A Bayesian Approach to Assessing the Risk of QT Prolongation

Suraj P. Anand* and Sujit K. Ghosh*

Institute of Statistics Mimeo Series 2600

Abstract

The standard approach to investigating a drug for its potential for QT prolongation is to construct a 90% two-sided (or a 95% one-sided) confidence interval (CI), for the difference in baseline corrected mean QTc (heart-rate corrected version of QT) between drug and placebo at each time-point, and to conclude non-inferiority if the upper limit for each CI is less than a pre-specified constant. One alternative approach is to base the non-inferiority inference on the largest difference in population mean QTc (baseline corrected) between drug and placebo. In this paper, we propose a Bayesian approach to resolving this problem using a Monte Carlo simulation method. The proposed method has several advantages over some of the other existing methods and is easy to implement in practice. We use simulated data to assess the performance of the proposed approach, and illustrate the method by applying it to a real data set obtained from a definitive QT study.

Keywords: Bayesian inference ; ICH; Monte Carlo simulation; QT/QTc interval; QTc prolongation; Thorough QT/QTc study

*Suraj Anand is a Graduate Student and Sujit Ghosh is an Associate Professor in the Department of Statistics at NC State University, Raleigh, NC 27695-8203, USA.

1 Introduction

An undesirable property associated with some non-antiarrhythmic drugs is their ability to delay cardiac repolarization, more generally known as QT prolongation (for more details see Appendix A). QT values are correlated with heart rate and hence a corrected equivalent (QTc) is used for data analysis. The International Conference of Harmonization (ICH) E14 guidelines (available at `www.fda.gov/cder/guidance/`) recommend conducting a 'Thorough QTc study' to determine whether the drug has a threshold pharmacological effect on cardiac repolarization, as detected by QTc prolongation. The study is typically carried out in healthy volunteers (as opposed to individuals at increased risk of arrhythmias) and is used to determine whether or not the effect of a drug on the QTc interval in target patient populations should be studied intensively during later stages of drug development. The standard approach to investigating a drug for its potential for QT prolongation is to construct a 90% two-sided (or a 95% one-sided) confidence interval (CI), for the difference in mean QTc between drug and placebo at each time-point, and to conclude non-inferiority if the upper limits for all these CIs are less than a prespecified constant. Usually, a positive control that is known to have an increasing effect on the QTc interval is included to assess the sensitivity of the study in detecting the QT effects. If the positive control demonstrates an effect as compared to the placebo, then the assay is considered reasonably sensitive to detection of QTc prolongation. Under the standard normality assumptions, this approach corresponds to both intersection-union test and the likelihood ratio test but is conservative and biased (Patterson and Jones, 2005). However, this approach is conceptually simple to understand and easy to implement, and being conservative it goes well with the regulatory agencies like the FDA. But, by the same virtue this approach is very stringent and may result into a high false positive rate. Hence, the industry needs a better approach and this drives the quest for new methodologies that check the false positive rates while safeguarding the interests of the consumer and regulatory agencies.

The ICH E14 guidance document defines a negative 'Thorough QT/QTc Study' as the one in which the upper bound of the 95% one-sided confidence interval for the largest time-matched mean effect of the drug on the QTc interval excludes 10 ms. This definition is chosen to provide reasonable assurance that the mean effect of the study drug on the QTc interval is not greater than 5 ms. When the largest time-matched difference exceeds the threshold, the study is termed positive. Hence, one alternative approach is to base

the non-inferiority inference on the largest difference in population mean QTc between drug and placebo. Coming up with an interval estimate for the maximum difference however is a non-trivial task. Since the current FDA guidelines for QT studies were put in place in November 2005, there has not been much research done in this area.

In a recent study, Eaton et al. (2006) obtained an asymptotic confidence interval for the parameter of maximum difference by approximating it with a smooth function and then using the 'delta method' to obtain an approximate confidence interval. The method developed by Eaton et al. (2006) is quite ingenious in that the parameter of interest is not a smooth function of the other parameters which invalidates the use of traditional delta method based on the maximum likelihood estimates of the other parameters. However, these two levels of approximation (first, of the smooth function for the parameter of interest and second, for the use of large samples) may not be satisfactory when the sample sizes are small and/or when the number of time points at which QTc is measured is large. This method is also fraught with computational problems, and in its current form its utility is limited to situations where the individual components of the time-matched mean difference between drug and placebo are small.

A good approach to analyzing a thorough QT study would be one that protects the interests of both the industry and the consumers, and is keeping in line with the ICH guidance. In this paper, we propose a Bayesian approach to obtaining an interval estimate for the maximum time-matched mean difference. We provide a thorough description of the Bayesian framework for analyzing QTc in Section 2. In Section 3, we use simulated data to assess the appropriateness of this approach, and then in Section 4 we apply the proposed method to a real data set on QTc obtained from a definitive QT study run at GlaxoSmithklime (GSK). We compare the results with those obtained from the standard method and the method proposed by Eaton et al. (2006). Finally, in Section 5, we present some discussions of the future research in this area.

2 A Bayesian Framework for QTc Data

We develop a general statistical model for the parallel study design for our data. Suppose n_1 and n_2 denote the number of subjects receiving drug and control (placebo) respectively, with QTc (baseline corrected) calculated at p time points. We denote the vector of p measurements taken on the subjects on drug by $\boldsymbol{x}_i,\ i = 1, \ldots n_1$, and those on control

by $\boldsymbol{y}_j,\ j = 1, \ldots n_2$, respectively. The measurements are assumed to be distributed identically and independently (iid) from p-variate normal populations (denoted by $\mathcal{N}_p$) and can be parameterized as follows:

$$\begin{aligned} \boldsymbol{x}_i &\overset{iid}{\sim} \mathcal{N}_p(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}), \quad i = 1, \ldots, n_1 \\ \boldsymbol{y}_j &\overset{iid}{\sim} \mathcal{N}_p(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}), \quad j = 1, \ldots, n_2 \end{aligned} \tag{1}$$

where $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ are the population means for the drug and the control groups, respectively, and the common covariance matrix $\boldsymbol{\Sigma}$ is assumed to be unknown. Let $\boldsymbol{\delta} = \boldsymbol{\mu}_1 - \boldsymbol{\mu}_2$ be the vector of differences in mean QTc between drug and control. Then the parameter of interest which is the maximum difference between the two groups is denoted by

$$\theta = \max_{1 \leq\ k \leq\ p} \delta_k \equiv \max_{1 \leq k \leq p} (\mu_{1k} - \mu_{2k}). \tag{2}$$

The goal of a QT study is to obtain an interval estimate of θ. Notice that θ is a non-smooth function of $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ and hence we can not use the delta-method (Casella and Berger, 2002, p.245) based on the maximum likelihood estimate (MLE) $\hat{\boldsymbol{\mu}}_1 = \bar{\boldsymbol{x}}$ and $\hat{\boldsymbol{\mu}}_2 = \bar{\boldsymbol{y}}$, where $\bar{\boldsymbol{x}} = \sum_i \boldsymbol{x}_i / n_1$ and $\bar{\boldsymbol{y}} = \sum_j \boldsymbol{y}_j / n_2$ denote the sample means. It is well known that (Anderson, 1984, p.167)

$$\sqrt{n} \boldsymbol{S}^{-1/2} (\hat{\boldsymbol{\delta}} - \boldsymbol{\delta}) \xrightarrow{d} \mathcal{N}_p(\boldsymbol{0}, \boldsymbol{I}_p)$$

where $n = n_1 n_2 / (n_1 + n_2)$, $\hat{\boldsymbol{\delta}} = \hat{\boldsymbol{\mu}}_1 - \hat{\boldsymbol{\mu}}_2$ is the MLE of of $\boldsymbol{\delta}$ and $\boldsymbol{S} = (\sum_{i=1}^{n_1} (\boldsymbol{x}_i - \bar{\boldsymbol{x}}) \boldsymbol{x}_i - \bar{\boldsymbol{x}})^t + \sum_{j=1}^{n_2} (\boldsymbol{y}_i - \bar{\boldsymbol{y}}) \boldsymbol{y}_i - \bar{\boldsymbol{y}})^t) / (n_1 + n_2)$ is the MLE of $\boldsymbol{\Sigma}$.

In order to obtain the posterior distribution of θ we first derive a class of conjugate priors for $\boldsymbol{\mu}_1, \boldsymbol{\mu}_2$ and $\boldsymbol{\Sigma}$, and then derive the posterior distribution of θ from the closed form posterior distribution of $\boldsymbol{\mu}_1, \boldsymbol{\mu}_2$ and $\boldsymbol{\Sigma}$. It is easy to verify that the following class of prior distributions is conjugate for the two-sample problem (Box and Tiao, 1973, chap.8),

$$\begin{aligned} \boldsymbol{\mu}_1 \mid \boldsymbol{\Sigma} &\sim \mathcal{N}_p(\boldsymbol{\mu}_{10}, \frac{1}{n_{10}} \boldsymbol{\Sigma}), \\ \boldsymbol{\mu}_2 \mid \boldsymbol{\Sigma} &\sim \mathcal{N}_p(\boldsymbol{\mu}_{20}, \frac{1}{n_{20}} \boldsymbol{\Sigma}) \ \text{ and} \\ \boldsymbol{\Sigma} &\sim \mathcal{IW}_p(a_0, \boldsymbol{B}_0), \end{aligned} \tag{3}$$

where $\mathcal{IW}_p$ denotes a p-dimensional inverse Wishart distribution with degrees of freedom

a_0 and scale matrix $\boldsymbol{B}_0$. In the absence of prior information we can use the reference prior (Bernardo, 1979) which turns out to be the Jeffreys prior given by

$$\pi(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \boldsymbol{\Sigma}) \propto |\boldsymbol{\Sigma}|^{-(p+1)/2}. \tag{4}$$

Notice that the Jeffreys prior (4) can be obtained as a limiting case of the conjugate prior described in (3) by setting $\boldsymbol{\mu}_{k0} = 0$ and letting $a_0 \to 0$ and $n_{k0} \to 0$ for $k = 1, 2$.

Thus, it is enough to obtain the posterior density of θ under the conjugate prior (3) and then take suitable limits to obtain the posterior distribution under the Jeffreys prior (4). By conjugacy it follows that the joint posterior distribution of $\boldsymbol{\mu}_1, \boldsymbol{\mu}_2$ and $\boldsymbol{\Sigma}$ is given by

$$\begin{aligned} \boldsymbol{\mu}_1 \mid \boldsymbol{\Sigma}, \boldsymbol{X}, \boldsymbol{Y} &\sim \mathcal{N}_p\left(\frac{n_{10}\boldsymbol{\mu}_{10} + n_1\bar{\boldsymbol{x}}}{n_{10}+n_1}, \frac{1}{n_{10}+n_1}\boldsymbol{\Sigma}\right) \\ \boldsymbol{\mu}_2 \mid \boldsymbol{\Sigma}, \boldsymbol{X}, \boldsymbol{Y} &\sim \mathcal{N}_p\left(\frac{n_{20}\boldsymbol{\mu}_{20} + n_2\bar{\boldsymbol{y}}}{n_{20}+n_1}, \frac{1}{n_{20}+n_2}\boldsymbol{\Sigma}\right) \\ \boldsymbol{\Sigma} \mid \boldsymbol{X}, \boldsymbol{Y} &\sim \mathcal{IW}(a_0 + n_1 + n_2, \boldsymbol{B}) \end{aligned} \tag{5}$$

where $\boldsymbol{B}^{-1} = \boldsymbol{B}_0{}^{-1} + (n_1+n_2)\boldsymbol{S} + \frac{n_{10}n_1}{n_{10}+n_1}(\bar{\boldsymbol{x}} - \boldsymbol{\mu}_{10})(\bar{\boldsymbol{x}} - \boldsymbol{\mu}_{10})^t + \frac{n_{20}n_2}{n_{20}+n_2}(\bar{\boldsymbol{y}} - \boldsymbol{\mu}_{20})(\bar{\boldsymbol{y}} - \boldsymbol{\mu}_{20})^t$ and $\boldsymbol{X}$ and $\boldsymbol{Y}$ denote the list of observed $\boldsymbol{x}_i$'s and $\boldsymbol{y}_j$'s, respectively.

Next, we use a Monte Carlo sampling scheme to obtain iid samples from the posterior distribution of θ as follows:

MC algorithm to obtain iid samples from $\theta|\boldsymbol{X}, \boldsymbol{Y}$

For $l = 1, 2, \ldots$ generate

step 1: $\boldsymbol{\Sigma}^{(l)} \mid \boldsymbol{X}, \boldsymbol{Y} \sim \mathcal{IW}(a_0 + n_1 + n_2, \boldsymbol{B})$,

step 2: $\boldsymbol{\mu}_1^{(l)} \sim \mathcal{N}_p(\frac{n_{10}\boldsymbol{\mu}_{10}+n_1\bar{\boldsymbol{x}}}{n_10+n_1}, \frac{1}{n_{10}+n_1}\boldsymbol{\Sigma}^{(l)})$,

step 3: $\boldsymbol{\mu}_2^{(l)} \sim \mathcal{N}_p(\frac{n_{20}\boldsymbol{\mu}_{20}+n_2\bar{\boldsymbol{y}}}{n_20+n_1}, \frac{1}{n_{20}+n_2}\boldsymbol{\Sigma}^{(l)})$, and

step 4: compute $\theta^{(l)} = \max_{1 \le k \le p}(\boldsymbol{\mu}_1^{(l)} - \boldsymbol{\mu}_2^{(l)})$.

It follows easily that $\theta^{(l)}$'s are iid samples from the posterior distribution of θ given $\boldsymbol{X}$ and $\boldsymbol{Y}$. Thus, we can compute any desired posterior summary as an estimate of θ using

the set of Monte Carlo (MC) samples $\theta^{(l)} : l = 1, 2, \ldots$. In particular, we compute the 90% (and one sided 95%) highest posterior density (HPD) interval estimate for θ and use it for our QT data. We use the method proposed by Hyndman (1996) to compute the HPD interval estimate based on the MC sample $\{\theta^{(l)} : l = 1, \ldots, N\}$, where the MC sample size N is chosen so that the overall MC error is less than some pre-assigned value $\epsilon > 0$.

3 A Simulation Study

We ran a Monte Carlo simulation study to assess the coverage properties of our proposed approach for different combinations of the operating characteristics drivers, namely, n_1, n_2, σ, and ρ, where we assume that the true $\boldsymbol{\Sigma} = \sigma^2[\rho \boldsymbol{I} + (1-\rho)\mathbf{1}\mathbf{1}^t]$ for illustration. Notice that our proposed method does not require that $\boldsymbol{\Sigma}$ to have specific correlation structure. For the purpose of simulations, we used a parallel study setup with $p = 6$, and true values of $\boldsymbol{\mu}_1 = (450, 460, 483, 479, 471, 467)$ and $\boldsymbol{\delta} = (1, 4.8, 5, 4.5, 1, 1)$. These values were chosen similar to a simulation study performed by Eaton et al. (2006). We vary $n = n_1 + n_2 \in \{30, 60, 80\}, n_1 = n_2 = n/2$, $\rho \in \{0, 0.2, 0.4, 0.6, 0.8\}$ and $\sigma \in \{8, 10, 12\}$. These operating characteristics provide a $3 \times 5 \times 3$ experimental design and we compute the 90% HPD interval for all possible 45 combinations of the operating characteristics. We used a class of normal conjugate priors for $\boldsymbol{\mu}_1, \boldsymbol{\mu}_2$ with the hyperparameters $\boldsymbol{\mu}_{k0} = 0$, and tuning parameters $n_{k0} = 0.001$ for $k = 1, 2$, and an inverted wishart prior for $\boldsymbol{\Sigma}$ with the hyperparameters $a_0 = p + 2$ and $\boldsymbol{B}_0 = \boldsymbol{I}$ to obtain the corresponding closed form posterior distribution and then derived the posterior distribution of θ using an MC algorithm. We used the algorithm outlined at the end of Section 2 to obtain iid samples from the posterior distribution of θ. To calculate posterior interval estimate, we generated 1000 samples from the posterior distribution using the algorithm described in the previous section, and computed the 90% highest posterior density (HPD) interval estimate for θ for the given combination of the operating characteristic parameters. We repeated this procedure 1000 times to obtain an estimate of the coverage probability as the proportion of the HPD regions containing the true parameter $\theta(= 5)$. To calculate the HPD regions, we used the method proposed by Hyndman (1996) in `R` package `hdrcde` (available at `www.robhyndman.info/Rlibrary/hdrcde`) .

The results based on simulations show that the nominal coverage of the 90% HPD

interval estimate of θ for most of the setting is close to the desired 90% although the observed patterns suggest that coverage (a) increases with ρ (b) decreases with σ, and (c) increases with n. Figures 1 and 2 summarize the simulation study results.

[Figure 1 about here]

In Figure 1, we fix $\sigma = 10$ and vary n and ρ. From Figure 1(a) it is clear that as the sample size increases ($n = 30, 60, 80$), the nominal coverage probability of a 90% HPD interval approaches the desired value of 0.90. However, it is also observed that the nominal coverages are close to the target value of 0.90 for large values of ρ when $\sigma = 10$ across all sample sizes.

[Figure 2 about here]

In Figure 2, we fix $n = 80$ and vary σ. It is evident from Figure 2 that as σ drops ($\sigma = 12, 10, 8$), the nominal coverage of a 90% HPD interval approaches to the target value of 0.90. Again we observe that when the true values of ρ is relatively larger the nominal coverage probabilities are closer to the target value regardless of the σ values. Overall, we find that our approach performs quite well given the fact that we have not used a specified structure of the common covariance matrix $\mathbf{\Sigma}$. If we use the true structure of the $\mathbf{\Sigma}$ the results will get even better (not reported here). However, in practice a specific parametric structure of the covariance may not be available and any such parametric assumption may increase the bias of the interval estimator of θ.

4 Application to a Definitive QT Study

The primary end-point of interest in a QT study is usually the change from baseline in QTc values, where baseline is defined according to the study design, the dosing pattern and other study specific considerations.

In Table 1 we present the data schematics of a typical QT study. Usually, a positive control that is known to have an increasing effect on the QTc interval is also included to assess the sensitivity of the study in detecting the QT effects. If the positive control demonstrates an effect as compared to the placebo, then the assay is considered reasonably sensitive to detection of QTc prolongation.

Subject	Time (Hrs)	1	2	3	.....	24
1	Baseline QTc	X1	X2	X3	.....	X24
1	Post-Dose (Drug)	Y1	Y2	Y3	.....	Y24
1	Change from baseline (Drug)	Y1-X1	Y2-Y1	Y3-X3	.....	Y24-X24
2	Baseline QTc	X1	X2	X3	.....	X24
2	Post-Dose (Placebo)	Y1	Y2	Y3	.....	Y24
2	Change from baseline (Placebo)	Y1-X1	Y2-Y1	Y3-X3	.....	Y24-X24

Table 1: Data schematics of a typical QT study.

The data used in this paper were obtained from a GSK conducted definitive QT study with placebo, a positive control (Moxifloxacin), and multiple drug concentration arms for the study drug, with QTc (Bazett) measurements taken at baseline and multiple time-points ($p = 10$). Although the actual study design was crossover, only data from a single period of the entire dataset were used and only data on placebo and the positive control arm (called drug henceforth) were included to mimic a parallel study design, with 36 subjects in each arm, for the purpose of illustration. The end-point of interest was change from baseline QTc value and the parameter of interest was maximum time-matched mean difference between the drug and the placebo for baseline corrected QTc. The observed maximum for the time-matched mean difference across the 10 timepoints is 17.18ms.

Figure 3 present the box plots for the change from baseline QTc values for the placebo and the drug arms respectively. It is evident that the baseline corrected QTc measurements of the drug group are generally higher than those in the placebo group, indicating a definitive QT study.

[Figure 3 about here]

The primary aim of this research was to obtain an interval estimate of θ using a Bayesian approach. We used a class of normal conjugate priors for $\boldsymbol{\mu}_1, \boldsymbol{\mu}_2$ with the hyperparameters $\boldsymbol{\mu}_{k0} = 0$, and tuning coefficients $n_{k0} = 0.001$ for $k = 1, 2$, and an inverted wishart prior for $\boldsymbol{\Sigma}$ with the hyperparameters $a_0 = p + 200$ and $\boldsymbol{B}_0 = \boldsymbol{I}$ to obtain the corresponding closed form posterior distribution and then derived the posterior distribution of θ based on it. To obtain an interval estimate for θ, we generated $N = 5000$ samples from the posterior distribution of θ and calculated a 90% HPD interval estimate

that came out to be $(15.47, 19.53)$. Figure 4 presents a plot for the histogram of MC samples and posterior density of θ using a kernel density estimate based on $N = 5000$ MC samples and a bandwidth of 0.2033.

[Figure 4 about here]

We compared our results to those obtained by using some of the existing methods in the literature. For instance, the standard method based on a mixed effect model with drug, time-point and drug by time-point interaction as fixed effects, and subject nested in drug as a random effect, yielded the largest 90% two-sided upper confidence limit to be 22.98. A comparatively simpler model with drug as the only factor with the confidence intervals calculated for each time-point, yielded the largest 90% two-sided upper confidence limit to be 22.36. The method proposed by Eaton et al. (2006) resulted into a 90% two-sided confidence interval of (14.09, 20.14) for θ. However their version of the superior biased corrected method failed to produce any results possibly as a result of computational problems owing to large components of the time-matched mean difference between drug and placebo.

Thus, the Bayesian approach appears to be resulting into a marginally less conservative upper bound, thereby decreasing the sponsor's risk of failing the 'Thorough QT study' while adequately controlling for the chances of a false favorable result. It is easy to implement and can be used efficiently to make inferences about a definitive QT study. The inference is not based on any large sample theory and can be used for any small sample size (n) and large number of time points (p).

5 Discussion and Future Research

A Bayesian (HPD based) approach appears to be a promising alternative to the standard approach for analyzing a 'Thorough QT Study'. It goes with the spirit of the ICH guidelines that advocate making inferences based on an interval estimate of θ, and is free of approximations and the use of of results from the large sample theory. A natural extension to this work will be to assess the feasibility of a Bayesian approach and its implementation for crossover designs or a more general regression model with more study specific factors.

Another Bayesian type interval estimate can be constructed using the quantiles of the posterior density of θ though we have observed bias related issues with this approach that warrant further research. A cautious reader may question our selection of priors for this approach. We selected conjugate priors for simplicity and the current calculations use very less information from the priors. We did not find the posterior intervals to be sensitive to the slight modifications of the priors in our simulation studies. The prior parameters can be updated to elicit available information based on historical data and scientific knowledge that may be available from the past and the concurrent investigations, to yield better coverage properties and estimates. Though we did not explore it in this project, it would be interesting to see how the posterior density of θ changes with changes to the prior parameters elicited from historical data. It would also be of interest to study in greater details how our Bayesian method compares with other recently proposed methods based on a direct interval estimation of θ that claim to be superior to the standard method. Finally, the model in (1) assumes a common $\boldsymbol{\Sigma}$ for both arms (i.e., for both drug and placebo groups). Within our Bayesian framework, we can relax this assumption and allow different $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$ for drug and placebo groups, respectively. The inference will proceed with the priors, $\boldsymbol{\Sigma}_1,\ \boldsymbol{\Sigma}_2 \overset{iid}{\sim} \mathcal{IW}_p(a_0, \boldsymbol{B}_0)$. The four-step MC sampling algorithm described in Section 2 can be extended to obtain MC samples from the posterior distribution of θ. Although we have not explored such possibilities, the extension appears to be quite straightforward. Last but not the least, one may explore the possibility of relaxing the normality assumption in (refmodel1). This appears to be a nontrivial extension, especially if we choose to develop a non-asymptotic method to obtain a 90% interval estimate of θ.

References

International Conference on Harmonisation (2005). Guidance for Industry: E14 Clinical Evaluation of QT/QTc Interval Prolongation and Proarrythmic Potential for Non-Antiarrythmic Drugs. Available at: `www.fda.gov/cber/gdlns/iche14qtc.htm`

Eaton, M. L., Muirhead, R. J., Mancuso, J. Y., Lolluri, S. (2006). A confidence interval for the maximal mean QT interval change due drug effect, *Drug Inf J.*, **40**, 267-271.

Patterson, S. (2005). Investigating Drug-Induced QT and QTc Prolongation in the Clinic: A Review of Statistical Design and Analysis Considerations: Report from the Pharmaceutical Research and Manufacturers of America QT Statistics Expert Team, *Drug Inf J.*, **9**, 243-266.

Patterson, S., Jones, B. Zariffa, N. (2005). Modeling and interpreting QTc prolongation in clinical pharmacology studies, *Drug Inf J.*, **39**, 437-445.

Boos, D., Hoffman, D., Kringle, R., Zhang, J. (2007). New confidence bounds for QT studies, *Statist. Med.*, **in press**.

Patterson S. and Jones B. (2005) *Bioequivalence and Statistics in Clinical Pharmacology*, Chapman and Hall, CRC Press; London.

Temple R. (2005). Overview of the concept paper, history of the QT/TdP concern; regulatory implications of QT prolongation. Presentations at Drug Information Agency/FDA Workshop. Available at: `www.diahome.org`

Acknowledgements

The authors would like to thank Dr. Sharon Murray (Clinical Pharmacology Statistics and Programming, GSK,RTP, NC) for her help with making the GSK data available and for providing insightful comments on the manuscript. Although it has been reviewed by the GSK corporation and approved for publication, it does not necessarily reflect GSK's policies and views.

Appendix: A brief background of QT Interval

[*Compiled from the Wikipedia, available online at* `http://en.wikipedia.org/`]

An electrocardiogram (ECG or EKG, abbreviated from the German Elektrokardiogramm) is a graphic produced by an electrocardiograph, which records the electrical voltage in the heart in the form of a continuous strip graph. It is the prime tool in cardiac electrophysiology, and has a prime function in the screening and diagnosis of cardiovascular diseases.

The ECG has a wide array of uses:

- Determines whether the heart is performing normally or suffering from abnormalities (eg. extra or skipped heartbeats - cardiac arrhythmia).
- May indicate acute or previous damage to heart muscle (heart attacks/myocardial infarction).
- Useful for detecting potassium, calcium, magnesium and other electrolyte disturbances.
- Allows the detection of conduction abnormalities (heart blocks and in bundle branch blocks).
- Provides information on the physical condition of the heart. A typical ECG tracing of a normal heartbeat consists of a P wave, a QRS complex and a T wave.

The QT interval is measured from the onset of the QRS complex to the end of the T wave. It represents the time between the start of ventricular depolarization and the end of ventricular repolarization. It is useful as a measure of the duration of repolarization. Figure 4 provides a schematic representation of normal ECG trace, with waves, segments, and intervals labeled.

[Figure 5 about here]

A normal QT interval is usually about 0.40 seconds. The QT interval as well as the corrected QT interval are important in the diagnosis of long QT syndrome and short QT

syndrome. The QT interval varies based on the heart rate, and various correction factors have been developed to correct the QT interval for the heart rate. The most commonly used method for correcting the QT interval for rate is the one formulated by Bazett and published in 1920. Bazett's formula is

$$QTc = \frac{QT}{\sqrt{RR}},$$

where QTc is the QT interval corrected for rate, and RR is the interval from the onset of one QRS complex to the onset of the next QRS complex, measured in seconds. An undesirable property associated with some non-antiarrhythmic drugs is their ability to delay cardiac repolarization, more generally known as QT prolongation.The long QT syndrome (LQTS) is a heart disease in which there is an abnormally long delay between the electrical excitation (or depolarization) and relaxation (repolarization) of the ventricles of the heart. It is associated with syncope (loss of consciousness) and with sudden death due to ventricular arrhythmias.

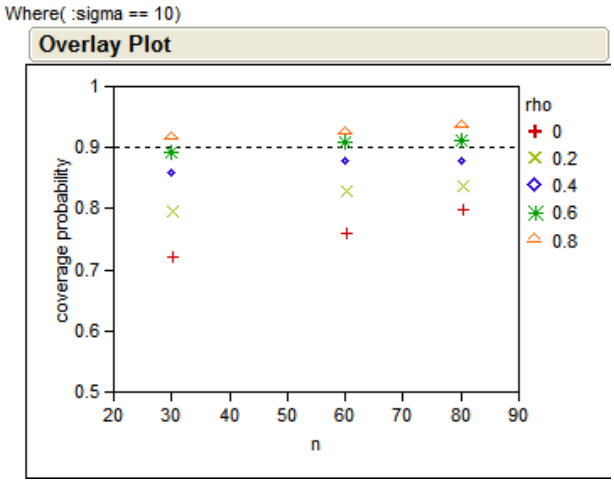

Figure 1: Plot of the nominal coverage probabilities vs. total sample size (n) and different ρ values when $\sigma = 10$.

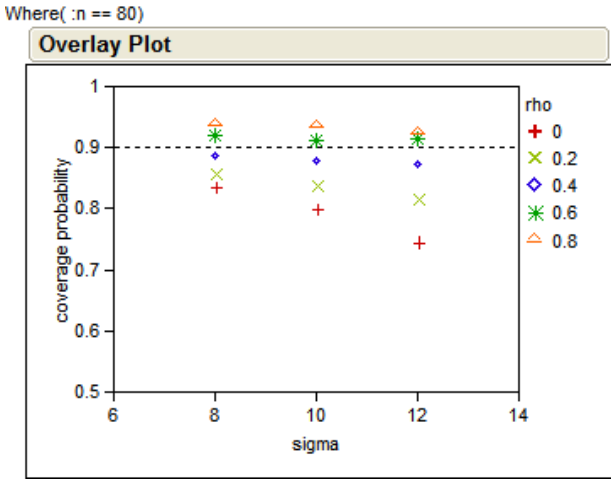

Figure 2: Plot of the nominal coverage probabilities vs. σ and different ρ values when $n = 80$.

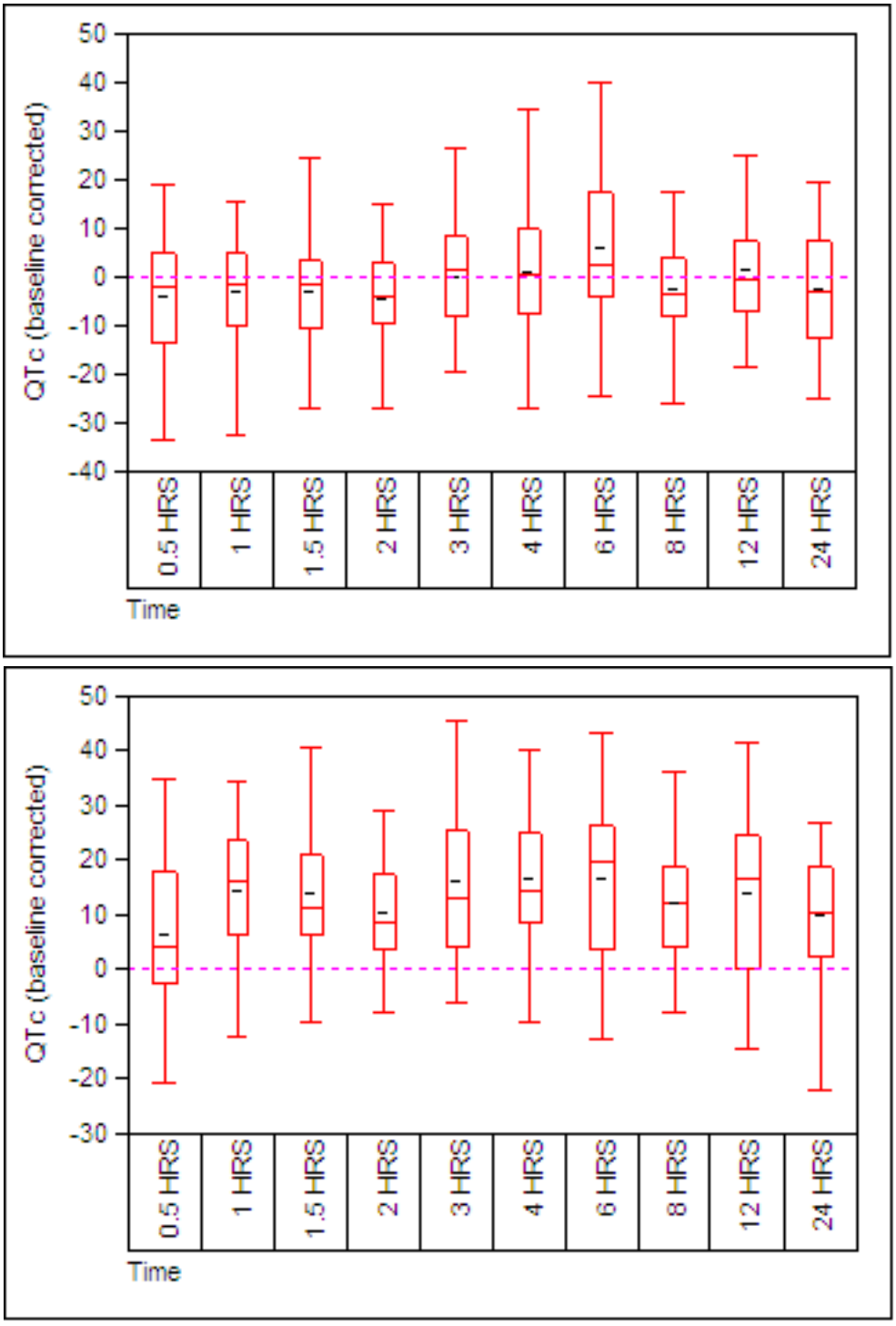

Figure 3: Box Plots for QTc (Baseline Corrected) data for the **placebo** (upper panel) and the **drug** (lower panel) groups at different timepoints ($n_1 = n_2 = 36$).

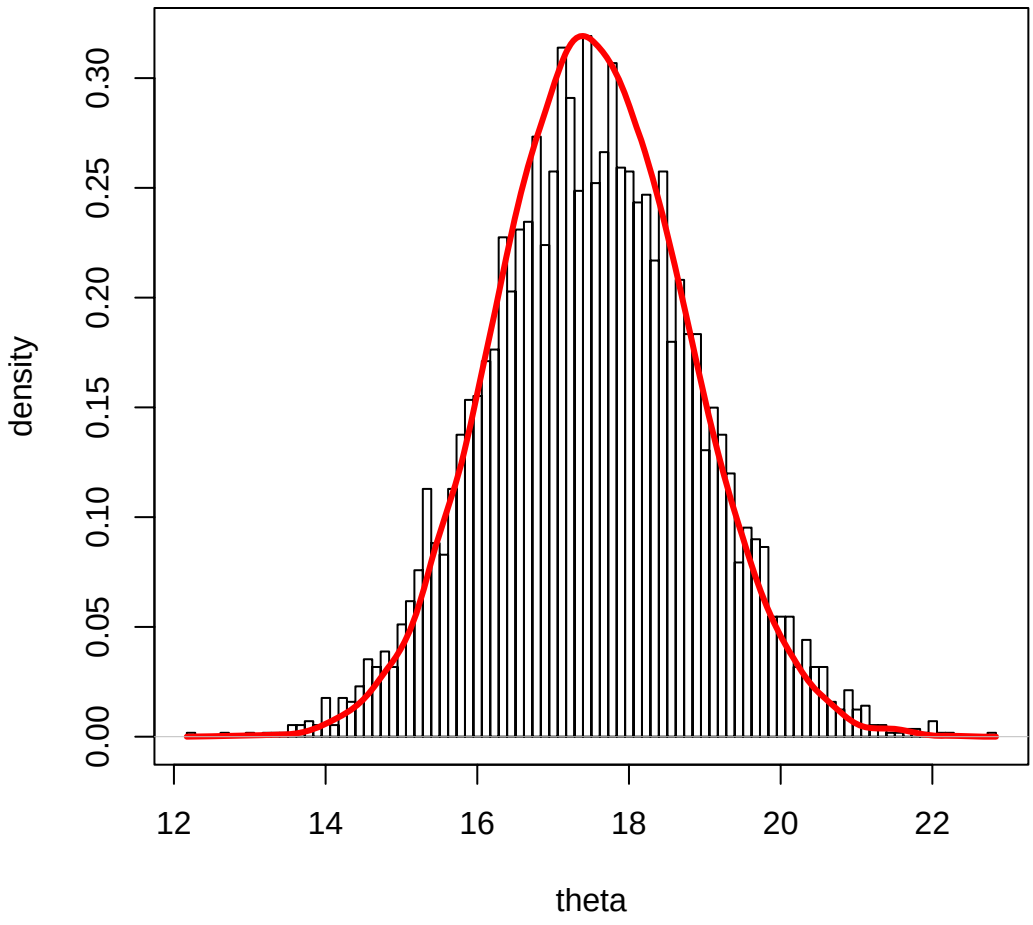

Figure 4: Histogram and kernel density estimate obtained from 5000 MC samples from the posterior density of θ.

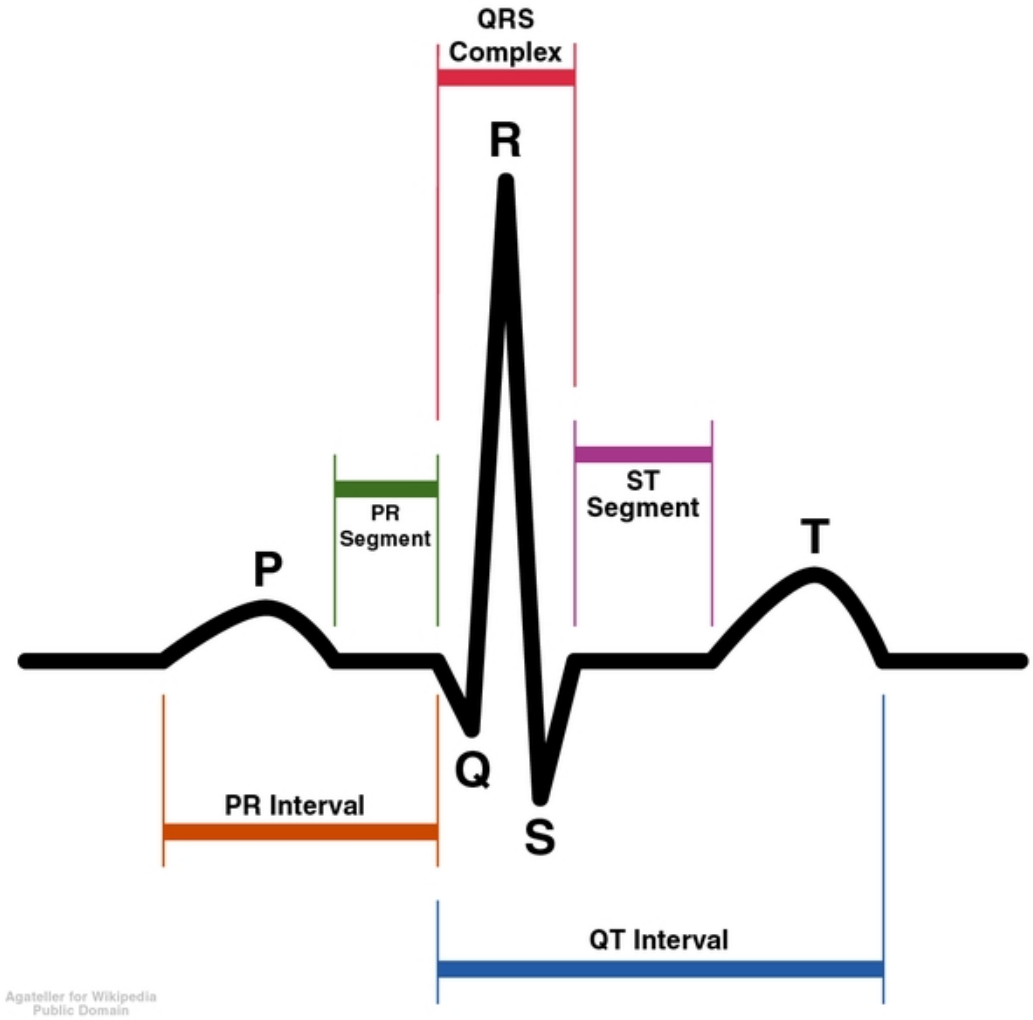

Figure 5: Normal ECG Trace: waves, segments, and intervals