Leveraging data standards and the QT technical specifications document to break down data silos, enable AI research, and advance regulatory science

John Zoscak, Rejaul Karim, Lars Johannesen, Jose Vicente,

CDER/OND/OCHEN/DCN

Abstract

- **Variability in clinical QTc data for moxifloxacin, a drug that prolongs the QTc interval in the electrocardiogram (ECG), has been previously reported across studies using different ECG methods but the impact of ECG methods in this variability is unknown.**
- **Clinical QTc datasets were recoded following FDA QT technical specifications document (QT-TSD). An artificial intelligence (AI) ECG model was trained using the associated annotated ECG waveforms (aECGs).**
- **The AI ECG model and a previously published ECG algorithm (ecglib) were tested in a set of studies not included in training and compared to the submitted data.**
- **Results suggest that different ECG methods do not significantly contribute to between-study variability in moxifloxacin QTc data.**
- **The use of QT-TSD datasets breaks down data silos, allowing efficient comparison of ECG methods and development and evaluation of AI ECG models.**

Introduction

- **QTc interval prolongation in the electrocardiogram (ECG) is associated with a potentially fatal heart rhythm.**
- **Annotated ECG waveforms (aECGs) from drug-safety clinical studies conducted to assess QTc prolongation potential of new drugs are submitted to FDA for review.**
- **There is no gold standard for measuring the QT interval and different ECG methods to measure QT can result in different but still valid QT measurements (Figure 1).**

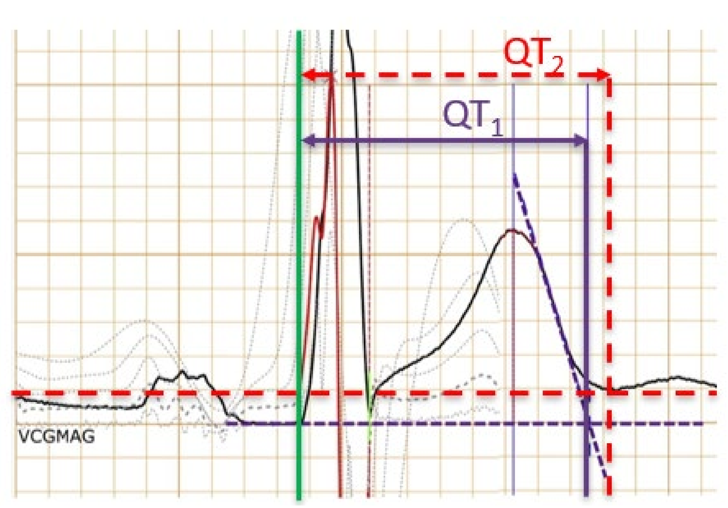

Figure 1. Different ECG methods can result in different QT interval values. Vertical lines show QT interval measured from the beginning of QRS (green) to the end of the T-wave using two different methods (QT1: purple; QT2: red). Adapted from Vicente J, J Electrocardiol 2018.

- **Variability in clinical QTc data for moxifloxacin, a QTc prolonger, has been previously reported across studies using different ECG methods, but the impact of ECG methods in this variability is unknown.**
- **To investigate this question:**
  - Clinical QTc datasets were recoded following FDA QT technical specifications document (QT-TSD) and an artificial intelligence (AI) ECG model was trained on the associated aECGs.
  - The AI ECG model and a previously published ECG algorithm (ecglib) were tested in a set of studies not included in training and compared to the submitted data.

Materials and methods

- **A previously developed AI ECG model capable of learning from multiple ECG annotations methods at the same time (Figure 2) was trained on 400,000+ aECGs from 45 QT studies.**
  - aECGs from 5% of subjects from each study held out for cross validation
  - AI ECG model input: 8-lead ECG 1kHz 1,350 msec long waveform
  - AI ECG model output: QT interval coded as one-hot-coded (softmax) annotations for start (Qonset) and end (Toffset) of the QT interval
  - AI ECG model network: 1 dense + 1 bidirectional long short-term memory (BiLSTM) + 3 dense layers
  - AI ECG model implemented in python (tensorflow and keras)

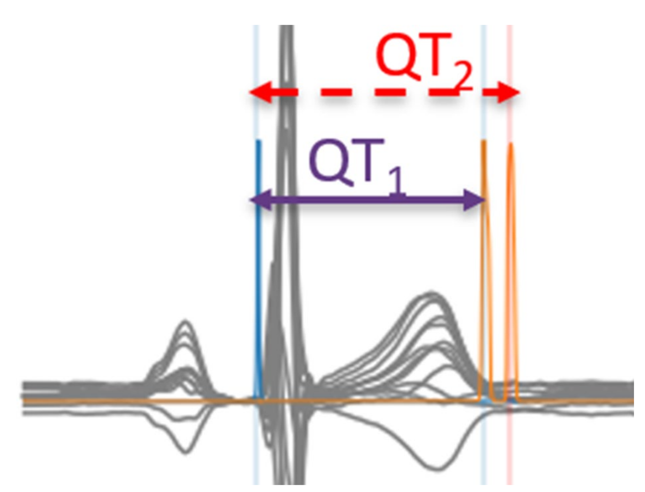

Figure 2. Blue and orange traces show AI ECG model predictions for Qonset and Toffset, respectively, vs. the original annotations (vertical lines). Adapted from slides by Vicente J, FDA Scientific Computing Days 2019.

- **Test dataset: 38,000+ aECGs and associated clinical datasets from 6 QT studies**
  - Submitted datasets (e.g., CDISC SDTM/ADAM) recoded to QT-TSD ADEG, ADPC, and ADSL
  - ADEG including reference to associated aECGs per ECG method (i.e., sponsor submitted annotations, ecglib, and AI ECG model output)
  - Analysis-ready datasets (i.e., QTPK) derived from QT-TSD datasets

- **AI ECG performance and ECG methods comparison**
  - QT bias = AI ECG QT – Sponsor submitted QT values
  - Comparison of the slope of the concentration-QTc (CQTc) relationship
    - CQTc analysis following the scientific white paper on CQTc modeling (Garnett et. al, J Pharmacokinet Pharmacodyn 2018)

Results and discussion

- **Results showed that the mean difference between the original and AI QT values (QT bias) was around -4 msec in the test dataset, with larger variability in the test vs. the training datasets (Table 1).**

Dataset	Mean QT Bias (msec)	Standard Deviation (msec)	ECGs (n)
Training	0.2	10.8	407,408
Crossvalidation	0.2	7.9	40,423
Test	-3.9	29.9	38,180

Table 1. QT bias (AI ECG QT – Sponsor submitted QT values) assessment results

- **Overall, by-time and CQTc results were similar across the three ECG methods. See Figures 3, 4, and 5.**

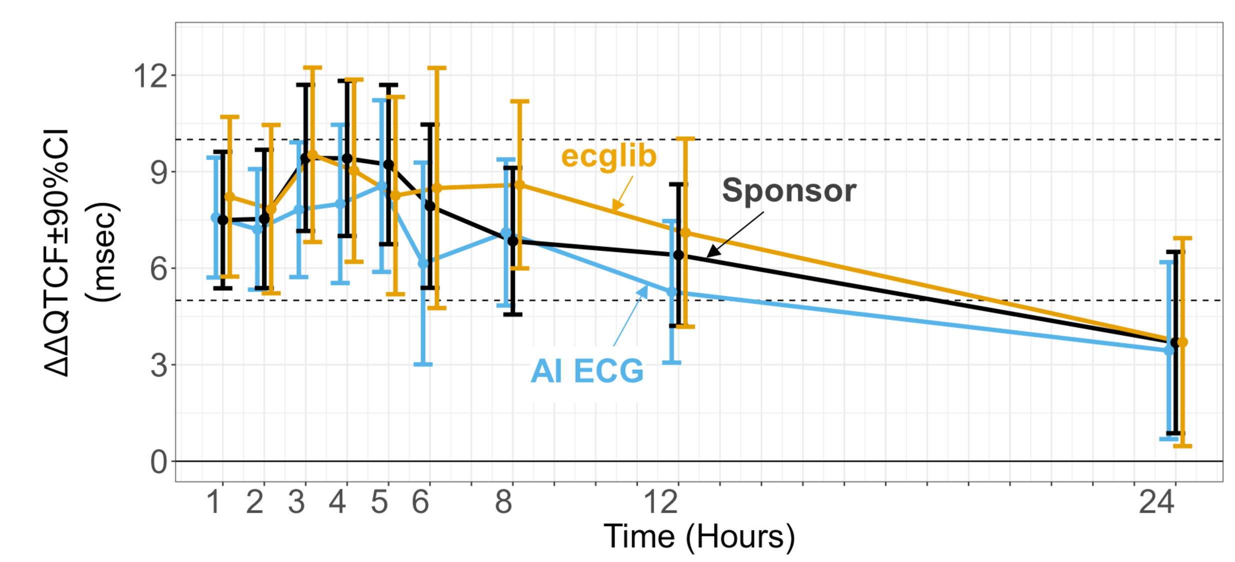

Figure 3. Moxifloxacin mean (dots) and 90% confidence intervals (error bars) placebo corrected changes from baseline (ΔΔ) QTc by time point and ECG method in one of the studies in the test dataset.

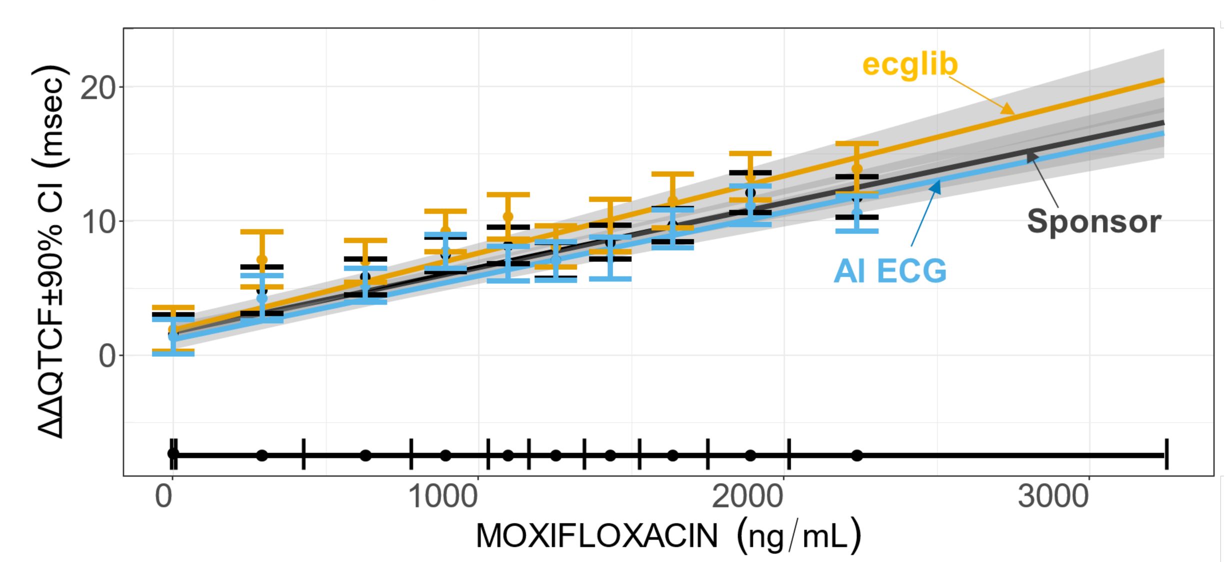

Figure 4. Goodness of fit plot for QTc in one of the studies in the test dataset. Mean (dots) and 90% confidence intervals (error bars) placebo corrected changes from baseline (ΔΔ) QTc vs. moxifloxacin concentration per ECG method. Solid line and gray areas are the mean and 90% CI predictions from the CQTc models for each ECG method.

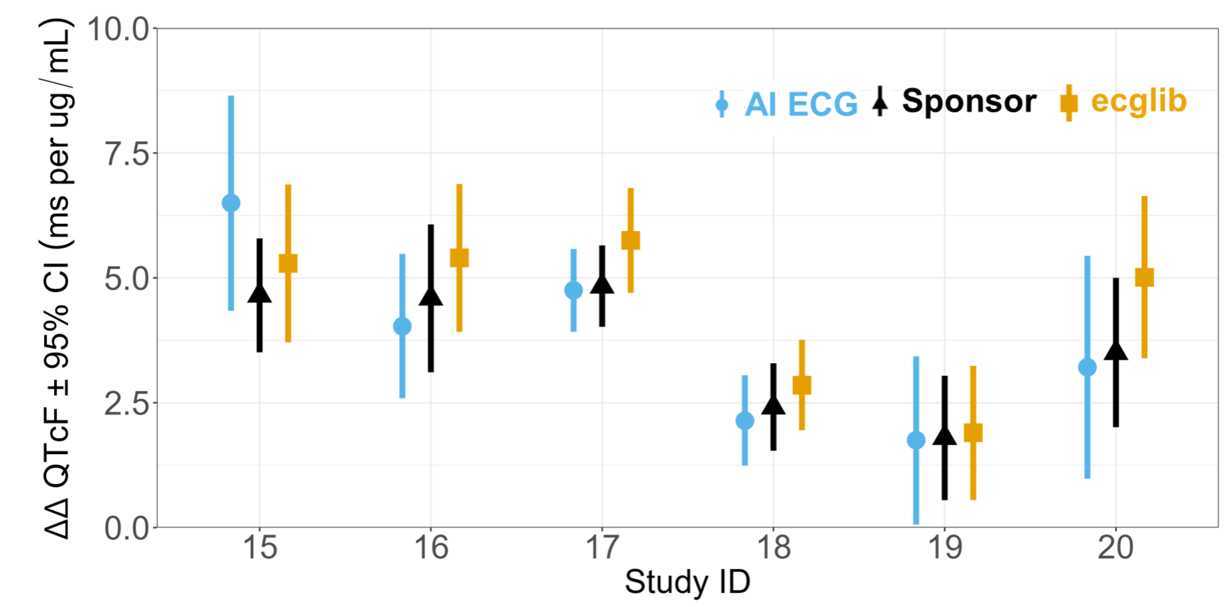

Figure 5. CQTc model slope estimates and associated 95% CIs for moxifloxacin from the 6 studies in the test dataset.

- **Large QT bias values (> 2xSD=59.7 msec) were observed in some individual waveforms (0.5 %) in the test dataset. This tended to be in cases where the ECG waveforms were difficult to annotate (e.g., low amplitude, Figure 6).**

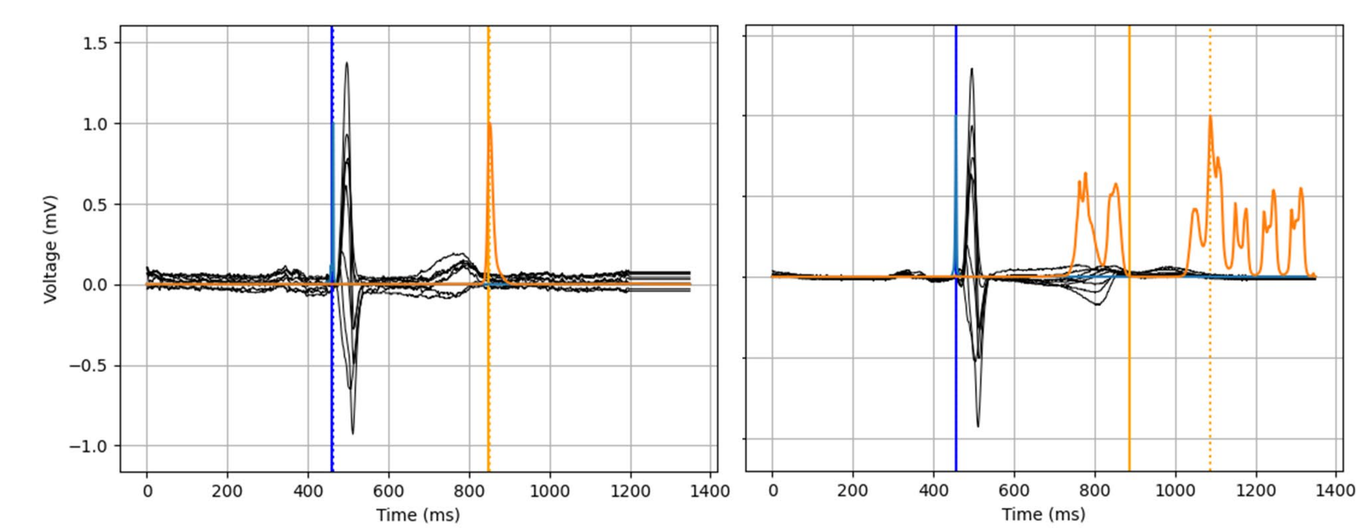

Figure 6. AI ECG model outputs in presence of small Twaves. The softmax output for Toffset (orange) on the right panel shows more uncertainty than the left.

Conclusion

- **Standardized data using the QT-TSD enabled:**
  - **Development of an AI ECG method that can reliably measure the QTc interval.**
  - **Efficient comparison of ECG methods showing that ECG methods do not significantly contribute to between-study variability.**
- **Standardizing datasets breaks down data silos and enables AI research and advances regulatory research.**

Acknowledgments: This project was supported in part by appointments to the Research Fellowship Program at the OCHEN/OND/CDER and DARS/OCP/OTS/CDER U.S. Food and Drug Administration, administered by the Oak Ridge Institute for Science and Education through an interagency agreement between the U.S. Department of Energy and FDA.