

# The Maximum Mean Difference

---

Statistical Problems in Assessing Cardiac Safety

Georg Ferber, Novartis Pharma AG, Basel

ROeS Seminar Bern, 10-13 Sep 2007

# Overview

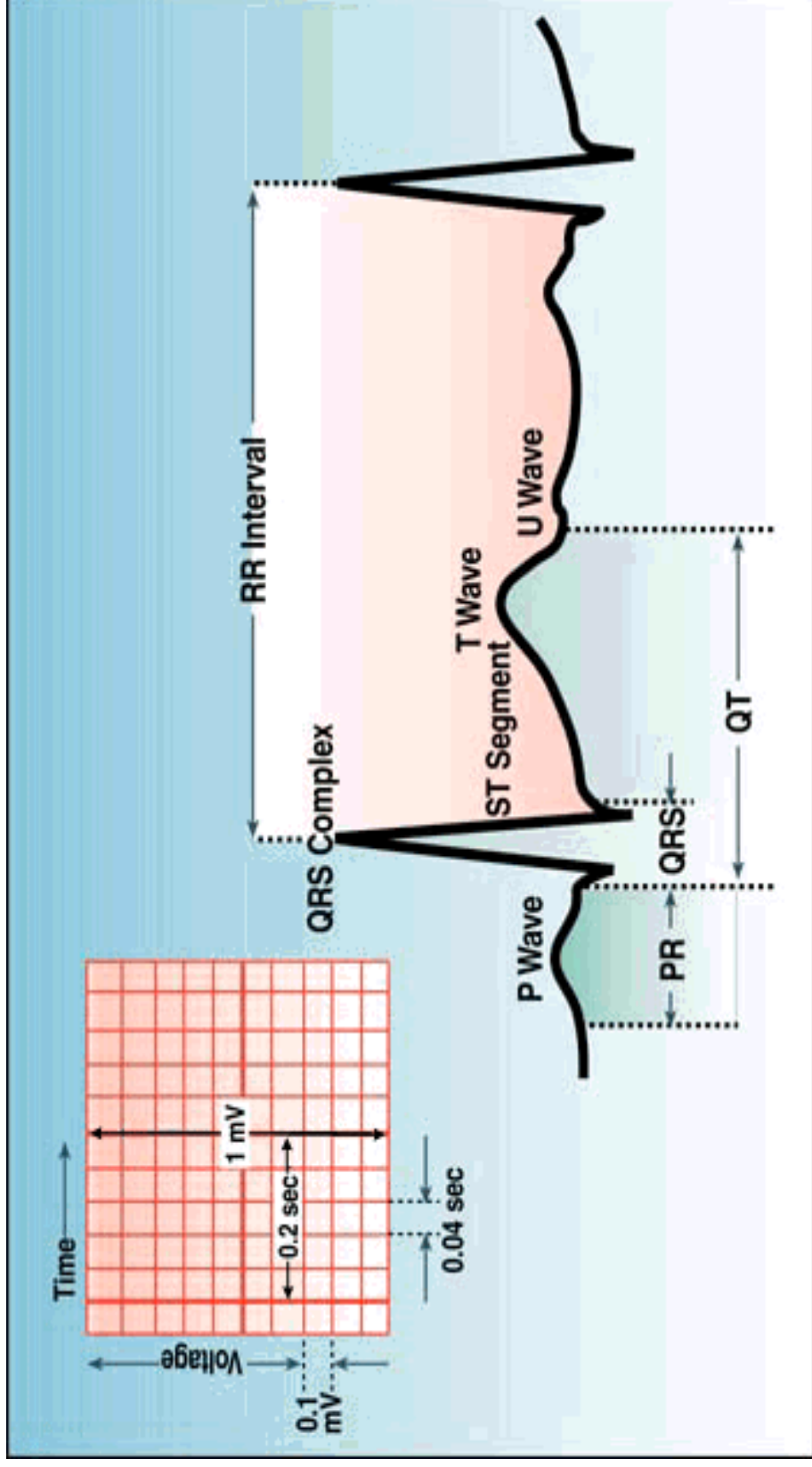
---

- QT prolongation – some background
- The Thorough QT/QTc Study and its endpoint
- Some solutions to the Max-of-Mean-Difference problem
- The "Bias" of the Max-of-Mean-Difference
- Other aspects of the Thorough QT/QTc study

# Background

- *Torsade de Pointes is a rare disturbance of the heart rhythm. It can lead to sudden death.*
- *TdP has been associated with a number of drugs, and some of them have been withdrawn from the market or restricted in their use.*
- *Since TdP is rare and – if not fatal – transient, it is difficult to observe. Prolongation of the QT interval in the ECG is considered a biomarker for TdP.*
- *However, the relationship between QT prolongation and TdP is rather loose – QT prolongation can be considered a good example for a poor biomarker.*
- *Despite of intense research, QT prolongation still is considered the best clinical predictor for the pro-arrhythmic potential of non-antiarrhythmic drugs.*

# What is the QT interval?



## ... and what is QTc?

- There is a relationship between the duration of the QT interval and that of the duration of one beat, i. e. the RR interval, or, inversely, the heart rate.
- Many attempts have been made to correct the QT interval for heart rate. QTc is the "corrected" QT interval, i. e. the interval one would expect at an RR-interval of 1 sec.
- Fridericia's correction

$$QTc = QT/RR^{1/3}$$

is usually a good one.

# History of ICH E14

---

- The relationship between drugs and TdP became evident in the late 80es
- CPMP has issued a guidance on QT assessment in 1997
- In 2002, the FDA joined an initiative of Health Canada (TPD) for a guidance.
- In early 2003, this draft guidance became ICH E14.
- After a few meetings, this guidance was promoted to Step 4 in May 2005 and became effective in the EU in Nov 2005. To my knowledge it is still not effective in Japan.

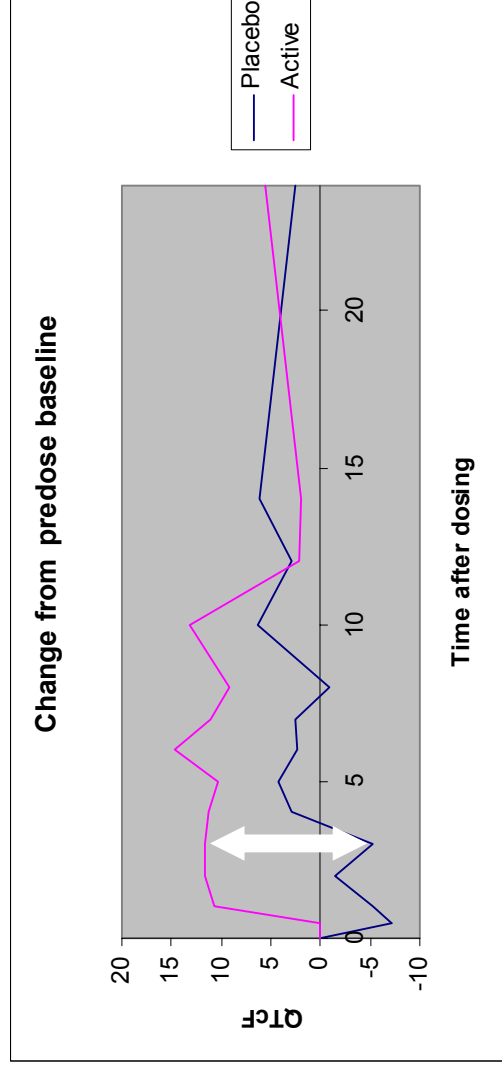
## The Thorough QT/QTc study (TQT study)

---

- This is a placebo and active controlled study
- usually performed in healthy volunteers
- measuring the ECG during baseline
- and at a number of time points during a day under study medication.
- Of key interest is the comparison of the QT/QTc interval – after adjustment for baseline – between test and placebo at the same time points.
- Active control is to ensure assay sensitivity.

## Definition of primary endpoint

... 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 definition is chosen to provide reasonable assurance that the mean effect of the study drug on the QT/QTc interval is not greater than around 5 ms.



# Is this a statistically reasonable criterion?

Naive statistical interpretation:

$X = (X_1, \dots, X_T)^T$  – vector of time matched mean differences

$Z := \text{Max}\{X_i\}$  is our random variable of interest

Determine a (one sided 95 %) CI for  $Z$ .

- Problem 1: What do we know about the distribution of  $X$ ?
- Problem 2: Even if we assume a benign distribution of  $X$ , what do we know about distribution of  $Z$ ?
- Problem 3: By definition,  $Z - X_i \geq 0$  for  $i = 1, \dots, T$  ("Bias")

# How can we solve the Max-of-mean problem?

---

- Literal interpretation
  - Approximation
  - Bootstrapping
- Simplification
  - Replacement by a intersection-union-test that is at least as conservative as the original condition

## Approximation (Eaton et al. 2006)

Start with a smooth  $r$ -norm like function for the vector of differences  $\mathbf{x}$ :

$$f_r(\mathbf{x}) := \left( \frac{1}{T} \sum_{x_i \geq 0} x_i^r \right)^{1/r}$$

and construct a confidence interval for  $f_r$  based on the  $\delta$ -method.

Approximate the maximum by letting  $r \rightarrow \infty$ .

# Bootstrapping

---

Boos et al. (2007) use bootstrapping to obtain confidence intervals for the maximum mean difference:

- parametric ones based on a multivariate normal model
- nonparametric ones

## Simplification

---

The Max-of-Mean-Difference condition will be met if  
for each time point,  
the confidence interval for the mean difference  
between the drug and placebo (baseline-adjusted)  
excludes an effect  $>10$  ms.

In this (somewhat more stringent) reformulation as an Intersection-Union Test the problem is amenable to conventional multiple noninferiority hypothesis testing.

# What does the IUT do?

- It opens a conventional way to analyse the TQT study according to the guidance
- It is conservative in the sense that it protects public health
- It does not remove the "bias" inherent in the endpoint definition
- In particular, the sponsor is still punished for adding a time point
- Simple sample size calculations can lead to unrealistically large trials if assumptions are simplified too much.
- It is with this option in mind that the guideline was promoted to step 4.

## How are TQT studies analysed in practice?

---

- It seems that virtually all sponsors use the intersection-union approach.
- Boos et al. recommend their nonparametric bootstrapping as more powerful.
- The Eaton et al. approach is rather of theoretical interest.
- Within the intersection-union approach, there are many variants that are being used.

# Bias

## Problem:

- Conceptually, the guideline looks for inference on the maximum of the "true" difference between the test drug and placebo.
- With this in mind, we would expect that
  - if the drug does not have any effect on QT,
    - our estimator should have a mean of 0
    - adding a time point should not change the location of our estimator systematically.
- In other words, conceptually we are looking for the expected value of the maximum difference

## Bias continued

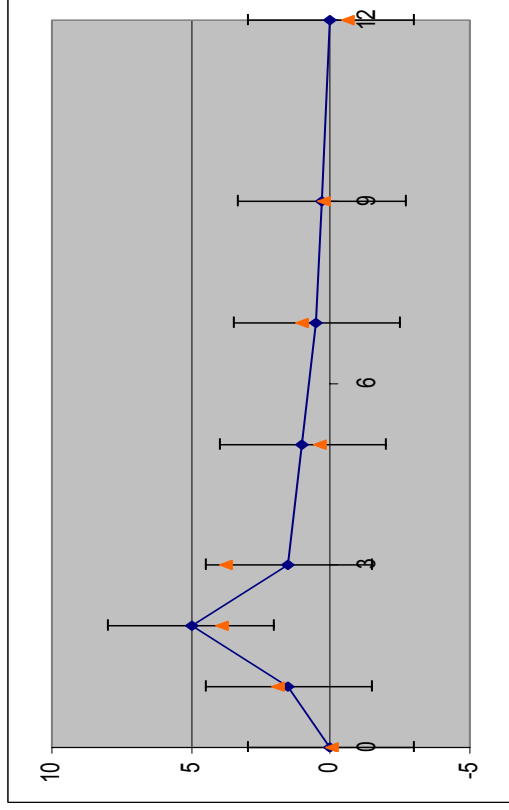
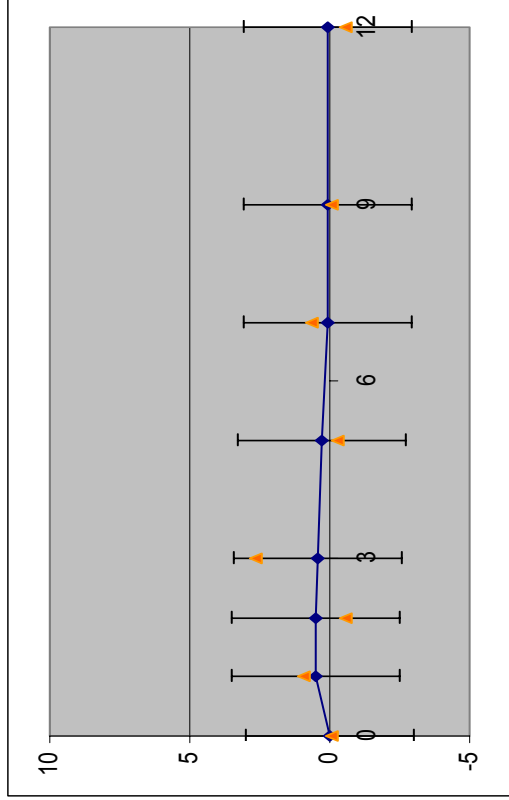
---

However,

- in practice, we can only look at estimators, e. g. the differences of mean values.
- Even if the drug has no effect on QT,
  - the maximum mean difference will be  $> 0$  with high probability
  - adding a time point will shift its distribution further upwards
- This is because we are looking at the maximum of estimates rather than an estimator of the maximum.

# What does the difference (Bias) look like?

Contribution of random variability of mean difference



Obviously, the degree of correlation between time points also plays a role.

## Bias - consequences

---

- Simulations by several authors suggest a bias of  $\approx 3$  ms if there is no true difference.
- It was with this in mind that the ICH E14 working group moved the noninferiority margin from 8 to 10 ms.

# Other statistical challenges

---

- **Correction method:**
  - In general, Fridericia's method for correction for heart rate works well, but it is known that there is between subject variability in the relationship. This can be estimated.
- **Baseline:**
  - Should we use a time-matched baseline or rather use an average over all pre-dose values we measure?
- **Repeats:**
  - The variability of our measurement can be reduced by doing more than one measurement per time point. This can be optimised with respect to costs.

# Conclusion

---

- The use of inferential statistics to show safety is new.
- Traditionally, we strive for robustness
- If we look for safety, robustness must not translate to missing a signal.
- This is the new statistical challenge behind the ICH E14 guideline.

Thank you for your attention!

# References

Boos DD, Hoffman D, Kringle R, Zhang J: New confidence bounds for QT studies. Statistics in medicine, electronically published Feb 2007.

Eaton ML, Muirhead RJ, Manusco JY, Kolluri S: A confidence interval for the maximal mean QT interval change caused by drug effect. Drug information journal **40** (2006), pp 267 - 271

International conference on harmonisation: The clinical evaluation of QT/QTc interval prolongation and proarrhythmic potential for non-antiarrhythmic drugs. Step 4, May 2005  
([http://www.ich.org/UrlGrpServer.jsr?@\\_ID=475&@\\_TEMPLATE=272#E14](http://www.ich.org/UrlGrpServer.jsr?@_ID=475&@_TEMPLATE=272#E14))