Analyzing treatment effects on cardiovascular safety (QT)

Setia Pramana

promotor :
De heer Tom JACOBS, Dr. An
Vandebosch, Dr. Filip DE RIDDER

Universiteit Maastricht

universiteit hasselt

Eindverhandeling voorgedragen tot het bekomen van de graad
Master of Science in Biostatistics

Universiteit Hasselt

Center for Statistics

Analyzing Treatment Effects on Cardiovascular Safety (QT)

By:

Setia Pramana

Supervisors:

Tom Jacobs
An Vandebosch, PhD
Filip De Ridder, PhD

Thesis Submitted in Partial Fulfillment of The Requirements for The Degree of Master of Science in Biostatistics

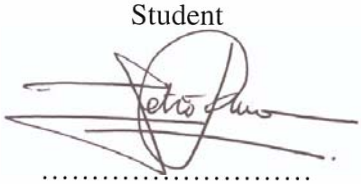
August 2007

Certification

This is to certify that this project was carried out by Setia Pramana under our thorough supervisions and reflects his true research ability.

Setia Pramana

Student

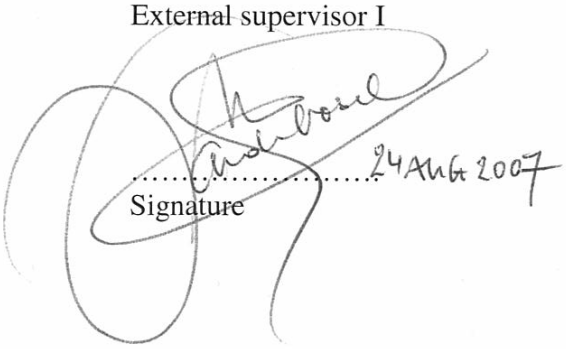


.....

Signature

An Vandebosch, PhD

External supervisor I

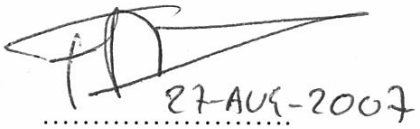


.....

Signature

Filip De Ridder, PhD

External supervisor II



.....

Signature

Tom Jacobs

Internal Supervisor

.....

Signature

Acknowledgements

I am highly delighted to give thanks to Almighty Allah who has given me the opportunity to come and study in the Hasselt University. I have to show my sincere appreciation to Vlaamse Interuniversitaire Raad (VLIR) for granting me scholarship to study here.

I wish to express sincere appreciation to my internal supervisor Mr. Tom Jacobs and my external supervisors An Vandebosch, PhD and Filip De Ridder, PhD for their assistance in the preparation of this thesis. In addition I am also grateful to my lectures and colleagues for their advice and assistance when needed.

I have to praise the patience and perseverance of my adored wife who accompany me and take care of our beloved daughter and son, Aisya and Danish, during the hard period in Biostatistics. Last in this list but first in my heart, I would like to thank my parents, brother and sisters for the emotional support and encouragement.

Abstract

All medications are expected to undergo cardiovascular safety evaluations in the beginning of development as early as possible to evaluate their effect on cardiac repolarization, an effect that can be measured as prolongation of the QT interval on the surface electrocardiogram (ECG). QT prolongation is considered to be the best available predictor of risk for dysrhythmia and sudden death.

This thesis primarily focuses on assessing the effect of a drug X at steady state compared with the placebo group on the length of the QT interval corrected for heart rate using linear derived method (QTcLD). The second objective is to study the effect of baseline imbalances on the estimated treatment effect.

In this randomized placebo-controlled trial, imbalances in the QTcLD interval distribution at baseline values and covariates between randomized groups were observed. To take this into account four approaches were performed. Properties and challenges of each method were compared and discussed. The analyses were carried into two parts: first the observations at 4 hour 30 minutes (time point 4.5) after dosing were analyzed. Then repeated measures analysis was performed to account for all time point measurements.

At time point 4.5 both doses (12 mg and 18 mg) of drug X statistically and clinically prolonged the QTcLD interval. However when all time points were considered neither 12 mg nor 18 mg of drug X at steady state had effect on the prolongation of the length of the QTcLD interval. Baseline imbalance between treatment groups in baseline was influent the results, hence adjusting in the statistical analysis for baseline variables was performed.

Key words: *Cardiovascular Safety, QTc Interval, Baseline Adjustment*.

Table of Contents

List of Tables

List of Figures

1. Introduction

Sudden death, a death from a cardiac cause within a short time (minutes to hour) after symptoms initially appear often without warning, is a major public health problem. One of several cardiac measures that have been used to predict sudden death and other arrhythmias is a prolonged QT interval (increased time taken for the hearth to recover from previous contraction) or QTc interval (Jindal, et.al, 2005).

The International Conference on Harmonisation (ICH) E14 guideline recommends to evaluate the cardiovascular safety of all drug in the beginning of development and preferably as early as possible. The commonly used clinical tool to assess the cardiovascular safety is electrocardiogram (ECG). One of the primary parameters of the ECG is the QT interval since it is considered as an important biomarker for potential lethal cardiac effects. Specifically the guideline recommends to perform a 'thorough QT/QTc study' to study the effects of the drug under investigation on the heart and on QT interval in particular. QT interval is a measure of the time between the start of the Q wave and the end of the T wave in the heart's electrical cycle as shown in figure 1. The QT interval is dependent on the heart rate. The faster the heart rate the shorter the QT interval. It has to be adjusted for heart rate to aid interpretation referred as QTc interval.

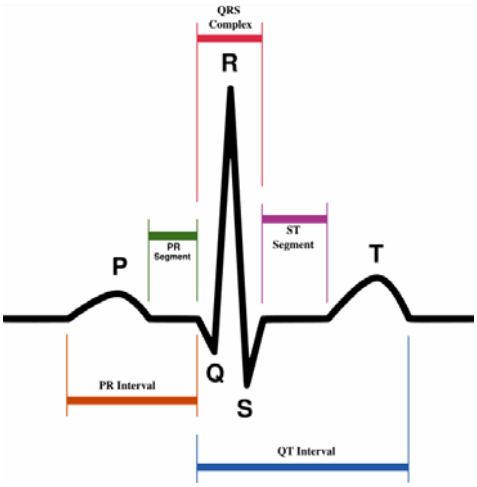


Figure 1. Schematic representation of normal ECG trace (sinus rhythm)

For the compound of interest, a thorough QT study was carried out and the drug was associated with a modest increase in QTc interval compared to placebo was established. This effect was generated at supratherapeutic plasma concentration, which is a plasma concentration which is much higher than observed in clinical practice. To obtain a

better understanding of the clinical relevance, a second QT study was carried out to compare the maximum recommended dose levels with an active comparator from the same therapeutic class. This comparator is perceived to have a good cardiovascular safety profile and showed a mean QTc prolongation of approximately 5.7 ms. However this prolongation is not thought to be clinically significant. As a secondary objective the study compared each treatment with placebo.

This thesis mainly focuses on assessing the effect on the length of the QT interval corrected for heart rate (QTc) interval of a drug X at steady state compared with placebo. The study is a randomized placebo-controlled trial. The second objective is to study the effect of baseline imbalances in the estimated treatment effect since imbalances in QTcLD distribution at baseline values and covariates were observed between randomized groups,

In a controlled trial, randomization is performed to ensure that allocation of subjects to treatments is left purely to chance. The characteristics of subjects that may influence outcome are distributed between treatment groups so that any difference in outcome can be assumed to be due to the intervention. However, imbalance between groups in baseline variables that may influence outcome (such as age or disease severity) can bias statistical tests, a property sometimes referred to as chance bias. Some protection against chance bias is given by stratified randomization and by adjusting in the statistical analysis for baseline variables (Roberts and Torgerson, 1999).

In section 2 an overview of the study design and results of an exploratory data analysis will be presented. Four possible methods of analysis will be compared and discussed in section 3. Section 4 presents the data analysis and results. In section 5 a general discussion will be given explaining the results observed in the data analysis, and finally concluding remarks and recommendation are given in section 6.

2. Study Overview and Exploratory Data Analysis

2.1 Study Objective

The study is a randomized, double-blind, drug X, active comparator and concurrent placebo, multi center study designed primary to determine whether the effect on QTc interval is comparable between a maximum recommended dose (12 mg) of drug X and that of 400 mg active comparator at steady state with twice daily dose administration. The secondary objectives are to explore the relationship between the pharmacokinetics of drug X and the ECG parameter of interest, and to evaluate the cardiovascular safety and tolerability of the maximum released dose (18 mg) of drug X at steady state.

2.2 Study Design

The study consisted of three phases: a screening phase up to 5 days (days –11 to –7), a placebo wash out phase of 6 days (days -6 to –1), and a treatment phase consisting of a one day open-label moxifloxacin period (day 1) and 12 days double blind period (days 2 to 11), including an end of study evaluation on day 12 or at early withdrawal from study. During the screening phase, the standard safety and screening evaluations were performed. Eligible subjects who gave their consent to participate in the study were admitted to the study site on the evening of day-7 to enter 6-day open label placebo wash-out period to wash-out concomitant medications and during which received placebo once daily. On the last two days of this period (day –2 and –1), serial triplicate 12-lead ECG recordings were collected and measure as baseline assessments.

On day 1 all subjects received a single dose of moxifloxacin administered in the morning to establish assay sensitivity. Moxifloxacin was selected to rule out an insensitive measurement methodology as it has a well-documented and predictable effect of prolonging QTc by approximately 5 milliseconds.

Then the subject was randomly allocated to receive one of the following three treatments during the double-blind treatment period according to the random assignment: drug X, active comparator or placebo. No stratification in randomization was performed. The study flow diagram can be seen in figure 2.

The subjects received study drug as per administration schedule from days 2 to 11 during the double-blind period after having a standardized breakfast. Serial intensive ECG measurements included multiple time points during placebo wash-out, on day 1, 6, 7, 11 and 12 were performed to take into account diurnal variation and also to obtain

maximum observed plasma concentration during a dosing interval on day 6 and 11. The measurements were taken on protocol specified time points (hour): -0.001 (pre-dose), 1, 1.5, 2.5, 3.5, 4.5, 6, 12 and 23.5. Therefore changes from baseline for each time point were analyzed. Hence each subject had repeated measurements both between and within days.

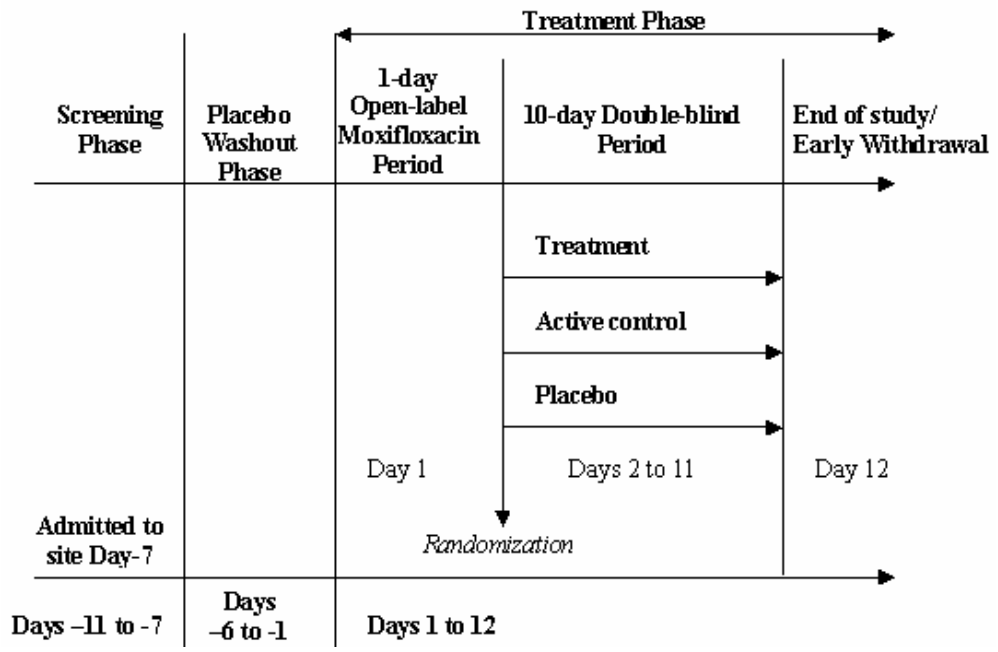


Figure 2. Study Flow Diagram

During the double-blind period, subjects in the drug X and the active comparator group received study drug according to the dosage schedule in table 1.

Table 1. Dosage Administration

Drug X	Active Comparator
12 mg on days 2 to 6.	200 mg on day 2
15 mg on day 7	400 mg on day 3
18 mg on days 8 to 11	600 mg on day 4
	800 mg on days 5 to 11

To get steady state condition and since it is not allowed to administer the dose 800 mg of active comparator directly, the drug was uptitrated from 200 mg on day 2 to 800 mg on day 5 until day 11 as shown in table 1. It is allowed to give subjects a maximum recommended dose of a drug X (12 mg) from the beginning of study, however it needs to be uptitrated by giving 15 mg of a drug X on day 7 to increase to the maximum

released dose. The maximum released dose (18 mg) is 50% higher than the maximum recommended dose (12 mg).

Subjects who withdrew from the study voluntarily or were withdrawn by the investigator before completion of all study evaluation were replaced to ensure that at least 40 subjects in the drug X and the active comparator group, and 20 subjects in the placebo group completed the study up to and including the pre-dose assessment on day 7. The minimum duration of the study for each subject including the screening phase was 19 days and the maximum duration was 23 days.

2.3 Sample Size Determination

A sample size of 40 completed subjects per treatment (drug X and active comparator) were considered to be sufficient to conclude no inferiority of 12 mg of drug X to 400 mg of active comparator twice daily with power of 81% using upper one-sided 95% confidence interval and a non inferiority criterion of 10 milliseconds (ms), when the true difference in means between drug X and active comparator equals 2 ms.

A sample size of 40 subjects in each treatment groups and 20 subjects in the placebo were considered be sufficient to estimate the difference in mean change from baseline between each drug and placebo to within 6.4 ms of the true value with 90% confidence.

It needs to point out that in this study statistically mean different from zero did not mean that clinically the groups were different. The mean difference was clinically considered to be different if it had upper bound of 95% confidence interval larger than 10ms. It is because the “thorough QT/QTc study’ is intended to determine whether the drug has a threshold pharmacologic effect on cardiac repolarization as detected by QT Interval prolongation. The threshold level of regulatory concern of the mean difference is around 5 ms. It is evidenced by an upper bound of the one-sided 95% confidence interval around the mean effect on QT interval of below 10 ms.

2.4 Study Population

Subjects between 18 and 50 years of age, a normal 12-lead ECG at screening, a body mass index between 18 and 35 kg/m^2 and other inclusion criteria were included. After completion of the trial 110 subjects had participated: 44 subjects with age ranged from 18 to 51 year in the drug X group, 44 subjects with age between 21 and 48 years in the active comparator group and 22 subjects with age between 23 to 49 years in the placebo group. Ninety-one subjects completed the study until day 11 and only 6 and 7 subjects

were left after day 1 and day 6 respectively. Six subjects were withdrawn due to protocol deviation.

Table 2. Number of Subjects withdrawn by Protocol Violation, Number of Subject left at day 1 and days 6-7, Number of Subjects remain in days 11-12

Treatment	Protocol Deviation	Day 1	Days 6-7	Days 11-12	Total
Drug X	2	3	4	35	44
Active Comparator	4	1	3	36	44
Placebo	0	2	0	20	22

The dataset used in this thesis comes only from drug X and placebo group (65 subjects). There are 20 males and 45 females aged between 18 and 51 years.

2.5 QT Correction Methodology

Since the QT interval is correlated to the length of heart rate (RR) as shown in figure 3, it has to be corrected. There are several methods to calculate the QT hearth rate corrected interval:

1. The historical clinical correction is to use Bazett's formula: $QT_B = \frac{QT}{\sqrt{RR}}$, where $RR = \frac{60}{HR(bpm)}$. Length of heart rate (RR) or duration of 1 heartbeat is the interval from the onset of one QRS complex to the onset of the next QRS complex, measured in milliseconds. Heart Rate (HR) describes the frequency of the cardiac cycled as "beats per minute" (bpm). Bazett's over-corrects at high heart rates and under-corrects at low heart rates. Figure 3 show that Bazett's Correction (QT_B) is still highly correlated with RR interval.
2. Fridericia's correction: $QT_F = \frac{QT}{\sqrt[3]{RR}}$. This method seems to have better correction as shown in figure 3. However it is still suspected to over-correct at high heart rate.
3. Linear derived method is considered as the most appropriate correction method as it incorporates all of the drug free QT/RR interval of the study in the linear model to derive the study-specific correction formula.

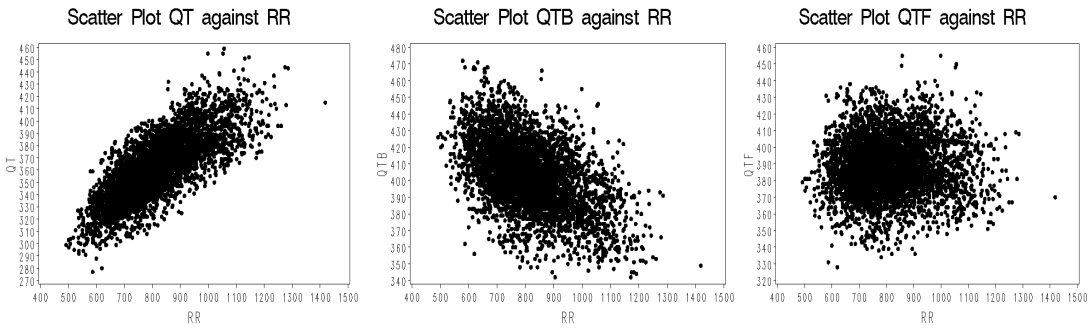


Figure 3. Scatter Plot for QT, QTB, QTF against RR in milliseconds

In this study a correction using the linear derived method (QTcLD) was used. To get QTcLD the following steps were performed:

1. The data from placebo group in all phase except moxifloxacin period on day 1 and the data from drug X group in screening and washout phase were used to estimate a model: $QT_{ij} = a + bRR_{ij} + c_i + \varepsilon_{ij}$, (i)

   where:

   i = 1,…, n is an index for subjects,

   j = 1,…, 9 is an index for time points,

   QT_{ij} is the QT interval at time point j of subject i,

   a is the intercept,

   b is the regression coefficient of RR,

   c_i is a subject specific to take into account correlation within subject.

   ε_{ij} is random disturbance with $\varepsilon \sim N(0, \sigma^2)$

2. Then QTcLD was calculated using the estimate $\hat{b}$ from model (i) to the formula: $QTcLD = QT + b(1 - RR)$

   In this study, the value of b was found equals 0.144.

Scatter plot of QTcLD against length of heart rate (RR) in figure 4 shows no correlation between them.

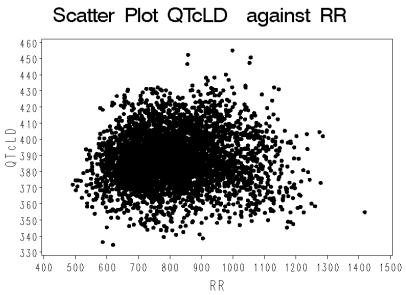


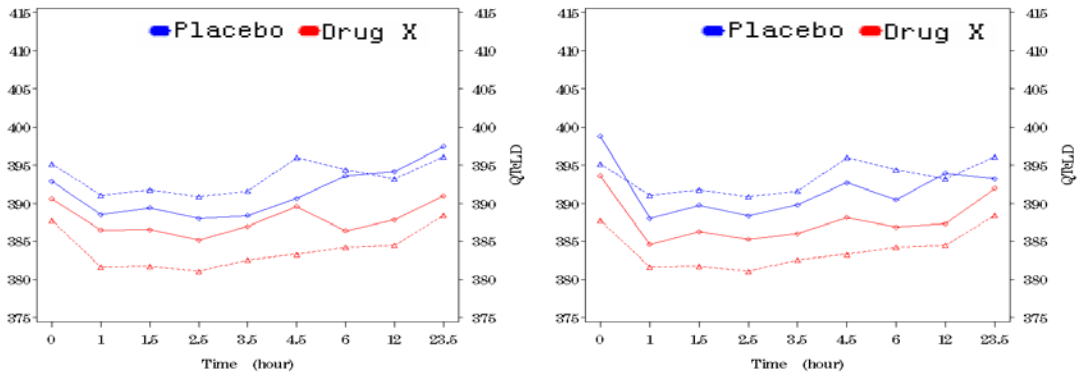
Figure 4. Scatter Plot for QTcLD against RR in milliseconds

2.6 Exploratory Data Analysis

The main focus of this thesis is to determine whether the effect on QTc interval at steady state is comparable between 12 mg drug X and placebo. The secondary focus is between18 mg drug X and placebo. It can be seen from figure 5(a) that the mean QTcLD interval for the placebo group at day 6 (blue solid line) was longer than that of the drug X group (red solid line) thorough the day. Similar situation occurred for day 11 as shown in figure 5(b).

However it is noteworthy that both groups had different QTcLD interval at the beginning of study (baseline). At baseline, placebo subjects (blue dashes line) had higher mean QTcLD interval than subjects in the drug X group (red dashes line). This led to the need of attention to the baseline values and covariates.

Diurnal variation was observed for both treatment groups in day 6 and 11. It is due to duration of the QT interval varies widely during normal daily activities and one of the determinants contribute to this variation is heart rate.



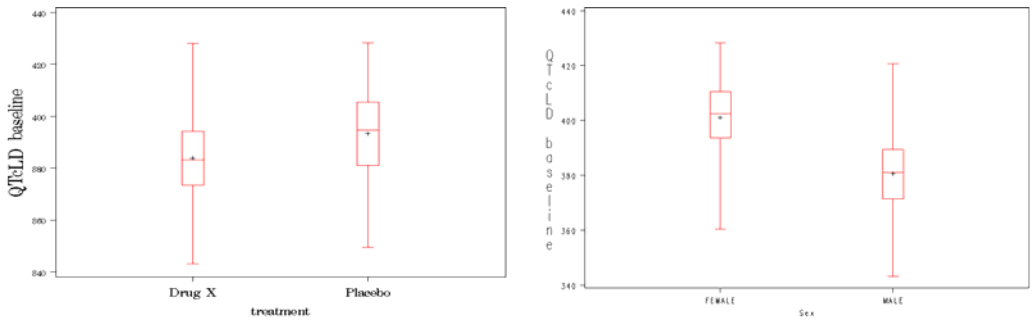
(a) Day 6 *(b) Day 11*

Figure 5. Mean profile of QTcLD for Drug X and Placebo (solid line) and their baseline (dashes) in each day

2.7 Baseline

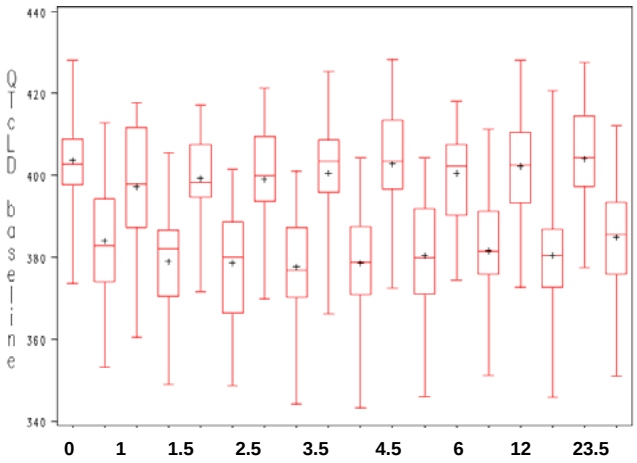
At the baseline the overall mean QTcLD interval for the treatment (drug X) group was lower than for the placebo (Figure 6(a)). Figure 7 shows that this condition occurred in all time points at day 6 and day 11. It can also be observed that there was variability for mean of QTcLD interval over time point. Females are expected to have longer QT interval than males, which was observed in this study (figure 6(b) and figure 7).

The fact that drug X group has a shorter QTcLD interval might be explained by a gender imbalance since the number of male subjects, who have shorter QTcLD interval, was higher in drug X group than in placebo group (Table 3).



(a) *(b)*

Figure 6. Box-plots for The Overall Mean of QTcLD Interval at baseline against (a) Treatment and (b) sex



** First box plot in each time point presenting Females, and the second for Males*

Figure 7. Box-plots for QTcLD Interval at baseline for each Sex group

Table 3. Number of Subjects by Sex and Treatment group

Treatment	Sex		Total
	Female	Male	
Drug X	10	34	44
Placebo	11	11	22
Total	21	45	66

At baseline there was a larger than 10 ms overall mean QTcLD interval difference between the drug X and placebo (Table 4). However, when we observed by Sex, there was no noticeable QTcLD interval mean difference between the drug X and placebo for both males and females.

Table 4. Mean (milliseconds) and standard deviation of QTcLD interval by Days, Treatment, and Sex

Days	Placebo			Drug X		
	Male	Female	Overall	Male	Female	Overall
Baseline	383.31 (13.64)	403.41 (12.14)	393.36 (16.35)	379.68 (13.63)	398.38 (14.66)	383.93 (15.92)
6	382.73 (14.69)	400.19 (15.06)	391.46 (17.22)	383.62 (15.99)	400.55 (18.07)	387.85 (18.07)
11	382.64 (14.32)	400.98 (13.41)	391.71 (16.62)	382.86 (14.89)	402.50 (18.70)	387.79 (18.04)

2.8 Time Point

Table 4 shows that at the treatment phase (days 6 and 11), the drug X group and the placebo in overall and by gender had similar mean QTcLD interval. However when we look at figure 5 we can see that for some time points they were different. In the mean profiles shown in figure 8, the variability between time points (diurnal variation) in figure 5 also occurred for each gender by treatment group. It suggested that besides sex, variability due to time point also need to be considered.

Moreover, figure 1 in the appendix shows that even though the data was split up by Sex and time point, there were still differences larger than threshold (5 milliseconds) for QTcLD interval at baseline, i.e. for female at time point 6. It suggested including baseline and time points in the analysis and adjusting for sex.

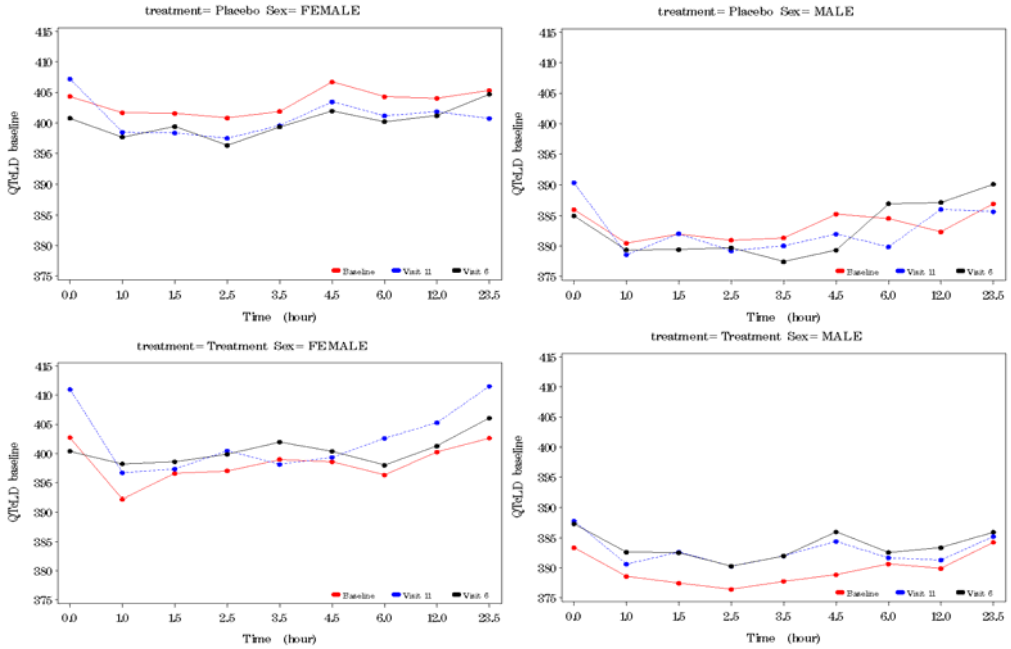


Figure 8. Mean profile of QTcLD Interval at Baseline, Day 6 and 11 by Treatment group and Sex

3. Methodology

3.1 Baseline Measurements

Since a single measurement will not be sufficient to establish the true condition of subjects as in many clinical trials the discussed study had more than one baseline. The measurement in day –2 and day –1 were considered as baseline to reduce measurement error. Then a single baseline time point matched was obtained from averaging two baselines (day –2 and day –1) in each time point. Baseline time point matched was done to account for diurnal variation.

There are three possible approaches can be used if we have more than one baseline. First one can use the last measurement if this is likely to be most predictive of all measured baselines. However in practice it will not be efficient unless there is no correlation between the baseline values and post treatment values. The second approach is to use the mean of all baseline measurements as the baseline value. This is a reasonable strategy if the variance of the baseline measurements is equal and if the correlations of all baseline values with post treatment values are the same. Third approach, we could fit all baselines simultaneously in the model. This last approach can be used if there are appreciable differences between the various correlations between baselines and pot treatment values. (Senn, Stevens, Chaturvedi, 2000).

3.2 Statistical Analysis

As mentioned in the exploratory data analysis part we have baseline imbalances and repeated measurements in this study. Hence it needs an adjusting in the statistical analysis for baseline variables in the repeated measures analysis. A summary measure analysis is commonly used in randomized clinical trials that have measurements before (baseline) and after commencing treatment (post). Matthews *et al*. (1989), and Senn, Steven, and Chaturvedi (2000) proposed and discussed the use of summary measures to analyze repeated measures. There are several summary measures available. However to avoid any temptation to choose a particular summary measure because it shows maximal differences between groups, summary measures have to be chosen before the data are collected. In this thesis we intend to find the most appropriate summary measures can be used. Therefore four possible approaches were performed and compared in two parts.

First, a single analysis is performed for observation at 4 hour 30 minutes after dosing in each day (time point 4.5). It is for simplicity and since the data shows that at this time point the largest differences were occurred. Then repeated measures analysis will be performed to account for all time point measurements. Since it was observed that Sex might confound the result, this variable will be included as the second model for each analysis.

The repeated measures analysis was performed using Proc MIXED in SAS® with satterwhite degree of freedom. Several variance covariance structures were compared using likelihood ratio test to get the most appropriate structure.

This study was not powered to see the difference since small number of sample size. Hence we were performed only exploratory study. As mentioned before the threshold level of regulatory concern of the mean difference is around 5 ms. It is evidenced by an upper bound of the one-sided 95% confidence of mean difference on QT interval of below10 ms. To see this clinically difference a two-sided 90% confidence interval for each mean difference in all models will be presented since the upper bound of it will be the same the upper bound for one-sided 95% confidence interval

The results of single analysis and repeated measure analysis for unadjusted (post treatment only) and adjusted (change from baseline and analysis of covariance) were compared as a sensitivity analysis as suggested by CPMP document (2003) for linear models.

3.2.1 Post Treatment Values Only (POST)

It has been proposed that, if the intra-individual variation and/or measurement error is large, using post-treatment values only can improve the precision and study power (Blomqvist and Dahlén, 1985), because the baseline values are supposed to be balanced. Analysis of Variance (ANOVA) was used to test differences in means of post-treatment values between the treatment (drug X) and the placebo group. Models developed were as follows:

Single Time point (4.5)

Model 1: $Y_{ij} = \mu + \tau_j + \varepsilon_{ij}$ (ii)

Model 2: $Y_{ijk} = \mu + \tau_j + \pi_k + (\tau\pi)_{jk} + \varepsilon_{ijk}$ (iii)

All time point (Repeated measures)

Model 3: $Y_{ijl} = \mu + b_i + \tau_j + \delta_l + (\tau\delta)_{jl} + \varepsilon_{ijl}$ (iv)

Model 4: $Y_{ijkl} = \mu + b_i + \tau_j + \pi_k + \delta_l + (\tau\pi)_{jk} + (\tau\delta)_{jl} + \varepsilon_{ijkl}$ (v)

where:

i = 1,…, n_j is an index for subjects within treatment j,

j = 1,2 is an index for treatment group,

k = 1,2 is an index for sex group,

l =1,…,9 is an index for time point,

Y_{ijkl} is the response observed in time point l, treatment j for patient i with gender k,

μ : the overall mean,

τ_j : the effect of treatment j,

π_k : the effect of gender k,

$(\tau\pi)_{jk}$: the effect for group $j \times$ gender k interaction,

δ_l : the effect for time point l,

$(\tau\delta)_{jl}$: the effect for group $j \times$ time point l interaction,

b_i : subject specific,

ε_{ijkl} : random disturbance term with expectation zero.

3.2.2 Change from Baseline (CHANGE)

The change from baseline value was derived by subtracting post-treatment values from the pre-treatment (baseline) values. By subtracting the post treatment values with the baseline values the variability due to diurnal variation and baseline values difference were expected can be taken into account. ANOVA was used to test the differences in the means of change value between drug x and placebo groups. Model developed are as follows:

Single Time point (4.5)

Model 1: $Y_{ij} - Y_{0ij} = \mu + \tau_j + \varepsilon_{ij}$ (vi)

Model 2: $Y_{ijk} - Y_{0ijk} = \mu + \tau_j + \pi_k + (\tau\pi)_{jk} + \varepsilon_{ijk}$ (vii)

All time point (Repeated measures)

Model 3: $Y_{ijl} - Y_{0ijl} = \mu + b_i + \tau_j + \delta_l + (\tau\delta)_{jl} + \varepsilon_{ijl}$ (viii)

Model 4: $Y_{ijkl} - Y_{0ijkl} = \mu + b_i + \tau_j + \pi_k + \delta_l + (\tau\pi)_{jk} + (\tau\delta)_{jl} + \varepsilon_{ijkl}$ (ix)

Other definitions are the same with models in POST with an additional information: Y_{0ijkl} is the baseline value in time point *l*, treatment *j* for patient *i* with gender *k*,

3.2.3 Analysis of Covariance (ANCOVA)

In this approach, between subjects variation in the baseline measurements are taken into account by using the mean baseline measurements for each subjects as a covariate in a linear model for treatment comparison of the post treatment mean. This approach is actually known in statistical environment as Analysis of Covariance, although it is not always called this explicitly within the clinical literature, (Tu Y.K, et.al, 2005). The similar models with previous approaches were developed as follows:

Single Time point (4.5)

Model 1: $Y_{ij} = \mu + \tau_j + \gamma_j + \varepsilon_{ij}$ (x)

Model 2: $Y_{ijk} = \mu + \tau_j + \gamma_j + \pi_k + (\tau\pi)_{jk} + \varepsilon_{ijk}$ (xi)

All time point (Repeated measures)

Model 3: $Y_{ijl} = \mu + b_i + \tau_j + \gamma_j + \delta_l + (\tau\delta)_{jl} + \varepsilon_{ijl}$ (xii)

Model 4: $Y_{ijkl} = \mu + b_i + \tau_j + \gamma_j + \pi_k + \delta_l + (\tau\pi)_{jk} + (\tau\delta)_{jl} + \varepsilon_{ijkl}$ (xiii)

Other definitions are the same with previous models with additional information: where γ_j is the effect of baseline.

3.2.4 Within Comparison

This last approach was rather different with the previous approaches. This approach was carried out to perform simple analysis to see the treatment effect only in the treated group. With within comparison we neglected the placebo and considered the observation on screening period of drug X as a control for the post drug X observation. In this approach only observations from drug X subjects were used. QTcLD value at baseline and value after treatment phase are considered as a control group and treatment group respectively. The paired *t* test is used to test the differences in the means of value between treatment at post-baseline phase and control groups in baseline phase and ANOVA for more than one covariate.

Single Time point (4.5)

Model 1: $Y_{ij} = \mu + \tau_j + \varepsilon_{ij}$ (xiv)

Model 2: $Y_{ijk} = \mu + \tau_j + \pi_k + (\tau\pi)_{jk} + \varepsilon_{ijk}$ (xv)

All time point (Repeated measures)

Model 3: $Y_{ijl} = \mu + b_i + \tau_j + \delta_l + (\tau\delta)_{jl} + \varepsilon_{ijl}$ (xvi)

Model 4: $Y_{ijkl} = \mu + b_i + \tau_j + \pi_k + \delta_l + (\tau\pi)_{jk} + (\tau\delta)_{jl} + \varepsilon_{ijkl}$ (xvii)

where:

i = 1,…, n_j is an index for subjects within group j,

j = 1,2 is an index for group (before and after treatment),

the other definitions are similar with previous approaches.

3.2.5 Model Comparison

In term of a single time point several papers, i.e by Senn (2004) and Vickers and Altman (2001), discussed and compared these first three approaches. These three approaches measure the same thing and there is no question of choosing between them on the basis of clinical relevance. The three approaches can be summarized in table 5. One can choose between them on the basis of either of variance or statistical philosophy. We can see in the treatment effect estimator in table 5 that POST and CHANGE are special case of ANCOVA. When $\beta = 1$ then CHANGE will the same with ANCOVA. When $\beta = 0$, the POST is the same as ANCOVA. Where β is the regression coefficient of baseline value (Y_{0i}) of the regression post treatment (Y_i) on the baseline value (Y_{0i}).

Table 5. Model and Treatment Effect Estimator for POST, CHANGE and ANCOVA

Analysis	Model	Treatment Effect Estimator
POST	$Y_i = \beta_0 + \beta_x X_i$	$\hat{\tau}_{POST} = \bar{Y}_T - \bar{Y}_C$
CHANGE	$Y_i - Y_{0i} = \beta_0 + \beta_x X_i$	$\hat{\tau}_{CHANGE} = (\bar{Y}_T - \bar{Y}_C) - (\bar{Y}_{0T} - \bar{Y}_{0C})$
ANCOVA	$Y_i = \beta_0 + \beta_B Y_{0i} + \beta_x X_i$	$\hat{\tau}_{ANCOVA} = (\bar{Y}_T - \bar{Y}_C) - \beta(\bar{Y}_{0T} - \bar{Y}_{0C})$

Where:

X_i : Treatment group of subject *i*,

Y_{0i} : Baseline value of subject *i*,

Y_i : Post treatment value of subject *i*,

$\bar{Y}_T$: mean post treatment value of the drug X group,

$\bar{Y}_C$: mean post treatment value of the placebo group,

$\bar{Y}_{0T}$: mean baseline value of the drug X group,

$\bar{Y}_{0C}$: mean baseline value of the placebo group.

In the case that the average baseline values are the same in each group the estimated treatment effect will be the same using POST and CHANGE. However when there is an imbalance in baseline, for example when treatment is effective the statistical significance of the treatment effect of the treatment effect by two methods will depend on the correlation between baseline and after treatment values. If the correlation is low using then CHANGE will add variation, and using POST is more likely to show significant result. Conversely, if the correlation is high, using only the raw data (POST) we will loose information and then CHANGE is more likely to be significant, (Vickers and Altman, 2001).

ANCOVA is the better approach for following reasons:

a. It adjust after treatment result for each subject with their baseline value
b. It yields the most appropriate p-value for the treatment difference.
c. It improves the precision of the estimated treatment difference thus increasing the statistical power of the trial, (Pocock, et.al, 2002)
d. It uses a value for β which minimizes the variance and makes it independent of the baseline, (Senn, Stevens, and Chaturvedi, 2000). Hence it robust with respect to imbalance in the baseline
e. It may generally be expected to have the lowest variance than other methods as shown in figure 9, (Senn, 2004).

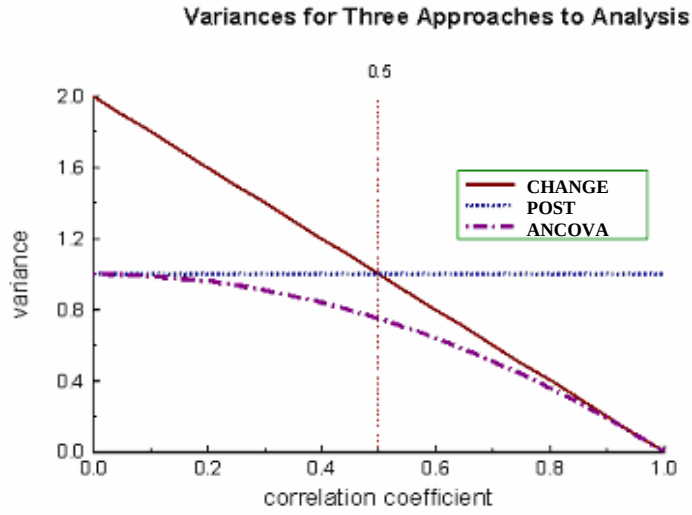


Figure 9. Variance for Three Approaches to Analysis, (Senn, 2004)

ANCOVA will have a low efficiency when there is high correlation between baseline and post treatment values. In this case, CHANGE is a reasonable alternative. However, as a general approach ANCOVA is preferred (Vickers and Altman, 2001).

Matthews *et al*. (1989), Senn, Steven, and Chaturvedi (2000), and Frison, L. and Pocock S,J (1992) proposed and discussed analyzing repeated measures using summary measure. The properties of three approaches in single point discussed above also apply in the repeated measure. Though the definition of repeated measure in their paper is rather different with what this study has. They consider days 0 as baseline and then the next dayss as repetition. It is different with this study that repetition is in each days (baseline, days 6 and days 11) and we want to compare the QTcLD interval between after treatment (days 6 and 11) with baseline. However we consider that the definition of baseline and repeated measure in their paper can be generalized for this study so that their paper argument can be used to discuss the result in this study.

4. Data Analysis and Results

Measurements in screening phase (day –2 and –1) were considered as the baseline. The QTcLD interval for day –1 and –2 seem to be the same as seen figure 2 in the appendix. The QTcLD intervals in the screening phase (day -2 and -1) were highly correlated with QTcLD in day 6 and 11 as shown in table 1 in the appendix. It also shows that correlation was similar for each baseline day. Therefore the use of mean of the baseline values (day -1 and –2) was appropriate to establish the true condition of the subjects.

Next in this chapter we will show the result of the analysis for 12 mg drug X at day 6 as the main objective and then followed by the result of the analysis for 18 mg as the secondary objective. Compound symmetry variance covariance structure was used after comparing several variance covariance structures using likelihood ratio test.

4.1 The 12 mg of Drug X versus Placebo at Day 6

4.1.1 Post-Treatment Value only (POST)

We start with only considering one measurement per day (time point 4.5). The result is shown in table 6. In the first two models, only post treatment value and post treatment value adjusted by Sex, we found no statistically difference between QTcLD interval mean for treated subjects and placebo at 5% level of significance. In addition as expected there was a large difference between QTcLD interval mean of males and females.

Then model 3 and model 4 were fitted to account for all measurements as repeated measures. When sex was not included in the model (model 3) and when it was included (model 4) we found no statistical difference between mean QTcLD intervals of two treatment groups. There were QTcLD interval mean differences across time points. However the differences for both groups were statistically similar. The inclusion of Sex (model 4) reduces the mean difference compared to when it was not included (model 3). At this day, female's mean QTcLD interval was higher than males.

The mean QTcLD interval differences between female and male subjects in the single analysis (model 2) and in the repeated measures analysis (model 4) were considered clinically difference base on the threshold mentioned in the methodology part.

In addition, the correlations between baseline and the post treatment value were high. We can observe the correlation coefficient in table 1 in the appendix.

Table 6. POST result for Day 6

Model	Parameter	P-value	Mean Difference	90% CI
Time point 4.5				
1	TrtGrp	0.8424	-1.055[a]	-9.883; 7.773
2	TrtGrp	0.6191	2.547[a]	-5.975; 11.070
	Sex	0.0006	18.523[b]	9.999; 27.046
	TrtGrp*Sex	0.4217		
All time points (Repeated)				
3	TrtGrp	0.4237	-3.604 [a]	-11.081; 3.873
	Timepoint	<0.0001		
	Trtgrp*Timepoint	0.2693		
4	TrtGrp	0.8815	0.627 [a]	-6.382; 7.636
	Timepoint	<0.0001		
	Trtgrp*Timepoint	0.2693		
	Sex	<0.0001	17.193 [b]	10.184; 24.202
	Sex* TrtGrp	0.9495		

a. QTcLD Interval Mean Difference (Drug X- Placebo)
b. QTcLD Interval Mean Difference (Female- Male)

4.1.2 Change from Baseline (CHANGE)

The same steps were also performed for the change from baseline approach and the result is presented in table 7. In model 1 the mean difference between QTcLD interval after treatment and before treatment (change from baseline value) at time point 4.5 for the drug X group was statistically higher than the placebo group. The estimated difference did not change much when adjusted by sex (model 2). The result was similar when all time points were considered. Moreover we found that the change value for both groups were different over time point. It can be seen in figure 10 that the largest change from baseline occurred at 4 hour 30 minutes after dosing (time point 4.5).

The mean change value of drug X group was statistically higher than the placebo in both single analysis on time point 4.5 and repeated measure analysis. However only the mean change value difference at the single analysis (model 1 and 2) were considered to be clinically different base on the threshold mentioned in the methodology part since their upper bound 90% confidence interval was larger than 10 ms.

The 90% confidence interval of mean difference (Drug X - placebo) for all time points adjusted for Sex is listed table 2 in the appendix. The table shows that the difference only occurred at some time points, especially at time point 4.5, which had the largest difference.

Table 7. CHANGE result for Day 6

Model	Parameter	P-value	Mean Difference	90% CI
Time point 4.5				
1	TrtGrp	0.0013	10.787	5.435; 16.140
2	TrtGrp	0.0064	9.607	3.936; 15.278
	Sex	0.7699		
	TrtGrp*Sex	0.2707		
All time points (Repeated)				
3	TrtGrp	0.0235	4.045	1.139; 6.949
	Timepoint	0.9284		
	Trtgrp*Timepoint	0.0412		
4	TrtGrp	0.0444	3.816	0.712; 6.920
	Timepoint	0.9284		
	Trtgrp*Timepoint	0.0402		
	Sex	0.4800		
	Sex* TrtGrp	0.8273		

QTcLD Interval Mean Difference: Drug X- Placebo

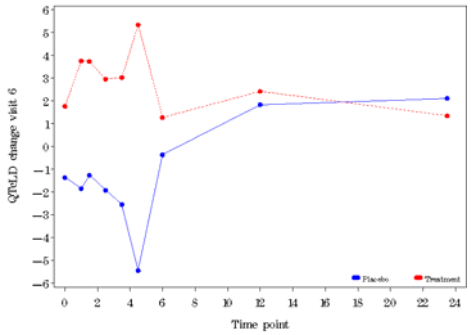


Figure 10. Change value for Day 6

4.1.3 Analysis of Covariance (ANCOVA)

The next approach was adjusting the baseline value by inclusion in the model. From the results shown in table 8 at time point 4.5 when adjusting only by the baseline value (model 1) we found that the mean QTcLD interval for drug X group was higher than placebo. The baseline value statistically had significant effect to the post treatment values. The QTcLD intervals seen in the post treatment depended on their QTcLD intervals at the beginning of study. Model 2 was adjusted for sex and its interaction but this did not change the result much.

When all time points were included as repeated measurements, in the models without Sex (model 3) and with Sex as a covariate (model 4) we found no significant effect of drug X on the QTcLD interval. In addition female subjects had significant higher mean QTcLD interval than male subjects. After adjusting for the baseline value, the 12mg of

the drug X have no effect to the QTcLD interval. This confirmed what was found in exploratory data analysis that placebo have longer QTCLD interval is due to more females subjects in this group.

Table 8. ANCOVA result for Day 6

Model	Parameter	P-value	Mean Difference	90% CI
Time point 4.5				
1	TrtGrp	0.0063	9.5304[a]	3.916; 14.279
	Baseline	<0.0001		
2	TrtGrp	0.0139	8.8038[a]	3.007; 15.145
	Baseline	<0.0001		
	Sex	0.7580		
	TrtGrp*Sex	0.2700		
All time points (Repeated)				
3	TrtGrp	0.8189	0.588[a]	-3.731; 4.907
	Baseline	<0.0001		
	Timepoint	0.2110		
	Trtgrp*Timepoint	0.09552		
4	TrtGrp	0.4356	2.1473[a]	-2.4468; 6.7292
	Baseline	<0.0001		
	Timepoint	0.1019		
	Trtgrp*Timepoint	0.1108		
	Sex	0.0060	8.4047[b]	3.5185; 13.2908
	Sex* TrtGrp	0.9845		

a. *QTcLD Interval Mean Difference (Drug X- Placebo)*
b. *QTcLD Interval Mean Difference (Female- Male)*

4.1.4 Within Comparison

Only observations from subject assigned to drug X were used in this approach and group effect was defined as after and before treatment. Using a paired t-test for time point 4.5 we found that the mean QTcLD interval after administering 12 mg of drug X was statistically different from mean QTcLD interval before treatment (model 1). It can be seen in table 9 that the mean QTcLD interval for females and males are statistically different. However the interaction between group and sex was not significant. Therefore the effect of drug X was similar for female and male subjects. This was also seen when repeated the measures analysis was performed on all time points (model 3 and 4). In addition, even though the mean QTcLD interval was statistically different over the time points, the effect of treatment was similar across time point.

Table 9. Within Comparison for Day 6

Model	Parameter	P-value	Mean Difference	90% CI
Time point 4.5				
1	Group	0.0055	5.346[a]	2.290; 8.410
2	Group	0.0483	4.238[a]	0.736; 7.740
	Sex	0.0036	17.308[b]	7.872; 26.744
	Group *Sex	0.2489		
All time points (Repeated)				
3	Group	<0.0001	2.849[a]	1.962; 3.737
	Timepoint	<0.0001		
	Group * Timepoint	0.7190		
4	Group	<0.0001	2.621[a]	1.596; 3.646
	Timepoint	<0.0001		
	Group * Timepoint	0.7194		
	Sex	0.0014	17.382 [b]	8.861; 25.904
	Sex* Group	0.4635		

a. QTcLD Interval Mean Difference (Before Drug X- After Drug X)

b. QTcLD Interval Mean Difference (Female- Male)

4.2 The 18 mg of Drug X versus Placebo at Day 11

4.2.1 Post Ttreatment Values Only (POST)

Similar results as day 6 were found for day 11 at time point 4.5 as shown in table 10. No statistical difference between treatment groups and no effect of sex was found. In the repeated measurement analysis of all time points, no group difference was found and there was a gender effect to the result. We found that the mean QTcLD interval at day 11 had significant diurnal variation however the effect of treatment was similar across time point. Similar as observed in day 6, the mean QTcLD interval was higher for females than for males. Similar with day 6, the correlations between baseline and the post treatment value at day 11 were high. We can observe the correlation coefficient in table 1 in the appendix.

The 90% CI of mean difference (drug X - placebo) for all time points adjusted for Sex is listed table 2 in the appendix. The table shows that the mean QTcLD interval different only occurred at some time points, especially at time point 4.5, which had the largest difference.

4.2.2 Change from Baseline (CHANGE)

The change from baseline value for drug X group was statistically different with placebo in the single analysis and when all time points were considered. The result is presented in table 11. However, only the mean change value differences at single

analysis (model 1 and 2) were considered to be clinically different. In contrast with day 6, the differences were similar over time point since the interaction with time point was not significant. It also can be observed in figure 11 that Sex statistically had no effect to the change from baseline values.

Table 10. POST result for Day 11

Model	Parameter	P-value	Mean Difference	90% CI
Time point 4.5				
1	TrtGrp	0.3348	-4.574[a]	-12.441; 3.293
2	TrtGrp	0.8496	-0.837 [a]	-8.189; 6.517
	Sex	0.0001	18.229[b]	10.877; 25.583
	TrtGrp*Sex	0.4585		
All time points (Repeated)				
3	TrtGrp	0.3516	-4.154 [a]	-11.551; 3.244
	Timepoint	<0.0001		
	Trtgrp*Timepoint	0.8803		
4	TrtGrp	0.8516	0.739 [a]	-5.846; 7.325
	Timepoint	<.0001		
	Trtgrp*Timepoint	0.8810		
	Sex	<0.0001	19.119 [b]	12.534; 25.705
	Sex* TrtGrp	0.9071		

a. QTcLD Interval Mean Difference (Drug X- Placebo)

b. *QTcLD Interval Mean Difference (Female- Male)*

Table 11 . CHANGE result for Day 11

Model	Parameter	P-value	Mean Difference	90% CI[a]
Time point 4.5				
1	TrtGrp	0.0035	8.237	3.723; 12.750
2	TrtGrp	0.0192	6.879	2.114; 11.645
	Sex	0.5038		
	TrtGrp*Sex	0.2224		
All time points (Repeated)				
3	TrtGrp	0.0288	4.422	1.127; 7.715
	Timepoint	0.1226		
	Trtgrp*Timepoint	0.6510		
4	TrtGrp	0.0366	4.5480	0.9978; 8.0981
	Timepoint	0.1226		
	Trtgrp*Timepoint	0.6511		
	Sex	0.9817		
	Sex* TrtGrp	0.7998		

QTcLD Interval Mean Difference: Drug X- Placebo

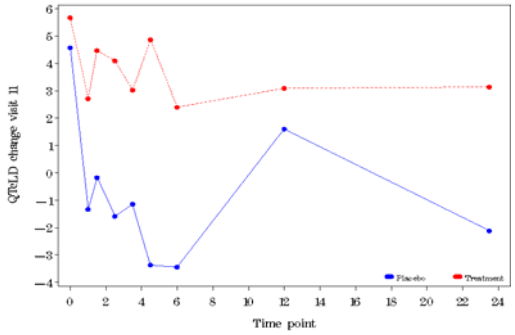


Figure 11. Change value for Day 11

4.2.3 Analysis of Covariance (ANCOVA)

After adjusted for baseline value, as shown in table 12 the drug X group was found to had higher mean QTcLD interval than the placebo at time point 4.5. Adding Sex and its interaction with treatment group as covariates to the model led to smaller mean QTcLD interval difference, which was not statistically significant. In the repeated measurements analysis this situation was changed. Considering all time point with and without inclusion of sex in the model, no effect of treatment was observed. In model 3 and 4, the mean QTcLD intervals were different over time point. However the effect of treatment was similar over time point since the interaction between treatment and time point was not statistically different. In addition, females had higher mean QTcLD intervals than male subjects.

4.2.4 Within Comparison

The paired t-test shows that for treated subjects at time point 4.5 the mean QTcLD interval after administering 18 mg of drug X was statistically different from the mean QTcLD interval at baseline. The result is presented in table 13. Inclusion Sex in the analysis led to a smaller treatment mean difference. The mean QTcLD intervals for females and males were significantly and clinically different (model 2). Since there was no significant interaction between gender and treatment, the effect of treatment was similar for female and male subjects. When repeated measures analysis was performed (model 3 and 4), we found that both treatment and sex had effect on the QTcLD interval. The effect of treatment was similar across time point and Sex.

Table 12. ANCOVA result for Day 11

Model	Parameter	P-value	Mean Difference	90% CI[a]
Time point 4.5				
1	TrtGrp	0.0378	5.760[a]	1.236; 10.286
	Baseline	<0.0001		
2	TrtGrp	0.0676	5.171[a]	0.532; 9.809
	Baseline	<0.0001		
	Sex	0.4263		
	TrtGrp*Sex	0.2044		
All time points (Repeated)				
3	TrtGrp	0.8701	0.424	-3.934; 4.781
	Baseline	<0.0001		
	Timepoint	0.0021		
	Trtgrp*Timepoint	0.8438		
4	TrtGrp	0.3484	2.472	- 1.919; 6.862
	Baseline	<0.0001		
	Timepoint	0.0007		
	Trtgrp*Timepoint	0.8627		
	Sex	0.0002	10.417[b]	5.678; 15.156
	Sex* TrtGrp	0.8478		

a. *QTcLD Interval Mean Difference (Drug X- Placebo)*

b. *QTcLD Interval Mean Difference (Female- Male)*

Table 13. Within Comparison for Day 11

Model	Parameter	P-value	Mean Difference	90% CI
Time point 4.5				
1	Group	0.0060	4.867[a]	2.057; 7.082
2	Group	0.0745	3.436[a]	0.278; 6.593
	Sex	0.0016	17.096[b]	8.578; 25.614
	Group *Sex	0.1659		
All time points (Repeated)				
3	Group	<0.0001	3.722[a]	2.749; 4.695
	Timepoint	<0.0001		
	Group * Timepoint	0.9250		
4	Group	<0.0001	3.848[a]	2.725; 4.970
	Timepoint	<0.0001		
	Group * Timepoint	0.9252		
	Sex	0.0007	19.325 [b]	10.591; 28.059
	Sex* Group	0.7119		

a. QTcLD Interval Mean Difference (Before Drug X- After Treatment)

b. QTcLD Interval Mean Difference (Female- Male)

5. Discussion

The main objective of this study was to assess the effect on QTcLD interval of 12 mg of drug X at steady state compared to the placebo. In the exploratory data analysis part, we discovered the imbalances at the baseline: the placebo group had higher mean QTcLD interval than drug X group. Four different approaches were performed with two parts. First step we focused only on the observations on 4 hours 30 minutes after dosing (time point 4.5), for simplicity and data driven. Then we continued with analyzing the values obtained from all time points. This last step was considered as repeated measures analysis. Four models were fitted in each approach.

Using only post treatment value at time point 4.5 we found neither in Day 6 nor Day 11 a statistically significant mean QTcLD interval difference between the treatment groups. From the repeated measures analysis, the mean QTcLD interval difference was also not significant and from the last model was found equals 0.627 and 0.739 for days 6 and 11 respectively. However using only the post treatment value (POST), we will loose information due to baseline imbalance and the fact that correlations between baseline and the post treatment value at days 6 and days 11 were high.

Conversely when the change value from baseline (post-baseline) was used, the mean change value at time point 4.5 was higher for drug X subjects than placebo in day 6 and 11. Adjusting for sex had small effect on the result since it was already incorporated the baseline value. Considering all time points revealed that the change from baseline value was different in the two groups. The mean difference of change value in the last model of repeated measure analysis equals 3.82 and 4.55 for day 6 and 11 respectively. Moreover only in day 6 that the differences varied over time point. However this result was expected. Since when there is imbalance in baseline and the correlation between baseline and the post treatment value is high, CHANGE is more likely to be significant. In gender, female and male subjects had similar change value in day 6 and 11.

The next approach was by adjusting the post treatment value by inclusion the baseline value in the model known as ANCOVA. After adjusted for baseline value the mean QTcLD interval at time point 4.5 was higher for drug X group than placebo at both day 6 and 11. Conversely when all time points were taken into account, the mean QTcLD intervals for both treatment groups were similar. The mean difference of QTcLD

interval in the last model of repeated measure analysis equals 2.15 and 2.47 for day 6 and 11 respectively. We found also that females had higher mean QTcLD interval than males.

In within comparison approach, the mean QTcLD intervals after treatment at day 6 and day 11 was higher than before treatment both the single analysis at time point 4.5 and the repeated measures analysis. The mean difference of QTcLD interval before treatment and after treatment in the repeated measures analysis equals 2.62 and 3.85 for day 6 and 11 respectively (model 4). The effect of treatment was similar across time point and Sex.

As mentioned in the methodology part, the CHANGE analysis is a special case of ANCOVA: i.e. when the regression coefficient of baseline value of regression post treatment on baseline value (β) equals 1 then CHANGE is the same with ANCOVA. It can be seen for this study in table 4 and 5 in the appendix that the mean difference obtained from CHANGE and ANCOVA tend to be similar as the β is close to 1. The standard errors of estimates using ANCOVA were bit higher than CHANGE for some time points that had high correlation. It showed that ANCOVA has low efficiency when there is high correlation between baseline and after treatment value.

The change from baseline was expected can be account for variability due to diurnal variation (time point). However in day 6, the interaction between treatment group and time point was statistically significant. It indicated that the change values are different over all time points in day 6. Hence variability due to time point still needed to take into account. It is the same with ANCOVA. It needed to be adjusted for baseline covariates since the baseline covariates were had statistically significant effect to the QTcLD interval after treatment. Hence both CHANGE and ANCOVA need to be adjusted for the possible influent covariates.

The within comparison approach using paired t-test would be simpler, but we lost information by neglecting the placebo group. Hence ANCOVA is better, because it uses all available information and allows us to observe whether the value of QTcLD interval after treatment depends on the QTcLD interval at baseline.

As we can see from the result that adjusting covariates generally improves the efficiency of the analysis either in single point or in repeated measures. It can be observed in narrower confidence interval for CHANGE and ANCOVA compared with

POST. Moreover the confidence interval for CHANGE was narrower than ANCOVA. It was expected since it also occurred in the single point analysis and also because of high correlation between baseline and post treatment value.

This thesis discusses the effect of imbalance in baseline to the analysis and also compares the result of four possible approaches that can be used. The baseline imbalances led to different results yielded by the four discussed approaches. Hence the appropriate method should be chosen. Considering the facts discussed previously, ANCOVA seems to be the best approach that can be chosen in advance for randomized clinical trials. ANCOVA is generally expected to have the lowest variance than other methods and also it robust with respect to imbalance in the baseline. If the regression coefficient of baseline value on regression post value on baseline values is known in advance to be close to one, CHANGE can be used.

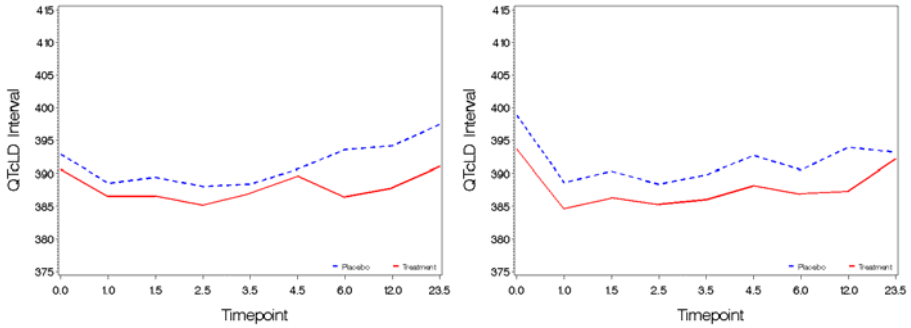
The CPMP document (2003) recommends that baseline imbalance should not be considered as an appropriate reason to include a baseline measure as a covariate. ICH E9 guideline advises to identify the covariates expected to have important influence on the primary outcome and to specify how to account for them in the analysis. It is also mentioned that the adjustment for the influence of covariates is an integral part of the planned analysis and hence should be set out in the protocol. Hence covariates to be included in the primary analysis must be pre-specified in the protocol or in statistical analysis plan. Senn (1994) suggest the covariates are chosen when trial is being design on the basis of previous studies covariates of prognostics value. The prognostic covariate should be fitted in an analysis of covariance or equivalent technique whatever the degree of imbalance.

6. Conclusion and Recommendation

6.1 Conclusion

From the four discussed approaches we found that using only data after treatment (POST) and change from baseline (CHANGE) can be regarded as special cases of summary statistics of the form: $S_i = Y_i - \beta Y_{oi}$. They depend on the baseline and the correlation between baseline and after treatment values. A simpler comparison such as within comparison might be simpler, but we lost information by neglecting the placebo and we cannot observe whether the result depend on the value at baseline or not. ANCOVA seems to be the best approach that can be chosen in advance for randomized clinical trials. It is a more general model of POST and CHANGE. It uses a value for β to minimize the variance of S_i and make it independent of the baseline.

From ANCOVA result it can be concluded that neither 12 mg nor 18 mg drug X at steady state have an effect on the prolongation of the length of the QT interval corrected for heart rate using linear derived method (QTcLD). The estimated mean profiles are shown in figure12. Imbalance between groups in baseline was influent the results. Hence adjusting in the statistical analysis for the baseline values and covariates was performed.



(a) Day 6 *(b) Day 11*

Figure 12. Estimated Mean Profile of QTcLD Interval for Placebo (blue dash line) and Drug X (solid red line) in day 6 and 11

6.2 Recommendation

To avoid imbalance in baseline, stratified randomization by baseline characteristics should be implemented. These baseline characteristics should be included in the analysis. The choice of baseline characteristics by which an analysis is adjusted should be determined by prior knowledge of an influence on outcome rather than evidence of imbalance between treatment groups in the trial.

Reference

Blomqvist N and Dahlén G. Analysis of Change: Are Baseline Measurements Needed? Some Statistical Comments on Common Experimental Design. *J Clin Periodontol* 1985; 12:877–881.

European Medicines Agency. *ICH Topic E9 Note for Guidance on Statistical Principles for Clinical Trials*; September 1998.

Frison L, Pocock SJ. Repeated Measures in Clinical Trials: Analysis Using Mean Summary Statistics and its Implication for Design. *Statistics in Medicine* 1992; 2:1685-1704.

Huck, S.W. and McLean, R.A. Using a repeated measures ANOVA to analyze data from a pretest-posttest design: A potentially confusing task. *Psychological Bulletin* **1975**; 82**:**511–518.

Jindal R, et.al. Cardiac Risk and Schizophrenia. *Journal Psychiatry Neurosci* 2005; 30(6):393-5.

Matthews JNS, et.al, Analysis of Serial Measurements in Medical Research. *British Medical Journal* 1990; 300:230-235.

Pocock SJ, et al. Subgroup Analysis, Covariate Adjustment and Baseline Comparison in Clinical Trial Reporting: Current Practice and Problem. *Statistics in Medicine* 2002; 21:2917-2930

Robert C, Torgerson DJ. Understanding Comparator led Trials: Baseline Imbalance in Randomized Comparator led Trials. *British Medical Journal* 1999; 319:185.

Senn SJ. Testing for Baseline Balance in Clinical Trials. *Statistics in Medicine 1994*; 13:1715-1726.

Senn SJ. Repeated Measures in Clinical Trials: Analysis Using Mean Summary Statistics and its Implication for Design [Letter, Comment]. *Statistics in Medicine* 1994; 13:197-198.

Senn SJ, Stevens L, Chaturvedi N. Repeated Measures in Clinical Trials: Simple Strategies for Analysis Using Summary Measures. S*tatistics in Medicine* 2000; 19:861-877.

Senn SJ. *Change from Baseline or Analysis of Covariance?: Lord's Paradox and Other Matters.* www.stochastik.math.uni- goettingen.de/colloquium/Lord's%20Paradox%20 Glasgow.ppt (2004). retrieved on17th august 2007.

Senn SJ. Change from Baseline and Analysis of Covariance Redaysed. *Statistics in Medicine 2006*; 25:4334-4344.

The European Agency for the Evaluation of Medicinal Products. Evaluation of Medicines for Human Use, Committee for Proprietary Medicinal Products (CPMP). *Point to Consider on Adjustment for Baseline Covariates*; May 2003.

Tu YK, et.al. Statistical Power for Analyses of Changes in Randomized Controlled Trials. *Journal of Dental Research* 2005; 84(3):283-287.

U.S. Department of Health and Human Serviced, Food and Drug Administration, Center for Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER); *Guidance for Industry E14 Clinical Evaluation of QT/QTc Interval Prolongation and Proarrhythmic Potential for Non-Antiarrhythmic Drugs*. International Committee of Harmonization, 2005; 1-20.

Verbeke, G. and Molenberghs, G. *Linear Mixed Models for Longitudinal Data*. New York: Springer – Verlag 2000.

Vickers AJ, Altman DG. Statistics Notes: Analysis Comparator led Trials with Baseline and Follow Up Measurements. *British Medical Journal* 2001; 323:1123-1124.

Vickers AJ. The Use of Percentage Change from Baseline as An Outcome in A Comparator led Trials is Statistically Inefficient: A Simulation. *BMC Medical Research Methodology* 2001;1: 6.

Vickers AJ. How Many Repeated Measures in Repeated Measures Design? Statistical Issues for Comparative Trials. *BMC Medical Research Methodology* 2003; 3: 22.

Appendix

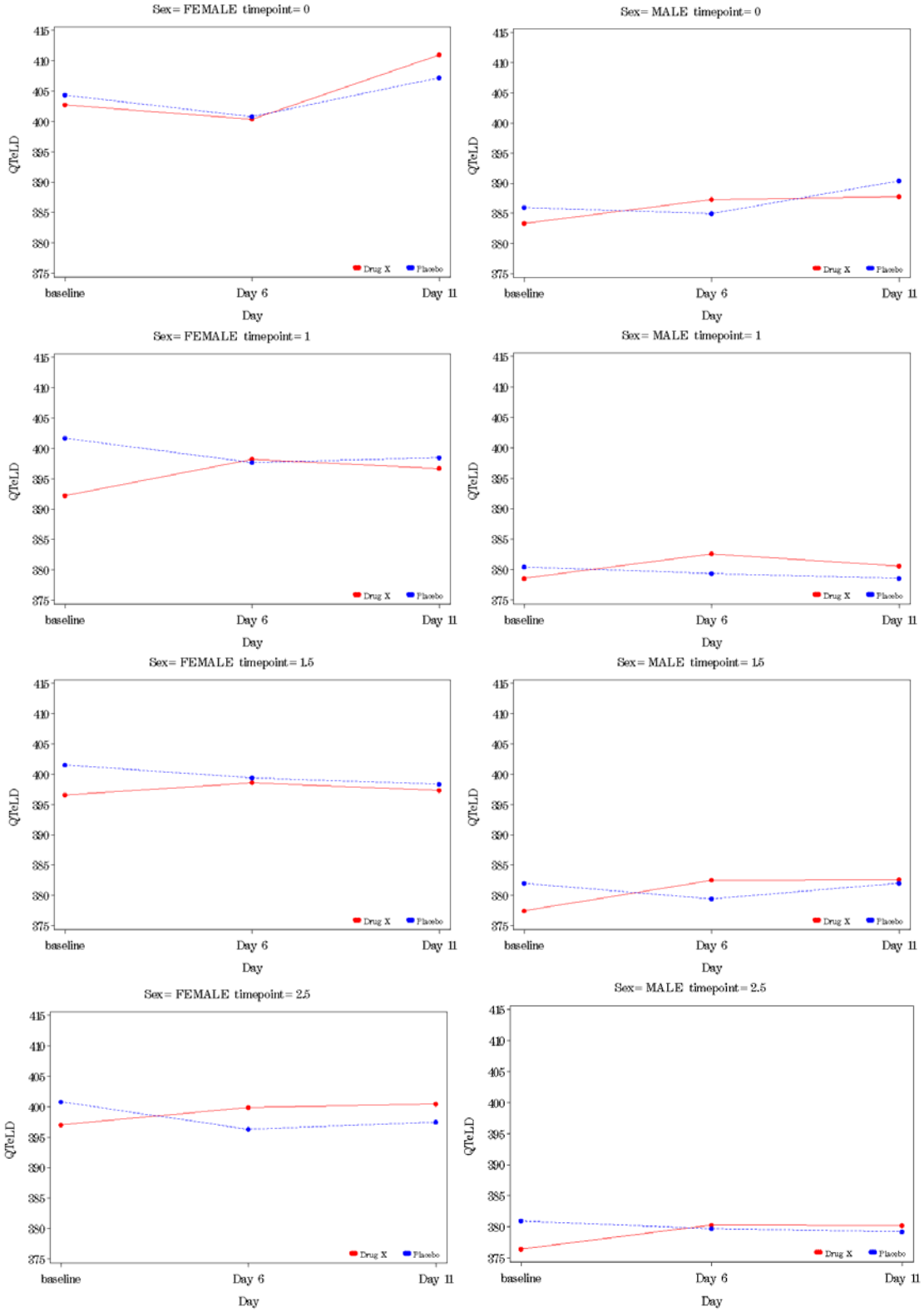


Figure 1. QTcLD Interval mean profile by sex and time point

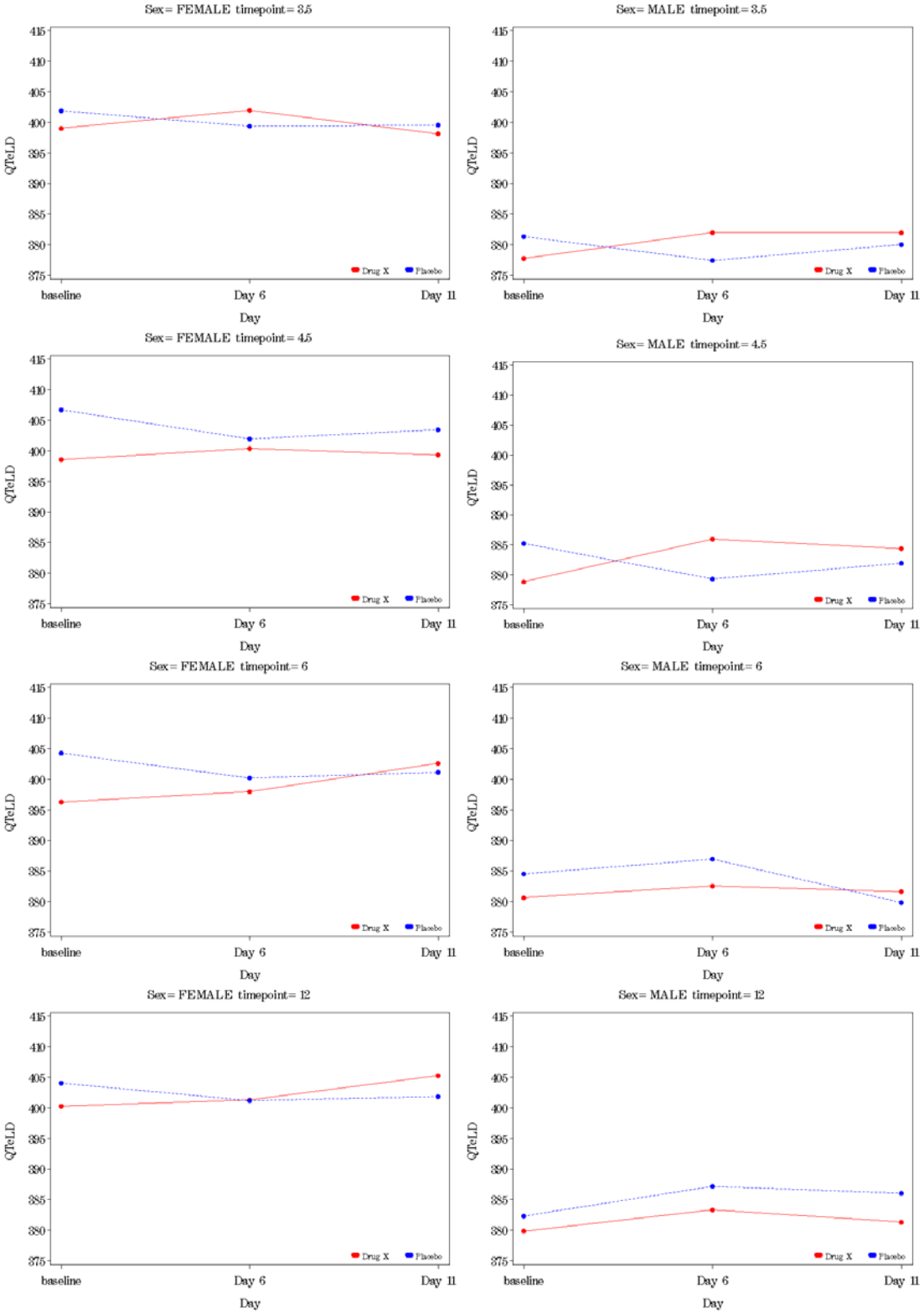


Figure 1. QTcLD Interval Mean Profile by sex and time point (continued)

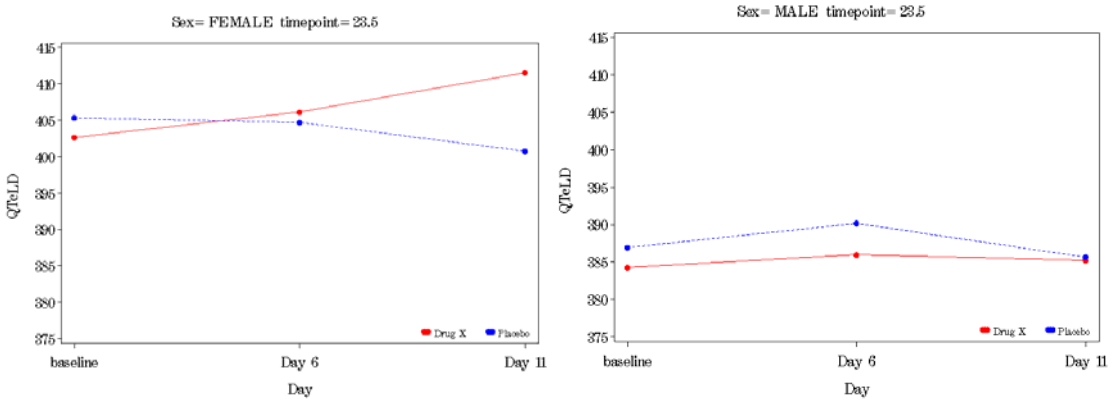


Figure 1. QTcLD Interval Mean Profile by sex and time point (continued)

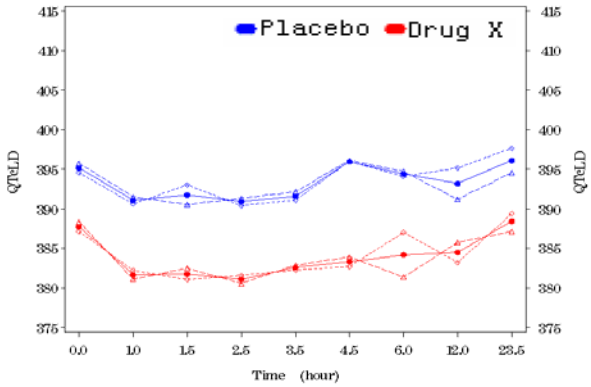


Figure 2. QTcLD Interval Mean Profile by time point for Day -1 (short dashed), Day -2 (long dashed) and their average/baseline (solid line) for each Treatment Group

Table 1. Correlations between QTcLD Interval in Day -1 and Day -2 and The Mean of Day -1 and Day -2 (Baseline) with QTcLD Interval in Day 6 and Day 11

Time point	Day 6			Day 11		
	-2	-1	Baseline	-2	-1	Baseline
0	0.81	0.83	0.84	0.76	0.73	0.769
1	0.8	0.76	0.81	0.71	0.77	0.769
1.5	0.79	0.76	0.8	0.75	0.7	0.748
2.5	0.87	0.76	0.84	0.78	0.75	0.792
3.5	0.81	0.8	0.84	0.77	0.71	0.781
4.5	0.74	0.76	0.76	0.82	0.81	0.833
6	0.77	0.77	0.8	0.81	0.78	0.824
12	0.78	0.81	0.83	0.73	0.76	0.783
23.5	0.76	0.73	0.77	0.74	0.74	0.777

Table 2. QTcLD Interval Mean difference (12 mg of Drug X - placebo), Standard error and 90% CI for all time points using CHANGE at Day 6

Time point	Lower bound	Mean diff	Upper bound	Std Err. Mean diff
0	-2.43	1.76	5.95	2.50
1	1.45	6.35	11.24	2.93
1.5	-0.87	4.42	9.71	3.17
2.5	0.32	4.83	9.35	2.70
3.5	-0.62	5.56	10.49	2.95
4.5	3.93	9.6	15.28	3.91
6	-3.36	1.76	6.9	3.07
12	-4.84	-0.11	5.07	2.97
23.5	-5.89	-0.07	5.75	3.48

Table 3. QTcLD Interval Mean difference (18 mg of Drug X - placebo), Standard error and 90% CI for all time points using CHANGE at Day 11

Time point	Lower bound	Mean diff	Upper bound	Std Err. Mean diff
0	-4.5	1.89	8.30	3.824
1	-1.14	4.59	10.34	3.426
1.5	-2.25	3.48	9.21	3.421
2.5	-0.22	5.36	10.99	3.329
3.5	-2.76	2.84	8.44	3.344
4.5	2.11	6.88	11.64	2.848
6	2.19	6.95	11.71	2.841
12	-3.34	2.51	8.36	3.494
23.5	1.63	7.13	12.63	3.283

Table 4. QTcLD Interval Mean difference (12 mg of Drug X - placebo), Standard error and 90% CI for all time points using ANCOVA and the Coefficient Regression of Baseline value **(** β **)** *at Day 6*

Time point	Lower bound	Mean diff	Upper bound	Std Err. Mean diff	Beta (β) (baseline)
0	-2.42	1.6	5.67	2.43	0.82
1	0.89	5.7	10	2.94	0.87
1.5	-0.96	4.42	9.80	3.22	0.99
2.5	0.001	4.53	9.05	2.70	0.89
3.5	0.56	5.55	10.54	2.98	0.99
4.5	3.01	8.8	14.60	3.47	0.88
6	-3.9	1.26	6.48	3.12	0.90
12	-5.1	-0.17	4.70	2.93	0.85
23.5	-6.06	-0.26	5.59	3.47	0.85

Table 5. QTcLD Interval Mean difference (18 mg of Drug X - placebo), Standard error and 90% CI for all time points using ANCOVA and the Coefficient Regression of Baseline value **(** β) *at Day 11*

Time point	Lower bound	Mean diff	Upper bound	Std Err. Mean diff	Beta (β) (baseline)
0	-4.684	1.696	8.076	3.808	0.844
1	-2.137	3.221	8.579	3.197	0.691
1.5	-2.928	2.781	8.49	3.407	0.811
2.5	-0.696	4.893	10.482	3.336	0.861
3.5	-3.188	2.273	7.734	3.260	0.780
4.5	0.532	5.707	9.809	2.769	0.779
6	1.232	5.937	10.637	2.807	0.809
12	-3.931	1.532	6.996	3.261	0.691
23.5	1.296	6.570	11.758	3.091	0.717

Auteursrechterlijke overeenkomst

Opdat de Universiteit Hasselt uw eindverhandeling wereldwijd kan reproduceren, vertalen en distribueren is uw akkoord voor deze overeenkomst noodzakelijk. Gelieve de tijd te nemen om deze overeenkomst door te nemen, de gevraagde informatie in te vullen (en de overeenkomst te ondertekenen en af te geven).

Ik/wij verlenen het wereldwijde auteursrecht voor de ingediende eindverhandeling:
Analyzing treatment effects on cardiovascular safety (QT)
Richting: **Master of Science in Biostatistics** Jaar: **2007**
in alle mogelijke mediaformaten, - bestaande en in de toekomst te ontwikkelen - , aan de Universiteit Hasselt.

Niet tegenstaand deze toekenning van het auteursrecht aan de Universiteit Hasselt behoud ik als auteur het recht om de eindverhandeling, - in zijn geheel of gedeeltelijk -, vrij te reproduceren, (her)publiceren of distribueren zonder de toelating te moeten verkrijgen van de Universiteit Hasselt.

Ik bevestig dat de eindverhandeling mijn origineel werk is, en dat ik het recht heb om de rechten te verlenen die in deze overeenkomst worden beschreven. Ik verklaar tevens dat de eindverhandeling, naar mijn weten, het auteursrecht van anderen niet overtreedt.

Ik verklaar tevens dat ik voor het materiaal in de eindverhandeling dat beschermd wordt door het auteursrecht, de nodige toelatingen heb verkregen zodat ik deze ook aan de Universiteit Hasselt kan overdragen en dat dit duidelijk in de tekst en inhoud van de eindverhandeling werd genotificeerd.

Universiteit Hasselt zal mij als auteur(s) van de eindverhandeling identificeren en zal geen wijzigingen aanbrengen aan de eindverhandeling, uitgezonderd deze toegelaten door deze overeenkomst.

Ik ga akkoord,

Setia Pramana

Datum: **27.08.2007**