Human Overread of Semi-Automated QT Measurements May Adversely Affect Final Results In Cardiac Safety Studies

Branislav Vajdic, PhD, Ihor Gussak, MD PhD FACC, Bosko Bojovic, PhD, Samuel George, MD FACC

NewCardio, Inc, Santa Clara CA 94054

Abstract

Background: Semi-automated QT interval measurement is the most commonly used method for "thorough QT studies" (TQTS) and other clinical drug development studies. In semi-automated methods, intervals are measured by computer and then every ECG tracing is overread by trained human readers. Typically, human readers change about 30% of the machine-placed QT annotations during the overread process. The impact on study results of these changes may be substantial: a recent TQT study of tolterodine showed that human alterations changed the overall mean Fridericia-corrected, drug-induced QT change (QTcF) by over 10 ms.

Methods: We determined QT and RR intervals on 1963 ECGs from a recent Phase I clinical drug trial by 3 methods: a global median-beat computer-assisted method performed by a core lab ("global," arbitrarily declared the "gold standard"), manual reading a highly qualified cardiologist ("manual") and fully automated determinations using the QTinno ECG analysis software ("QTinno"). To model the effect of human overread, we generated a mixed set of QT results by replacing 507 randomly-selected QTinno results (25.8% of all readings) with the corresponding 507 manual QT results ("mixed"). This mixed group is, to the first order a good approximation of what happens in the semi-automated method.

Results: The 3 methods returned virtually identical results for placebo-corrected, time-matched change in QTcF (ΔΔQTcF). Specifically, mean pairwise difference for three methods was: QTinno vs. global <0.3 ms, QTinno vs. manual <0.6 ms and global vs. manual <0.4 ms. However, the mixed set showed a marked change for drug–induced QT change when compared with QTinno results. This resulted in substantially different ΔΔQTcF (mean pairwise difference) for global vs. mixed: it increased substantially from <0.4 ms to 7.9 ms.

Conclusions: This model suggests that mixing human overread with machine generated QT measurements may introduce unexpected distortions in key cardiac safety metrics determined by semi-automated methods. The theoretical treatment offered here supports the results of the tolterodine TQTS. The root cause of the phenomenon lies in the fact that humans tend to read QT interval consistently shorter than modern computer algorithms. Impact of mixing results as well as influence of the percentage of ECG readings changed by manual readers will be discussed .

Introduction

Many drugs have been shown to prolong cardiac repolarization and increase risk of torsades de pointes arrhythmia. Early identification of such effects is a critical priority. The ICH E14 Guidance for Industry states that virtually all new chemical entities should have a "thorough QT study" (TQTS) - a single highly powered trial designed to identify drugs that prolong cardiac repolarization, by evaluating the effect on rate-corrected QT interval (QTc) at multiple drug doses and time points.

At present, cardiac interval measurement for drug development is performed using digital ECGs and electronic calipers on a high-resolution computer screen. Measurements may be fully manual, but are more commonly "semi-automated" - that is, automated QT determination by computerized algorithm, followed by human overread and correction.

It is well established that manual determinations by human readers are subject to substantial intra- and inter-reader variability that can significantly affect key cardiac safety metrics. Routine use of human overread and adjustment may have a similar impact on semi-automated measurements.

Recently, Malhotra et al (Clin Pharm Thera 81:377-385, 2007) documented this effect using positive-control ECGs from a Thorough QT Study. They showed that mean placebo-corrected change in Fridericia-corrected QT intervals (ΔΔQTcF) after a single dose of moxifloxacin was 8.9 ms using unadjusted semi-automated results, and 19.3 ms if the results were overread and adjusted by humans readers. This remarkable difference is a cause for concern as it may result in significant new regulatory requirements for cardiac safety data in Phase II and III clinical trials. Such a difference may add substantial time and cost to later-phase clinical trials.

Study Objective

We undertook this study to better understand the potential negative impact of human overreads in semi-automated methods, and to evaluate their possible effects in a novel experimental model distinct from Malhotra et al.

- Hypothesis: *Human adjustment of a subset of semi-automated results will distort determination of key cardiac safety metrics*
  - Introduces intra- and inter-reader variability
  - Unpredictably mixes results from 2 different approaches that produce different results for absolute QTcF and drug-induced change in QTcF

Study Design

- Initial study was placebo-controlled, parallel-group, double-blind, Phase I multiple ascending dose study with a quinolone antibiotic (1963 ECGs)
- Compared QT and RR intervals obtained by 3 methods:
  - Global median-beat semi-automated method, overread and adjusted by 3 core lab cardiologists (global)
  - careful manual reading by 2 highly qualified cardiologists on 2 occasions (manual) and
  - fully automated determinations using the QTinno ECG analysis software (QTinno)
- Mean absolute QTcF and time-matched change in QTcF (ΔQTcF) compared by Bland-Altman analysis

Experimental Model

- Generate "mixed-mode" dataset by randomly replacing 507 QTinno-generated QTcF determinations with corresponding manual determinations (about 25% of total datapoints)
- Determine absolute QTcF and ΔQTcF on the mixed-mode dataset and compare to original results from 3 methods

Results

1. Comparison of Absolute QTcF by Measurement Method

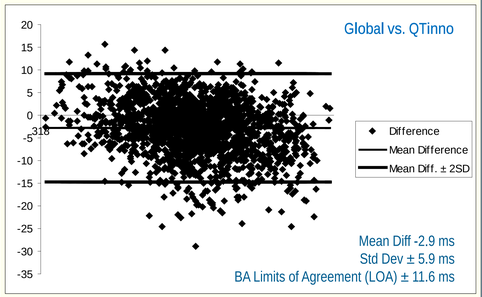

Fig.1. Bland-Altman plot of difference between Global and QTinno – Absolute QTcF values

Results

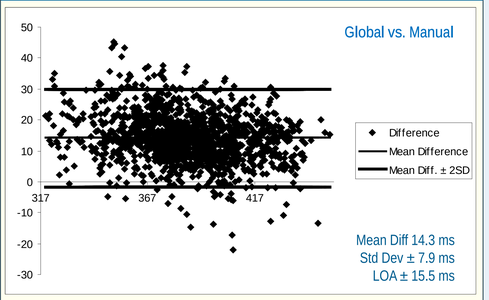

Fig. 2. Bland-Altman plot of difference between Global and Manual Methods – Absolute QTcF values

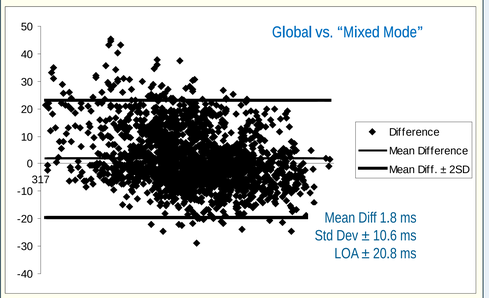

Fig. 3. Bland-Altman plot of difference between Global and "Mixed Mode" – Absolute QTcF values

Summary of Absolute QTcF Results by Measurement Method

- QTinno and Global results comparable, whereas manual results were substantially shorter
- Shows human readers read shorter when compared to fully automated and semi-automated methods
- The "mixed mode" set returned results that were similar to those of semi-automated and QTinno, showing that mixing randomly selected automated reads with manual reads does not have a marked impact on absolute QTcF determination

Results

2. Comparison of Drug-Induced Change in QTcF (ΔQTcF) by Measurement Method

- Mean ΔQTcF compared by Bland-Altman analysis:
  - QTinno vs. Global: mean difference -0.1 ± 6.5 ms
  - Global vs. Manual: mean difference -0.3 ± 6.2 ms
  - QTinno vs. Manual: mean difference -0.3 ± 6.6 ms
- Although absolute QTcF was shorter by manual method, results for ΔQTcF comparable by all 3 methods
- Thus, differences in measurement methods do not affect key cardiac metrics, *as long as measurement methods are consistent*

3. Effect of Mixing Manual and Automated Data on ΔQTcF

- Mean ΔQTcF compared by Bland-Altman analysis:
  - QTinno vs. Global: mean difference -0.1 ± 6.5 ms
  - Global vs. Manual: mean difference -0.3 ± 6.2 ms
  - QTinno vs. Manual: mean difference -0.3 ± 6.6 ms
  - Global vs. Mixed-Mode: mean difference 7.9 ± 10.6 ms

Discussion and Conclusions

- Mixing 2 methods – manual and automated QT determination – does not substantially affect absolute QTcF determination, but does produce a marked alteration in mean ΔQTcF, a key cardiac safety metric
- This is true even though either method would have yielded acceptable results if left alone
- We believe this is a reasonable first approximation of a possible negative impact of manual adjustment in semi-automated methods
- Our results are consistent with an earlier study by Malhotra et al., who showed that manual adjustment of semi-automated data markedly altered mean ΔQTcF and ΔΔQTcF in a recent Thorough QT Study
- Additional study is needed to quantify this effect and evaluate the impact of manual overreading