Industry View: Integration of non-clinical and clinical data to replace the TQT Study

Charles Benson, M.D., Ph.D.
Eli Lilly & Co
Can data integration improve our approach to QT assessment?

**Issues**

Currently – Only a Thorough QT study or positive early clinical study at therapeutic doses is acceptable as characterization of the risk of QT prolongation.

- Preclinical data does a good job at predicting risk of QT prolongation... but not good enough (false negative =15%?).
- Early clinical data does a good job at predicting risk of QT prolongation... but not good enough (false negative <10%).
- Can a ‘totality of evidence’ approach, integrating preclinical and early clinical data using M&S, appropriately characterize the QT interval?
Bayes' theorem

- Thomas Bayes' probability theorem involving prior knowledge and accumulated experience
- Combining the preclinical QT data with the early clinical data could result precise prediction of clinical risk
- This could obviate the need for a TQT trial
- Updating probabilities based on existing + new data is a good idea.

Thomas Bayes (1702 –1761) was a British mathematician and Presbyterian minister.
Contingency table and Sensitivity, Specificity, Type 1 & 2 Error

<table>
<thead>
<tr>
<th>Actual disease</th>
<th>Present</th>
<th>Absent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test Positive</td>
<td>TP</td>
<td>FP</td>
</tr>
<tr>
<td>Results Negative</td>
<td>FN</td>
<td>TN</td>
</tr>
</tbody>
</table>

TP = True Positive  
FP = False Positive  
FN = False Negative  
TN = True Negative

Sensitivity = $\frac{TP}{(TP + FN)} = 1$ - Type 2 error (beta) =1 - chance of concluding no disease when disease is present

Specificity = $\frac{TN}{(TN + FP)} = 1$ - Type 1 error (alpha) =1 - chance of concluding a disease when none exists
## Contingency table and positive and negative predictive values

<table>
<thead>
<tr>
<th>Actual disease</th>
<th>Present</th>
<th>Absent</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Test</strong></td>
<td><strong>Positive</strong></td>
<td><strong>TP</strong></td>
</tr>
<tr>
<td><strong>Results</strong></td>
<td><strong>Negative</strong></td>
<td><strong>FN</strong></td>
</tr>
</tbody>
</table>

**TP** = True Positive  
**FP** = False Positive  
**FN** = False Negative  
**TN** = True Negative

Positive Predictive Value  
\[
\text{Positive Predictive Value} = \frac{TP}{(TP + FP)}
\]

Negative Predictive Value  
\[
\text{Negative Predictive Value} = \frac{FN}{(FN + TN)}
\]
Bayes' theorem applied to drug development and QT data

What is the positive and negative predictive value of Phase 1 QT data, given that preclinical data does a fairly good (but not 100%) job of predicting risk?

For sake of argument, say there is a good margin to clinical concentrations (10x?):

Sensitivity of preclinical package hERG + dog = .90 (False negative rate = 10%)

Specificity of preclinical package hERG + dog = .60 (False positive rate = 40%) ****These compounds are usually KILLED

**Specificity decreases (FP increase) with increased cutoff for assay (i.e. from 10 to 30 fold margins)
Bayes' theorem applied to drug development and QT data

What is the positive and negative predictive value of Phase 1 QT data, given that preclinical data does a fairly good (but not 100%) job of predicting risk?

For sake of argument, say there is a good margin to clinical concentrations (10x?):

• Prevalence of a true QT prolonger making it to Phase 1 via a False negative set of preclinical studies is 10% (assumed for this exercise)
Bayes' theorem applied to drug development and QT data

What is the positive and negative predictive value of Phase 1 QT data, given that preclinical data does a fairly good (but not 100%) job of predicting risk?

For sake of argument, say there is a margin to clinical concentrations (2x):

- Sensitivity of Phase 1 PK/PD = .95 (False negative rate = 5%)
- Specificity of Phase 1 PK/PD = .90 (False positive rate = 10%)

We test a Preclinical NEGATIVE compound in Phase 1...
## Pretest and Posttest Probability

<table>
<thead>
<tr>
<th>Actual Disease</th>
<th>Present</th>
<th>Absent</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test Results</td>
<td>Positive</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>TP</td>
<td>FP</td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>FN</td>
<td>TN</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>10</td>
<td>90</td>
<td>100</td>
</tr>
</tbody>
</table>

Given actual prevalence (prior probability) of 10%

- Sensitivity = \( \frac{TP}{TP + FN} = 0.95 \)
- Specificity = \( \frac{TN}{TN + FP} = 0.90 \)
### Pretest and Posttest Probability

Given actual prevalence (prior probability) of 10%

<table>
<thead>
<tr>
<th>Actual disease</th>
<th>Present</th>
<th>Absent</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test Results</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>TP</td>
<td>FP</td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>FN</td>
<td>TN</td>
<td></td>
</tr>
<tr>
<td>total</td>
<td>10</td>
<td>90</td>
<td>100</td>
</tr>
</tbody>
</table>

**Sensitivity** = \( \frac{TP}{TP + FN} = .95 \)  
\( TP = 9.5 \)

**Specificity** = \( \frac{TN}{TN + FP} = .90 \)  
\( TN = 81.5 \)
### Pretest and Posttest Probability

<table>
<thead>
<tr>
<th>Actual disease</th>
<th>Present</th>
<th>Absent</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test Results</td>
<td>Positive</td>
<td>9.5</td>
<td>FP</td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td>FN</td>
<td>81.5</td>
</tr>
<tr>
<td>total</td>
<td>10</td>
<td>90</td>
<td>100</td>
</tr>
</tbody>
</table>

Given actual prevalence (prior probability) of 10%

- **Sensitivity** = \( \frac{TP}{TP + FN} = 0.95 \)  
  \( TP = 9.5 \)

- **Specificity** = \( \frac{TN}{TN + FP} = 0.90 \)  
  \( TN = 81.5 \)
## Pretest and Posttest Probability

<table>
<thead>
<tr>
<th>Actual disease</th>
<th>Present</th>
<th>Absent</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test Positives</td>
<td>9.5</td>
<td>18.5</td>
<td>10</td>
</tr>
<tr>
<td>Results Negative</td>
<td>0.5</td>
<td>81.5</td>
<td>90</td>
</tr>
<tr>
<td>Total</td>
<td>10</td>
<td>90</td>
<td>100</td>
</tr>
</tbody>
</table>

**Sensitivity** = \( \frac{TP}{(TP + FN)} = 0.95 \)  
TP = 9.5

**Specificity** = \( \frac{TN}{(TN + FP)} = 0.90 \)  
TN = 81.5

Given actual prevalence (prior probability) of 10%
Bayes' theorem applied to drug development and QT data

<table>
<thead>
<tr>
<th>Actual QT Effect present</th>
<th>Present</th>
<th>Absent</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase 1 Positive</td>
<td>9.5</td>
<td>9</td>
<td>18.5</td>
</tr>
<tr>
<td>Results Negative</td>
<td>.5</td>
<td>81</td>
<td>81.5</td>
</tr>
<tr>
<td>total</td>
<td>10</td>
<td>90</td>
<td>100</td>
</tr>
</tbody>
</table>

Positive Predictive Value = $\frac{TP}{TP + FP}$
$\frac{9.5}{9 + 9.5} = .51$ 51%

Negative Predictive Value = $\frac{TN}{FN + TN}$
$\frac{81}{81.5} = .994$ 99.4%

Given negative preclinical, prevalence (prior probability) is 10%
Bayes' theorem applied to drug development and QT data

What is the positive and negative predictive value of TQT data, given that preclinical and Phase 1 data does a fairly good (but not 100%) job of predicting risk?

For sake of argument, say it is a well powered study:

• Sensitivity of TQT study = .95 (false negative rate = .05)
• Specificity of TQT study = .90 (false positive = 10%)

We test a Preclinical and Phase 1 NEGATIVE compound in a TQT study...
Bayes' theorem applied to drug development and QT data

Given negative preclinical and phase 1 clinical, prevalence (prior probability) is .6%

<table>
<thead>
<tr>
<th>Actual QT Effect present</th>
<th>Present</th>
<th>Absent</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>TQT Positive</td>
<td>.57</td>
<td>9.94</td>
<td>10.51</td>
</tr>
<tr>
<td>Results Negative</td>
<td>.03</td>
<td>89.46</td>
<td>89.49</td>
</tr>
<tr>
<td>total</td>
<td>.6</td>
<td>99.4</td>
<td>100</td>
</tr>
</tbody>
</table>

Positive Predictive Value = \( \frac{TP}{TP + FP} \)
\[
\frac{.57}{9.94 + .57} = .054 \quad 5.4\%
\]

Negative Predictive Value = \( \frac{TN}{FN + TN} \)
\[
\frac{89.46}{89.49} = .999 \quad 99.9\%
\]

The addition of the TQT study increased the negative predictive value from 99.4% to 99.9%.

Charles Benson, M.D.Ph.D.  
Copyright © 2009 Eli Lilly and Company
Resolving Uncertainty in Predicting QTc

- 80% of the uncertainty about QTc outcome resolved by hERG assay
- 95% following the in vivo study
- >100% following FHD

Risk resolved for QT prolongation

- hERG
- In vivo QT
- Human Ph I QT data
- Human TQT

0% 50% 100%

CAN Phase I Phase II Registration

Slide courtesy of Derek Leishman, Ph.D.
Eli Lilly and Co
A dynamic population model to calculate the cost-effectiveness of ICH E14 for a prototype QT-prolonging antipsychotic drug entering the US and European markets.
Implementing ICH E14 vs. not implementing ICH E14 cost approximately $246,000 per life year gained, $4.4 million per sudden cardiac death prevented, and $343,000 per quality-adjusted life years (QALYs) gained.

This is a substantially higher sum than society is at present willing to pay per QALY gained for health-care programs in Europe (€20,000-80,000 per QALY).

In the United States, a threshold of $50,000–$100,000 per QALY gained is commonly cited.

Even when several of the assumptions in the model were varied, there were no results in favor of regulation.

Our study shows that cost-effectiveness analysis of drug regulatory measures is feasible and should be considered before developing such measures.
The model is conservative:

- It uses a very high estimate of QT liability (antipsychotics), if it had used a lower incidence the QALY would be higher.
- It doesn’t include the cost of wasted QT measurement on drugs that don’t survive to the market for reasons other than QT.
- It doesn’t include the cost of QT measurement other than TQT (upto 22% of Phase 1 costs)
- It doesn’t include cost of killing good drugs due to QT prolongation
- It doesn’t include preclinical costs
  - GLP hERG study is around $50K
  - dofetilide-binding assay or or astemizole-binding + analytical work
  - in vivo CV study (clearly we get additional CV information) cost around $200K
Conclusions

- No test is perfect (100% sensitive and specific), however, Bayesian combination of the preclinical QT data and the early clinical data results in negative predictive values which are very high, and likely exceed that of a TQT trial.

- Integration of preclinical QT data with the early clinical data results in precise prediction of clinical risk which could obviate the need for a TQT trial.

- Cost-effectiveness analysis of drug regulatory measures is feasible and should be considered before developing such measures Bouvy et al., Clinical pharmacology & Therapeutics, Vol 91, No 2. February 2012
Questions:

- What can M&S do to improve our approach to QT prolongation risk?
  - Are we ready to trust a combined, ‘totality of evidence,’ approach to QT assessment?
  - Are we ready to trust concentration response models to predict negative and positive early phase data?

- “…In a world of rising healthcare expenditures and increasing drug development costs, regulatory agencies and society at large should think carefully about what they are willing to pay for reassurance with respect to drug safety to the extent of determining the magnitude of very small risks.” Bouvy et.al, Clinical pharmacology & Therapeutics, Vol 91,No 2. February 2012
Back up Slides

- Acknowledgement: Thanks to Derek Leishman, Ph.D. Eli Lilly and Company for many of the slides in the back up deck.
Predicting TQT – HERG Margin

Optimal at 45-50 fold margin between hERG IC$_{50}$ and unbound plasma concentration

Predicting TQT – Using IC$_{10}$ Value

Pre-clinical data translation to man

- **hERG**
  - Wallis (2010)
  - Analysis of 19 compounds (11 with +ve TQTS; 8 with –ve TQTS)
  - Did hERG IC$_{10}$ predict TQTS outcome within 2-fold C$_{max}$ free?

<table>
<thead>
<tr>
<th>Pre-clin</th>
<th>Human</th>
<th>Neg</th>
<th>Pos</th>
<th>Predictive capacity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neg</td>
<td>6</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pos</td>
<td>2</td>
<td>9</td>
<td></td>
<td><strong>0.79</strong></td>
</tr>
</tbody>
</table>

If a compound is ‘negative’ in human, what is the probability that the animal model will correctly identify it?

- **Specificity**
  - 0.75

If a compound is ‘positive’ in human, what is the probability that the animal model will correctly identify it?

- **Sensitivity**
  - 0.82

The IC$_{10}$ value is approximately 1/9$^{th}$ of the hERG IC$_{50}$

• Based on a cross pharma initiative through ILSI/HESI
• In Vivo assay appears very predictive of which compounds might be associated with TdP
  • What about concentration where effects occur in animals and man?

Pre-clinical data translation to man

**In vivo** models – conscious telemetered dog

- Hanson et al. (2006) - ILSI-HESI
- Analysis of 12 compounds (6 with TdP risk; 6 without)
- Did effect on QTcF predict TdP risk?
- Significant prolongation = TdP risk; NS prolongation = no TdP risk

<table>
<thead>
<tr>
<th>Pre-clin Human</th>
<th>Neg</th>
<th>Pos</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neg</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Pos</td>
<td>0</td>
<td>6</td>
</tr>
</tbody>
</table>

Predictive capacity = 1.00

If a compound is 'negative' in human, what is the probability that the animal model will correctly identify it.

Specificity: 1.00  Sensitivity: 1.00
Pre-clinical data translation to man

- **In vivo** models – conscious telemetered dog
  - Toyoshima et al. (2005) - PRODACT
  - Analysis of 21 compounds (11 with TdP risk; 10 without)
  - Did effect on QTc predict TdP risk?
  - Significant prolongation = TdP risk; NS prolongation = no TdP risk

<table>
<thead>
<tr>
<th>Pre-clin</th>
<th>Human</th>
<th>Predictive capacity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pos</td>
<td>Neg</td>
<td>Pos</td>
</tr>
<tr>
<td>1</td>
<td>11</td>
<td>0.95</td>
</tr>
<tr>
<td>Neg</td>
<td>9</td>
<td>0</td>
</tr>
</tbody>
</table>

- Based on a cross pharma and CRO initiative in Japan called PRODACT
- Again trying to predict compounds associated with TdP
  - Likely requires a large QTc change in man? See Bednar et al 2002. These authors suggested most TdP associated with QTc of 500ms or greater
  - If one can only detect changes of reasonable magnitude in vivo the predictive value is still very high if attempting to predict dramatic effects in man
Statistical Power ILSI/HESI (c.f. Hanson et al 2006)

- n=8, 4 x 4 Latin Square, doubled
- 15 ECG complexes averaged at 7 time points
- 80% Power to detect 7% increment of QT, a 5% increment of QTcF, and a 4% increment of QTcQ
- Based on 8 animals and 4 treatments the data density is around 420 ECGs from around 400K possible complexes (0.1%)

### Pre-clinical data translation to man

**In vivo models – conscious telemetered dog**

- Wallis (2010)
- Analysis of 19 compounds (11 with +ve TQTS; 8 with –ve TQTS)
- Did effect on QTc predict TQTS outcome within 2-fold $C_{\text{max}}$ free
- $> 10$ ms prolongation = +ve TQTS; $< 10$ ms prolongation = -ve TQTS

<table>
<thead>
<tr>
<th></th>
<th>Human</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Neg</td>
<td>Pos</td>
</tr>
<tr>
<td>Pre-clin Neg</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>Pre-clin Pos</td>
<td>2</td>
<td>9</td>
</tr>
</tbody>
</table>

**Predictive capacity:** 0.79

**Specificity:** 0.88

**Sensitivity:** 0.82

- Recent publication based on 19 TQT (Thorough QT) type studies
- Attempting to predict a QTc change of between 5 and 10ms in man
  - Concentration and margin based since uses $C_{\text{max}}$ value from TQT study
  - Still good predictive capacity
  - Improved by using additional data in an integrated assessment as suggested in ICH S7B
Statistical Power Pfizer (c.f. Wallis, 2010)

- n=4, 4 x 4 Latin Square
- All available ECG complexes, superinterval average
- 80% power to detect the following changes: HR (10 bpm), LV+dP/dtmax (375 mmHg/s), MBP (5 mmHg) and QTc (4 ms; approx 2%)
- Based on 4 animals and 4 treatments the data density is likely around 300K ECGs from around 400K possible complexes (75%)

A HESI Consortium Approach to Assess the Human Predictive Value of Non-Clinical Repolarization Assays

Safety Pharmacology Society Annual Meeting Innsbruck, Austria September 21, 2011

- Presentation from John Koerner (FDA)
- Co-chair of this ILSI/HESI Group
- Strong partnership with FDA
Objectives

Project Objectives

1. To assess the concordance between non-clinical repolarization assays and clinical measures of QT interval prolongation

2. To investigate the mechanisms for any discrepancy identified between non-clinical and clinical results and to determine viable and successful alternative approaches to identify these compounds

3. To assess the proarrhythmic potential of such compounds

- Ultimate goal to eliminate TQT for genuinely negative compounds
- Potentially limit attention to compounds thought to be proarrhythmic
Concordance Phase

Stage 1 Strategy

- Generate datasets containing non-clinical and clinical QT information
- Public-private collaborative data collection
- 4 approaches
  1. Data submitted to FDA (76 drugs)
  2. Additional data input from Pharma companies
  3. Use of data from previous initiative (Hanson et al) (12 drugs)
  4. Data available from the literature (17 drugs)

- Focus of the work to date
- All 3 & 4 approaches have largely completed
- Approach 1 in progress
- Approach 2 on hold – desire was for Phase 1 data
- FDA have supported this well but competing demands are an issue

### FDA Database Contents

<table>
<thead>
<tr>
<th></th>
<th>TQT</th>
<th>hERG</th>
<th>APD</th>
<th>Non-Clinical ECG</th>
</tr>
</thead>
<tbody>
<tr>
<td># of Studies</td>
<td>76</td>
<td>73</td>
<td>41</td>
<td>73</td>
</tr>
<tr>
<td>Quality Assurance</td>
<td>22</td>
<td>68</td>
<td>41</td>
<td>30</td>
</tr>
<tr>
<td>% Post-ICH57B</td>
<td>100%</td>
<td>49%</td>
<td>10%</td>
<td>36%</td>
</tr>
</tbody>
</table>

Studies with all 4 endpoints: 28
Studies with hERG, ECG, & TQT: 57

*Preliminary analysis is discussed in subsequent slides*

- 22 drugs: 10 positive TQT, 12 negative TQT
Concordance to Date

- Encouraging concordance for the in vivo study to date
- The poorer specificity of hERG impacts the integrated assessment
- APD assays have a strong positive predictive value but low negative predictive value – been abandoned by many

### Preliminary Concordance Summary (22 Drugs)

<table>
<thead>
<tr>
<th></th>
<th>hERG</th>
<th>1X (n=13)</th>
<th>3X (n=13)</th>
<th>10X (n=13)</th>
<th>30X (n=12)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specificity</td>
<td>0.40</td>
<td>0.40</td>
<td>0.25</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>Sensitivity</td>
<td>0.63</td>
<td>0.63</td>
<td>0.78</td>
<td>0.89</td>
<td></td>
</tr>
<tr>
<td>Concordance</td>
<td>0.54</td>
<td>0.54</td>
<td>0.62</td>
<td>0.67</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>APD</th>
<th>1X (n=6)</th>
<th>3X (n=9)</th>
<th>10X (n=9)</th>
<th>30X (n=8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specificity</td>
<td>0.67</td>
<td>0.80</td>
<td>0.60</td>
<td>0.60</td>
<td></td>
</tr>
<tr>
<td>Sensitivity</td>
<td>0.0</td>
<td>0.0</td>
<td>0.25</td>
<td>0.33</td>
<td></td>
</tr>
<tr>
<td>Concordance</td>
<td>0.33</td>
<td>0.44</td>
<td>0.44</td>
<td>0.50</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>QTC</th>
<th>1X (n=8)</th>
<th>3X (n=15)</th>
<th>10X (n=13)</th>
<th>30X (n=10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specificity</td>
<td>1.0</td>
<td>1.0</td>
<td>0.83</td>
<td>0.75</td>
<td></td>
</tr>
<tr>
<td>Sensitivity</td>
<td>1.0</td>
<td>0.57</td>
<td>0.71</td>
<td>0.83</td>
<td></td>
</tr>
<tr>
<td>Concordance</td>
<td>1.0</td>
<td>0.80</td>
<td>0.77</td>
<td>0.80</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Integrated</th>
<th>1X (n=18)</th>
<th>3X (n=20)</th>
<th>10X (n=19)</th>
<th>30X (n=18)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specificity</td>
<td>0.56</td>
<td>0.64</td>
<td>0.44</td>
<td>0.38</td>
<td></td>
</tr>
<tr>
<td>Sensitivity</td>
<td>0.56</td>
<td>0.67</td>
<td>0.90</td>
<td>0.90</td>
<td></td>
</tr>
<tr>
<td>Concordance</td>
<td>0.56</td>
<td>0.65</td>
<td>0.68</td>
<td>0.67</td>
<td></td>
</tr>
</tbody>
</table>
In Vivo Study Quality an Issue

• Note the variability and shape of the in vivo QTc exposure-response relationship
• These changes were however all large and statistically significant
ILSI/HESI Data Included (Approach 3)

- Note the low variability in the ILSI/HESI data
- Sadly not representative of many submitted studies

Example of positive QTc compound: Cisapride

[Graph showing changes in hERG inhibition, in vitro APD, in vivo QTc, and human QTc with increasing concentration of the compound]

- Clinical data: van Haarst et al., Clin Pharmacol Ther 1998;64:542-6
1. Quantitative method to relate pre-clinical safety pharmacology data from dog telemetry (CVS), rodent Irwin/FOB (CNS) and rodent plethysmography (respiratory) models to Phase I clinical outcome in man
   • Valentin et al., 2009 JPTM 60 152-158

2. Retrospective analysis of small molecule compounds from 7 pharmaceutical companies under the umbrella of the Association of the British Pharmaceutical Industry (ABPI)
   • AstraZeneca, Amgen, Janssen, GSK, Pfizer, Eli Lilly, Novartis

3. Parameters from models will be mapped to a two dimensional framework
   • Confidence in the model
   • Confidence in the translation
The dog telemetry model adequately predicts QTc changes in man based on the data analysis of the 114 compounds in the Animal Model Framework.
How can Assay Sensitivity be Demonstrated in Early Clinical Studies?

- A more precise term for ‘Assay Sensitivity’ for experimental science is ‘assay validity’
- Positive controls for QT measurement were prescribed due to a lack of validation
- A positive control is unnecessary in a sufficiently valid study, and may worsen the predictive value of the study conclusions.
‘Assay Sensitivity, Failed Clinical Trials, and the Conduct of Science’

“Subsequent, [negative] findings are then assumed to be the result of some failure of ‘assay sensitivity’ of the trial.”

“This reasoning, however, has the potential of distorting the scientific process, such that the adequacy of the trial is judged not by the design but, instead, by the results of the trial itself.”

Data from the FDA IRT

- Based on 178 TQT studies we have reviewed, there were 9 studies which failed the assay test.
- Among those 9, most of them didn't show moxi time profile instead of failed 5 ms lower bound test.
- Many of these were over encapsulated (blinded) moxi

Sensitivity?

Increased false negative rate
Why is QT special?

- Other biomarkers are important, have significant variability and not required to have a positive control to demonstrate lack of an effect:
  - LDL, HCT, LFTs, etc.
  - We consistently employed ‘tried and true’ methods of assay validation
  - QT shouldn’t be a special case
  - Sensitivity (false negative rate) is well controlled
  - Risk/benefit information is not well supported by positive control or additional TQT
How can Assay Sensitivity be Demonstrated in Early Clinical Studies?

- Assay Validity has been demonstrated for many early phase clinical studies
  - Internal, Construct, External, Statistical conclusion validity have been established for QT studies
- Further ‘validation’ techniques
  - Malik method – variability, stability, reproducibility
  - Retrospective power analysis
    - A priori establishment of exclusion of 10msec
  - Empiric – past positive result


Slope of relationship between LY concentration and QTc = 0.0003 (90%CI: -0.0028 - 0.0034), p=0.891

Assuming above slope is the true mean relationship, a LY concentration of approximately 17000 ng/ml would be required to produce a mean QTc prolongation of 5 ms. This concentration is approximately 15-fold higher than the mean Cmax.

Since slope is positive but very small and the confidence interval includes zero it can be concluded that there is no clinically significant or statistically significant relationship between LY concentration and QTc.
Relationship Between LY Concentration and QTcI Change from Baseline

Slope of relationship between LY concentration and QTcI change from baseline = -0.000055 (90%CI: -0.00205 - 0.00194), p=0.964

Since slope is very small, negative, and the confidence interval includes zero it can be concluded that there is no clinically significant or statistically significant relationship between LY concentration and QTcI change from baseline.

Probability statements based on confidence intervals:

- Concentration at which true mean change in QTcI is 95% likely to be less than 5ms: 3370 ng/ml
- At mean Cmax observed at 600mg (1300 ng/ml), probability that the true mean increase in QTcI is >5ms is <0.01%
- At 5x mean Cmax observed at 600mg (6500 ng/ml), probability that the true mean increase in QTcI is >5ms is 18.5%
Methods of Establishment of Assay sensitivity:

- How do we know the slope isn’t significantly greater than in your study?
- Or if the same experimental system demonstrates an approximately flat response –
- “[the positive control] is needed to ensure that the study is properly designed and conducted and able to detect small changes in QTc.”
- **TQT STUDY IS NECESSARY?**
Agreement of Pooled Phase I/II C-QT Models With Negative TQT Studies

- "The C-QT analysis was dominated by a single phase I study that had serial triplicate ECGs taken during periods with substantial drug concentrations"
- double-blind, randomized, parallel thorough QTc study evaluated AD1 at 3 mg and 10 mg (2× and 7× therapeutic dose), placebo, and moxi 400 mg for 7 days in 140 healthy participants (85 men and 55 women).
Is a Thorough QTc Study Necessary? The Role of Modeling and Simulation in Evaluating the QTc Prolongation Potential of Drugs*

- Agreement of Pooled Phase I/II C-QT Models With Negative TQT Studies

- AD1 3 mg and 10 mg were noninferior to placebo at every time point postdose (ΔΔQTcF <5 ms with upper 95% CI <10 ms).

Figure 8. Negative slope observed in thorough QTc (TQT) study for AD1.

Is a Thorough QTc Study Necessary? The Role of Modeling and Simulation in Evaluating the QTc Prolongation Potential of Drugs*

• Agreement of Pooled Phase I/II C-QT Models With Negative TQT Studies
  • “it is reasonable to ask whether a TQT is necessary in the context of a well-constructed C-QT data set and analysis. Pooled single and multiple ascending-dose (SAD/MAD) data sets will span a dose range sufficient to cover and exceed the “supratherapeutic” dose of the TQT design.”
  • “Two main arguments against this idea can be expected.”
    • active control arm of moxifloxacin
    • lack of power - findings from the pooled analysis could well be a “false negative” due to insufficient power
Is a Thorough QTc Study Necessary? The Role of Modeling and Simulation in Evaluating the QTc Prolongation Potential of Drugs*

Agreement of Pooled Phase I/II C-QT Models With Negative TQT Studies

• “it is reasonable to ask whether a TQT is necessary in the context of a well-constructed C-QT data set and analysis. Pooled single and multiple ascending-dose (SAD/MAD) data sets will span a dose range sufficient to cover and exceed the “supratherapeutic” dose of the TQT design.”
  • The first argument would maintain that without an active control arm of moxifloxacin, there would be no way to gauge the sensitivity of study participants to QT-prolonging drugs.
    • parallel-arm TQT studies are allowed, the issue of assay validation cannot be population sensitivity
• “At the study level, standard procedures for ECG assessment and manual reading at a central laboratory can be implemented. In practice, manually read ECGs have shown stable results for moxifloxacin, so there should not be a concern for labs with experienced ECG readers. Alternatively, a small moxifloxacin comparative arm may be added in a SAD or MAD study.”

*Rohatagi, Shashank, Carrothers, Timothy J., Kuwabara-Wagg, Jon, Khariton, Tatiana
September 4, 2009 J Clin Pharmacol
Methods of Establishment of Assay sensitivity:

- “In the case of a *negative or neutral* relationship between the drug concentration and QTc, a positive control is needed to assure assay sensitivity.”

- **What is meant by ‘Assay Sensitivity’?**
A more precise term for ‘Assay Sensitivity’ is ‘Validity’

- Validity
  - Internal validity
  - Construct validity
  - External validity
  - Statistical conclusion validity*
    - Type I Error
    - Power (sensitivity)

*Note: Some treat Type II error statistical conclusion validity as a separate topic called assay sensitivity.
Is a Thorough QTc Study Necessary? The Role of Modeling and Simulation in Evaluating the QTc Prolongation Potential of Drugs*

• Agreement of Pooled Phase I/II C-QT Models With Negative TQT Studies
  • lack of power - findings from the pooled analysis could well be a “false negative” due to insufficient power

*Rohatagi, Shashank, Carrothers, Timothy J., Kuwabara-Wagg, Jon, Khariton, Tatiana
September 4, 2009 J Clin Pharmacol
Simulation Case Study

“Assuming a drug indeed prolongs the QTc interval, what are the chances that a C-QT model based on a multiple ascending-dose data set will not detect the effect?”

- Compound with 5 msec prolongation (truth)
- Each QTc measurement, including baseline, with intrasubject (ie, residual) error with a mean of 0 ms and a standard deviation of 15 ms.
- MAD study, each arm with 10 participants, serial triplicate ECGs taken during periods with substantial drug concentrations.
- 5000 replicates of the trial were simulated

*Rohatagi, Shashank, Carrothers, Timothy J., Kuwabara-Wagg, Jon, Khariton, Tatiana
September 4, 2009 J Clin Pharmacol
Is a Thorough QTc Study Necessary? The Role of Modeling and Simulation in Evaluating the QTc Prolongation Potential of Drugs*

Figure 11. Distribution of estimated prolongation at typical $C_{\text{max}}$ based on modeled C-QT slope in the “5-ms positive” exercise. Note that the dotted line corresponds to 5 ms “true state of world.”

*Rohatagi, Shashank, Carrothers, Timothy J., Kuwabara-Wagg, Jon, Khariton, Tatiana
September 4, 2009 J Clin Pharmacol
Is a Thorough QTc Study Necessary? The Role of Modeling and Simulation in Evaluating the QTc Prolongation Potential of Drugs*

Figure 12. Distribution of 95% lower confidence limit (LCL) prolongation at typical $C_{\text{max}}$ based on modeled C-QT slope in the “5-ms positive” exercise.

*Rohatagi, Shashank, Carrothers, Timothy J., Kuwabara-Wagg, Jon, Khariton, Tatiana
September 4, 2009 J Clin Pharmacol
even in the case where the C-QT model data set consists of only a multiple ascending-dose study with 40 participants, the C-QT model estimated a statistically significant, positive slope 99% of the time

FN rate < 1%

“if a C-QT model based on a data set of similar or greater richness and size estimates a slope nonsuperior to 0, the results can be taken with confidence that a regulatory-meaningful effect does not exist at the dose studied”
Comments about the positive control

• “the specter of the positive control helps assure good trial performance”
• “... the trials became larger and more carefully conducted as a result of insisting upon the positive control.”
• “By having an operational assessment of study quality, [the positive control] frees sponsors to explore more efficient study designs and measurement technology.”


Charles Benson, M.D.Ph.D.
Copyright © 2009 Eli Lilly and Company