inhibitor Data Using Univariate Priors for $\mathbf{\Omega}$ (1,000 burn-in; 10,000 iterations kept).

			percentiles		
Parameter	mean	St.Dev.	2.5%	median	97.5%
σ^2_{BR}	0.048	0.013	0.028	0.047	0.078
σ^2_{BT}	0.034	0.011	0.018	0.032	0.060
σ^2_{D}	0.006	0.005	0.001	0.005	0.017
σ^2_{TR}	0.062	0.013	0.041	0.060	0.092
σ^2_{TT}	0.053	0.011	0.037	0.052	0.079
σ^2_{WR}	0.013	0.003	0.008	0.013	0.021
σ^2_{WT}	0.019	0.004	0.013	0.019	0.029
γ_{111}	-0.049	0.042	-0.131	-0.049	0.034
γ_{112}	-0.007	0.042	-0.088	-0.007	0.075
γ_{121}	-0.044	0.040	-0.122	-0.045	0.035
γ_{122}	-0.002	0.040	-0.081	-0.002	0.077
γ_{211}	0.065	0.042	-0.017	0.065	0.148
γ_{212}	-0.010	0.042	-0.092	-0.009	0.074
γ_{221}	-0.012	0.041	-0.092	-0.012	0.066
γ_{222}	0.059	0.041	-0.023	0.059	0.140
μ_1	5.640	0.037	5.567	5.641	5.711
μ_2	5.631	0.034	5.562	5.631	5.695
ρ	0.945	0.050	0.816	0.956	0.999

Table 2: Posterior Summaries of Model Parameters for MAO Inhibitor Data Using Univariate Priors for $\mathbf{\Omega}$ (1,000 burn-in; 10,000 iterations kept).

			1st Approach			2nd Approach		
w	k	$c_k - c_{k+1}$	$f(H_k\|\boldsymbol{y})$	MC Error	PO_{k4}	$f(H_k\|\boldsymbol{y})$	MC Error	PO_{k4}
194	2	50	0.505	0.020	3.30	0.467	0.021	2.84
	50	98	0.182	0.007	1.18	0.196	0.009	1.19
	98	146	0.160	0.007	1.05	0.172	0.008	1.04
	146	194	0.153	0.008	1.00	0.165	0.008	1.00
198	2	30	0.399	0.021	4.34	0.411	0.021	4.71
	30	58	0.123	0.006	1.33	0.121	0.006	1.38
	58	86	0.106	0.005	1.15	0.103	0.005	1.17
	86	114	0.096	0.004	1.04	0.098	0.004	1.11
	114	142	0.093	0.005	1.01	0.092	0.004	1.04
	142	170	0.092	0.005	1.00	0.087	0.004	0.99
	170	198	0.092	0.005	1.00	0.088	0.005	1.00

Table 3: MCMC Estimates of Posterior Probabilities for t-distributions defined on different intervals of degrees of freedom.

			percentiles		
Parameter	mean	St.Dev.	2.5%	median	97.5%
σ^2_{BR}	0.051	0.016	0.028	0.047	0.098
σ^2_{BT}	0.036	0.011	0.020	0.034	0.063
σ^2_{D}	0.006	0.004	0.001	0.005	0.017
σ^2_{TR}	0.061	0.016	0.039	0.058	0.108
σ^2_{TT}	0.050	0.011	0.033	0.049	0.078
σ^2_{WR}	0.011	0.003	0.005	0.010	0.018
σ^2_{WT}	0.014	0.004	0.006	0.014	0.024
γ_{111}	-0.046	0.043	-0.129	-0.047	0.041
γ_{112}	-0.004	0.042	-0.085	-0.005	0.080
γ_{121}	-0.036	0.039	-0.111	-0.037	0.042
γ_{122}	-0.006	0.040	-0.081	-0.007	0.076
γ_{211}	0.061	0.042	-0.023	0.062	0.142
γ_{212}	-0.011	0.043	-0.099	-0.010	0.071
γ_{221}	-0.012	0.040	-0.091	-0.012	0.065
γ_{222}	0.054	0.041	-0.029	0.056	0.132
μ_1	5.645	0.036	5.573	5.646	5.712
μ_2	5.630	0.033	5.565	5.630	5.693
ρ	0.946	0.050	0.819	0.959	1.000
Degrees of freedom (df)	13.11	7.73	2.86	11.49	28.91

Table 4: Posterior Summaries of $t - model$ Parameters for MAO Inhibitor Data (Prior of df: $Uniform(2, 30)$; 1,000 burn-in; 10,000 iterations kept).