Bias Evaluation as a Metric for Protection Against False Negative Results

- A Replacement for the Pharmacological Positive Control?

Borje Darpo MD PhD
Chief Scientific Officer, iCardiac
The positive control in QT studies serves as a protection against false negative results.

- When the study drug does not have a QT effect, it is reassuring to see that the QT effect of moxifloxacin can be detected in the same study.

- When supratherapeutic plasma levels cannot be achieved, the concern for false negative results increases.
Research Objective:
To evaluate whether bias in QTc data severe enough to cause false negative results with exposure response (ER) analysis can be detected by a proposed metric

Research group:
- Georg Ferber (Independent Consultant)
- Meijian Zhou (iCardiac)
- Corina Dota (AZ)
- Christine Garnett (FDA)
- Jim Keirns (Astellas)
- Marek Malik (St Paul’s)
- Norman Stockbridge (FDA)
- Borje Darpo (iCardiac)
**Underlying concept:**

*Compare data generated by the core lab with fully automated, ‘human independent’ background data from the same study*

- Most core labs extract ECGs from 5-minute time windows around each defined timepoint and then use a fully automated, underlying ECG algorithm to produce initial fiducial (measurement) points for the QT interval
  - Bravo (AMPS), Veritas (Mortara), COMPAS (iCardiac), 12SL (GE), Eclysis (AZ), and others
- The reported value for each subject and timepoint is then generated from extracted ECGs using a large variety of techniques with variable level of human intervention
  - Methods vary also within the larger categories, such as semi-automated
Methods

Metric

• Compare for each subject and timepoint:
  – QTc with the core lab’s method, the *reported value*
  with
  – Median QTc from 5-minute time window (all beats) from which ECGs are extracted, the ‘background’

• *Evaluate negative bias using Bland-Altman (BA) plots and evaluate the slope of the linear regression across data points*

Testing the impact of bias

• Artificially introduce negative bias into the core lab reported values
• Evaluate to what extent this resulted in false negative results using data from a small study with 5 QT↑ drugs
• Evaluate the BA-slope vs. rate of false negatives
Comparing Methods Using Bland Altman (BA) Plots

Bland Altman plot displays differences between methods vs. average of both, e.g., HPQT (iCardiac’s method) compared with COMPAS (automated algorithm)

<table>
<thead>
<tr>
<th></th>
<th>Mean difference (ms)</th>
<th>Slope* (Mean; 95% CI, ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original</td>
<td>-0.57</td>
<td>-0.004 (-0.023 to 0.015)</td>
</tr>
<tr>
<td>Biased dataset</td>
<td>-1.8</td>
<td>-0.102 (-0.133 to -0.071)</td>
</tr>
</tbody>
</table>

* Estimated from a linear model by robust regression using M-estimators
Data – IQ-CSRC Study

- 20 healthy subjects randomized to 3 treatment periods
- Each treatment period with two dosing days
  ✓ Dose on Day 1 intended to result in QT effect (\(\Delta\Delta QTcF\)) of ~9 to 12 ms
  ✓ Dose on Day 2 intended to give QT effect of ~ 15 to 20 ms
- Incomplete block design: 9 subjects received each active drug and 6 subjects received placebo

<table>
<thead>
<tr>
<th>Drug</th>
<th>Day 1</th>
<th>Day 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>Observed peak (\Delta\Delta QTcF), ms Mean</td>
</tr>
<tr>
<td>Ondansetron</td>
<td>9</td>
<td>12.2</td>
</tr>
<tr>
<td>Quinine</td>
<td>9</td>
<td>13.3</td>
</tr>
<tr>
<td>Dolasetron</td>
<td>9</td>
<td>6.5</td>
</tr>
<tr>
<td>Moxifloxacin</td>
<td>9</td>
<td>11.9</td>
</tr>
<tr>
<td>Dofetilide</td>
<td>9</td>
<td>14.2</td>
</tr>
<tr>
<td>Levocetirizine</td>
<td>8</td>
<td>1.8</td>
</tr>
<tr>
<td>Placebo</td>
<td>6</td>
<td>NA</td>
</tr>
</tbody>
</table>
**Bias Methods – Deterministic**

**Knee**: Values exceeding 400 ms were reduced by 30%, 50% or 70%

**Linear**: Symmetric reduction/increase around center value (400 ms)

**E_max**: Values above/below center value (400 ms) reduced/increased with E_max function
Compression of range of QTc:
The deviation from the central value of QTc₀ = 400 ms is reduced by a factor α:

\[ \text{QTc}_{\text{bias}} = \text{QTc}_0 + \alpha (\text{QTc} - \text{QTc}_0) \]

- Asymmetric random bias: the formula above is applied only to values > QTc₀
- Symmetric random bias: formula applied to all values.

α is selected from a uniform distribution in [a, 1] with a = 0.7, 0.5, 0.3 and 0

- 300 runs of each severity, i.e. 1200 in total
Slopes of regression* through data points very shallow (within -1.3 to +2.4ms/100ms)
* Estimated from a linear model by robust regression using MM-estimators
Evaluation of False Negatives

• Exposure response analysis was performed on each biased dataset using the same statistical methods as for the original (published) dataset
• The QT effect ($\Delta\Delta QTcF$) was predicted at the observed Cmax on Day 1, i.e. after the lower dose of the drugs
• In case the predicted effect (UB of 90% CI of $\Delta\Delta QTcF$) was $< 10$ ms, this was deemed a false negative results
Slopes of regression on biased data from ondansetron range from -0.03 to -0.67 (= -3 to -67 ms over a range of 100 ms)
Ondansetron: Results – Deterministic bias

Slope of exposure response relation

False negatives vs. bias severity
Dolasetron: Results – Deterministic bias

Slope of exposure response relation

False negatives vs. bias severity
Results: Random bias - BA slopes vs. False negatives

- 2400 runs of random asymmetrical and symmetrical bias
- Except for dolasetron, with only 7.4 ms peak effect on Day 1, the rate of false negatives was < 5% when negative bias did not exceed -0.20 (-20 ms per 100 exceed ms)
Conclusions

- Only severe bias resulted in false negative results when using ER analysis on data from mildly QT prolonging drugs
  ✓ ER analysis is a robust methodology with low likelihood of false negative results when sufficient amount of data used

- The regression slope from Bland Altman plots (BA slope) seems to be a useful metric to detect biased ECG methods
  ✓ Bias severe enough to cause false negatives with exposure response analysis was detectable using BA plots
If a BA-slope exceeding -10 ms per 100 ms (-0.10) is observed between fully automated ECG measurements and the core lab reported values, data should be further explored if used to support a claim of a negative QT assessment using ER analysis without inclusion of a positive control.
Limitations

- Bias was artificially created and severe; it can be questioned how ‘real’
  - The intention was however to create bias severe enough to cause false negatives with ER analysis of data from a small study, which we succeeded with

- Only one ECG method, the High Precision QT technique
  - HPQT uses more data (~100 of ~300 beats) as compared to most semi-automated techniques (9 of ~300 QT/RR)

- Only one data set and only 5 QT+ drugs

- We encourage others to perform the same type of analysis using other data sets and ECG methods
THANK YOU FOR YOUR ATTENTION

QUESTIONS?