Design and Analysis of QT/QTc Studies
Conceptional and Methodical Considerations
Based on Experience

Dr. Manfred Wargenau, M.A.R.C.O. Institute, Düsseldorf

OVERVIEW

• Clinical background
• The ICH E14 guideline / review of “deltas”
• Statistical issues concerning the ICH E14 decision criterion for concluding a negative trial
• Study objective and appropriate endpoints with examples
• Power and sample size considerations
• Logistic model for prediction of pathological QTc prolongations
• Summary
The normal ECG

- P wave = electrical signature of the current that causes atrial contraction
- QRS complex = current that causes contraction of the ventricles
  - Q wave = small current as the action potential travels through the inter-ventricular septum
  - R and S waves = contraction of the myocardium itself
- T wave = repolarization of the ventricles
QT and QTc

• The QT interval is the time between the start of the Q wave and the end of the T wave.

• It is dependent on the heart rate and has to be corrected.

• The standard clinical correction is Bazett's formula:
  \[ QTc = \frac{QT}{RR^{1/2}} \]
  where QTc is the QT interval corrected for rate, and RR is the interval between the R peaks (the length of one heart beat).

• However, this formula tends to over-estimate QTc at high heart rates and to under-estimate at low heart rates.
QTc prolongation and ventricular arrhythmia

- Some drugs have the potential to delay cardiac repolarization.
- The delay in repolarization favors the development of cardiac arrhythmias (e.g. torsade de points)
- Torsade de points can lead to ventricular fibrillation and sudden death.

QTc prolongation is therefore considered as a surrogate marker for cardiac arrhythmias.
Clinical relevance

- Current use of non-cardiac QTc prolonging drugs was associated with an almost 3-fold increased risk of sudden cardiac death
  Straus SM et al (2005)

- Risk factors with an impact on cardiac repolarization have to taken into account. These include diabetes, cardiac morbidity, gastro-intestinal medication and anti-psychotics, and female sex.

- TdP predominantly occurred in individuals with QTc values exceeding 500 ms
  Bednar MM et al (2001)
The clinical evaluation of QT/QTc interval prolongation and proarrhythmic potential for non-antiarrhythmic drugs (ICH E14)

“Drugs are expected to receive a clinical electrocardiographic evaluation, beginning early in clinical development, typically including a single trial dedicated to evaluating their effect on cardiac repolarization (“thorough QT/QTc study”).”

**Assay sensitivity:**

“The positive control should have an effect on the mean QT/QTc interval of about 5 ms. Detecting the positive control’s effect will establish the ability of the study to detect such an effect of the study drug.”
“Relevant delta(s)”

• (High) clinical relevance:  QTc > 500 ms
  (or QTc prolongation > 50-60 ms)

• “Clinical threshold”:  QTc prolongation of 30 ms

• “Drugs that prolong the mean QTc interval by more than 20 ms have a substantially increased likelihood of being arrhythmic.” (ICH guideline E14)

• “The regulatory threshold pharmacological effect on cardiac repolarization is a drug effect on QTc of 5 ms in healthy volunteers.” (ICH guideline E14)
The ICH E14 decision criterion

”A negative ‘thorough QT/QTc study’ is 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 provides reasonably assurance that the mean effect is not
greater than around 5 ms”.
The ICH E14 decision criterion

- Does it allow for one-sided testing to detect the threshold of 5 ms?
- Doesn’t it mean that the ”equivalence delta” is in effect 10 ms?
- Why is the study requested to be powered to detect 5 ms when 10 ms are acceptable?

- For example:
  Parallel group study, SD=15, power=90%, two-sided $\alpha=5\%$
  $\Rightarrow$ N=190 subjects per group to detect a delta of 5 ms

A mean difference of up to 7.5 ms may actually be observed to conclude for a negative study!!!
Statistical analysis issues

Parallel group design

• Subjects are randomized to receive one of K treatments.
• For each subject, a QTc time-profile is determined, e.g.: Predose, + 0.5, + 1, + 2, +4, +8, +12 hours post dosing.

Repeated measurement model:

\[ Y_{ijt} = \mu_i + \text{subj}_{j(i)} + \tau_t + (\mu \tau)_{it} + e_{ijt} \]

- \( \mu_i \) denotes the expected overall mean for treatment group \( i \)
- \( \tau_t \) the expected mean for time-point \( t \)
- \( (\mu \tau)_{it} \) denotes the interaction term between treatment group and time
- \( Y_{ijt} \) the QTc interval at time-point \( t \) for subject \( j \) in group \( i \)
- \( e_{ijt} \) the random error at time-point \( t \) for subject \( j \) in group \( i \)

\[ i = 1, ..., K; \ t=1,...,T; \ j = 1, ..., n_i \]
Analysis according to ICH E14 (presumably)

• Primary endpoint: QTc measurement at time-point \( t (= 1, \ldots, T) \)
• Effect: Difference at time-point \( t \) between the mean QTc after active drug and the mean QTc after placebo

SAS application of the repeated measurement model:

– PROC GLM;
– CLASS GROUP SUBJ TIME; /* T=2 for simplicity */;
– MODEL QTC = GROUP SUBJ(GROUP) TIME GROUP*TIME ;
– ESTIMATE ‘tp1: VERUM – PLAC’ GROUP 1 –1 GROUP*TIME 1 0 –1 0;
– ESTIMATE ‘tp2: VERUM – PLAC’ GROUP 1 –1 GROUP*TIME 0 1 0 –1;

Time-matched analysis always assumes and includes “interaction”.
Analysis according to ICH E14 (II)

“Largest time-matched mean effects” considers:

→ The maximum of T test statistics: \( V_{\text{Max}} = \max (V_1, \ldots, V_T) \)

→ Construction of the confidence interval for \( V_{\text{Max}} \) cannot be based on the “usual” quantile of the t-distribution (e.g. 2.00 for \( N=60 \))
Application of the studentized maximum modulus approach would yield: 2.90 for \( T=10 \) and \( N=60 \)

For example: Instead of (0.4, 9.6) one would obtain: (-1.7, 11.7)
Example of mean profiles for a „no effect“ situation

No “treatment*time” interaction Analysis of overall mean

Thus, there is no reason to regard T different test statistics at each time-point, and to select the largest one.
Example of mean profiles for an „effect“ situation

significant “treatment*time” interaction  \( \rightarrow \)  proof of drug effect
quantification of drug effect difficult

Actually: \( E_{\text{max}} \), i.e. the average of individual maxima of QTc = 423 ms
The use of the ICH criterion is therefore also questionable in this situation.
Other design issues concerning “time-matched analysis“

(1) Complete baseline QTc profiles

(2) Cross-over study with placebo period:
    Period effect has then to be ruled out.

In both cases:
    The ICH criterion is still based on the comparison of mean values
    at each time-point, and requires consideration of the distribution
    of the maximum of T test statistics.
Study objectives / Endpoints

(1) "No effect"
   QTc mean over post-dose values or
   QTc mean prolongation (change from predose / baseline) or
   AUC(T¹-T*)

(2) "Effect"
   Maximum QTc over post-dose values (or maximum QTc prolongation) or
   AUC(T¹-T*)

An appropriate time window (T¹, T*) with 0 ≤ T¹ < T* ≤ T needs to be defined. Primary endpoint most recently accepted by the FDA for a thorough QTc study:

   AUC(0.5-3 h)
Examples (I)

"No effect" Cross-over Study, N=24, healthy volunteers

Repeated dosing

<table>
<thead>
<tr>
<th></th>
<th>Maximum prolongation</th>
<th>Overall mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Verum</td>
<td>10.6</td>
<td>441</td>
</tr>
<tr>
<td>Placebo</td>
<td>12.2</td>
<td>443</td>
</tr>
<tr>
<td>Difference (97.5% CI)</td>
<td>-1.9 (-∞, 4.43)</td>
<td>-1.5 (-∞, 2.44)</td>
</tr>
</tbody>
</table>

Very similar result was obtained for single dosing.
Examples (II)

”No effect” Cross-over Study, N=24, CHD patients

Repeated dosing

<table>
<thead>
<tr>
<th></th>
<th>Maximum prolongation</th>
<th>Overall mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Verum</td>
<td>9.0</td>
<td>446</td>
</tr>
<tr>
<td>Placebo</td>
<td>4.9</td>
<td>443</td>
</tr>
</tbody>
</table>

Difference (97.5% CI) 4.2 ($\infty$, 9.14) 3.3 ($\infty$, 5.67)

Upper limit of the confidence interval < 10 ms
Power and sample size considerations

**Cross-over study** adequately powered (90%) with
- \( N=24 \) using QTcF or QTcI.

**Parallel group study:**
- \( N>130 \) per group, power 80%
- \( N>190 \) per group, power 90%.

**Most recently accepted by the FDA for thorough QT/QTc studies:**
- \( N=40 \) per group

assuming an effect of the positive control of 8-10 ms (”sensitivity”).
The logistic regression model

\[ \text{logit} (p) = \log \left( \frac{p}{1 - p} \right) = \alpha + \mathbf{x} \beta = \alpha + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k \]

\[ p = \frac{e^{(\alpha + x \beta)}}{1 + e^{(\alpha + x \beta)}} \]

- ordinal response variable \( Y \) \((< 480, 480-<500, \geq500 \text{ ms})\)
- explanatory variables \( x \):
  - continuous (dose) and categorical (gender)
- \( p \): probability for binary or ordinal outcome of variable \( Y \)
- \( \alpha \): the intercept parameter, \( \beta \): vector of slope parameters
The logistic regression model (example)

Estimated risks of experiencing QTc ≥ 500 ms depending on gender and dose

for males and females
Summary (I)

• Clinically relevant effects (TdP highly associated with QTc > 500 ms) and the pharmacological threshold value (5 ms) relate to completely different concepts

• The likelihood of clinical effects increases with QTc prolongation > 20 ms / 30 ms “clinical threshold”

• The guideline implicitly refers to an equivalence delta of 10 ms
Summary (II)

- The approach of “time-matched analysis” is questionable from the statistical point of view:
  - it complicates analysis for “no effect” objective
  - is likely to cause bias for the “effect” situation due to inter-subject variability of $E_{\text{max}}$
  - requires consideration of the distribution of the “maximum test statistic”
  - calculation of “standard” 95% confidence intervals inflates the type I error probability for (1) concluding an effect and (2) concluding a negative trial
Summary (III)

- Primary endpoints such as
  - maximum QTc prolongation or
  - AUC(T1-T*)

  are useful for addressing both study objectives of “effect” and “no effect”, and for quantifying QTc drug effects.

- Categorical analyses including application of the logistic model provides information on the risk of experiencing relevant QTc prolongations.
END

THANK YOU
References


Design a study with a drug affecting QTc

4-way cross-over (preferably)

Treatments

- Placebo
- Therapeutical dose x
- Dose x/2
- Dose 2x

Complete baseline ECG profile prior to each period