An Evaluation of Multiple Algorithms for the Measurement of the Heart Rate Corrected JTpeak Interval

Jean-Philippe Couderc, Shiyang Ma, Alex Page, Connor Besaw, Jean Xia, W. Brian Chiu PhD, Johan de Bie, Jose Vicente, Martino Vaglio, Fabio Badilini, Saeed Babaeizadeh, Cheng-hao Simon Chien, Mathias Baumert

PII: S0022-0736(17)30266-2
DOI: doi: 10.1016/j.jelectrocard.2017.08.025
Reference: YJELC 52485

To appear in: Journal of Electrocardiology

Please cite this article as: Couderc Jean-Philippe, Ma Shiyang, Page Alex, Besaw Connor, Xia Jean, Brian Chiu W, de Bie Johan, Vicente Jose, Vaglio Martino, Badilini Fabio, Babaeizadeh Saeed, Chien Cheng-hao Simon, Baumert Mathias, An Evaluation of Multiple Algorithms for the Measurement of the Heart Rate Corrected JTpeak Interval, Journal of Electrocardiology (2017), doi: 10.1016/j.jelectrocard.2017.08.025

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.
An Evaluation of Multiple Algorithms for the Measurement of the Heart Rate Corrected JTpeak Interval

Jean-Philippe Couderc, Shiyang Ma, Alex Page, Connor Besaw, Jean Xia, W. Brian Chiu, PhD; Johan de Bie, Jose Vicente, Martino Vaglio, Fabio Badilini, Saeed Babaeizadeh, Cheng-hao Simon Chien, Mathias Baumert

a Heart Research Follow-up Program, University of Rochester Medical Center, Rochester, NY 14618
b Department of Biostatistics and Computational Biology, University of Rochester Medical Center, Rochester, NY 14618
c School of Electrical and Electronic Engineering, The University of Adelaide, SA5005, Australia
d Mortara Instrument Europe s.r.l., Via Cimarosa 103, 40033 Casalecchio di Reno (BO), ITALY
e Division of Cardiovascular and Renal Products, Office of Drug Evaluation I, Office of New Drugs Center for Drug Evaluation and Research, U.S. Food and Drug Administration
f AMPS LLC, 418 Central Park W, Suite 21C, 10025 - New York NYC, U.S.A
g Advanced Algorithm Research Center, Philips Healthcare, Andover, Massachusetts

This article reflects the views of the authors and should not be construed to represent FDA’s views or policies.

Contact Information:
Jean-Philippe Couderc, PhD , MBA
Heart Research Follow-up Program, University of Rochester Medical Center, Rochester, NY 14618
USA
Phone: 585-275-1096;
Fax: 585-273-5283;
e-mail: heartjpc@heart.rochester.edu.
Keywords: electrocardiogram, cardiac safety, QT, long QT syndrome, CiPA, torsade de pointes.
Abstract

Interest in the effects of drugs on the heart rate-corrected JTpeak (JTpc) interval from the body-surface ECG has spawned an increasing number of scientific investigations in the field of regulatory sciences, and more specifically in the context of the Comprehensive in vitro Proarrhythmia Assay (CiPA) initiative. We conducted a novel initiative to evaluate the role of automatic JTpc measurement technologies by comparing their ability to distinguish multi- from single-channel blocking drugs. A set of 5,232 ECGs was shared by the FDA (through the Telemetric and Holter ECG Warehouse) with 3 ECG device companies (AMPS, Mortara, and Philips). We evaluated the differences in drug-concentration effects on these measurements between the commercial and the FDA technologies. We provide a description of the drug-induced placebo-corrected changes from baseline for dofetilide, quinidine, ranolazine, and verapamil, and discuss the various differences across all technologies.

The results revealed only small differences between measurement technologies confirming that the JTpc interval distinguishes between multi- and single-channel (hERG) blocking drugs when evaluating the effects of dofetilide, quinidine, ranolazine, and verapamil. In the case of quinidine and dofetilide, we noticed a poor consistency across technologies because of the lack of standard definitions for the location of the peak of the T-wave (T-apex) when the T-wave morphology is abnormal.
Introduction

The comprehensive in vitro proarrhythmia assay (CiPA) initiative proposes a new strategy to shift the emphasis away from QT prolongation and focus on predicting torsadogenic hazard at an early stage of drug development [1]. CiPA emerged to address the issue of false positive detection of drugs that combine hERG potassium channel block with additional multi-channel effects that have the potential to balance out the deleterious effect on the repolarizing current [2]. CiPA consists of four components: 1) the in-vitro assessment of the drug effect on multiple human cardiac currents, 2) the in-silico reconstruction of the ventricular cellular electrophysiology, 3) the evaluation of the in-vitro effects in human stem-cell derived ventricular myocytes, and 4) the clinical evaluation of electrocardiographic effects of the drug.

The work described below pertains to this fourth component, and more specifically to the evaluation of a proposed novel ECG marker [3-5], viz., the drug-induced JTpeak interval (JTp) prolongation. The JTp interval is a sub-interval of the QT interval. JTp measures the time lapse between the end of the QRS (J point) and the apex of the T-wave (Tpeak). JTp interval prolongation is not clinically used, and only a limited number of publications have investigated its value, heart-rate dependency, and clinical significance [3-6]. Importantly, its electrophysiological meaning remains to be fully understood [7].

In drug studies, the automation of the measurement of the QT/QTc interval has been widely adopted, and when combined with a proper adjudication process it has led to improved precision and an improved overall paradigm of QT safety assessment [8]. Following the introduction of CiPA, the automatic measurement of the JTp interval has gained some interest in this specific context. Therefore, one needs to understand better the strengths and weaknesses of
the technologies used for this measurement. We explored the concordance of JTp intervals when measured by various algorithms. To conduct such an exercise, we invited various stakeholders to an open initiative to apply their JTp measurement technique on a reference set of ECG recordings that was shared through the Telemetric and Holter ECG Warehouse [9,10]. This dataset was generated during a study conducted by the FDA to test the hypothesis that hERG potassium channel block prolongs QTc by prolonging both the heart-rate corrected JTp (JTpc) (early repolarization) and Tpeak–Tend (late repolarization) intervals on the ECG, whereas the addition of calcium and/or late sodium current block preferentially shortens JTpc prolongation caused by hERG block [4,5,11]. We report the drug concentration dependency of the JTpc intervals measured by the various technologies, and assess whether JTpc could distinguish between predominantly and strongly hERG-blocking drugs (quinidine and dofetilide) and drugs that have balanced ion channel (hERG and late sodium and/or calcium) blocking properties (ranolazine and verapamil).

**Methods**

**The ECG dataset**

A set of 5,232 ECGs was recorded during a randomized, crossover, double-blinded, placebo controlled clinical study (NCT01873950) [3]. The study involved 22 healthy subjects (11 women) with ages ranging from 18 to 35 years old. The subjects had 5 visits during which they were exposed to a single dose of dofetilide, quinidine, ranolazine, verapamil, or placebo. Sixteen time points were scheduled per visit, and three 10-second 12-lead ECGs were extracted at each time point while subjects were resting in a supine position. Additionally, blood samples
were acquired for plasma drug concentration (PK analysis). The Mason–Likar lead configuration was used. The ECG files were made available in ISHNE format from the Telemetric and Holter ECG Warehouse (thew-project.org) and in Physionet format from the Physionet website (physionet.org). Importantly, the participants in our initiative had access to the measurements of the QRS onset, J point, T-end location and T-apex locations measured and published by the FDA [3]. The method adopted by FDA for producing these measurements was semi-automatic, i.e., adjudication of the T-apex location was visually reviewed and manually adjusted by FDA personnel. The participants of the JTp initiative had the choice between using the original raw 10-second ECG signals or their corresponding median beat signals already available in the dataset. More information about the methods used by the participants to automatize JTp interval measurement is provided in this issue of the journal [12-14].

**ECG measures provided by the participants**

Participation in the initiative required the registrants to deliver a set of measurements defined as follows: one single time interval value for the RR, JTp, RJ and RTpeak (RTp) for each ECG file. The computation of the JTp intervals was based on the formula published by Johannesen et al. [3]: JTp = JTp/(RR)^{0.58}. The same heart rate correction formula was used to correct the RTp interval (RTpc). We assume that the R peak location is the most stable fiducial point in most tested technologies. Furthermore, the tested drugs are not known to have an effect on the QRS complex at the concentrations in the study [3]. We therefore used the R peak as a reliable anchor point to study the variability of the J and T-apex locations using the reported RJ and RTp intervals.
Three leads were required to be measured: leads II and V5, and the vector magnitude lead (VM). We limited this report to the analyses of lead VM. The VM lead was defined as the square root of the sum of the squares of the orthogonal X, Y, and Z signals which were computed by applying the Guldenring transformation matrix to the 12-lead signals [15]. It was expected that measurements were done on median/representative beats.

Assessment of consistency of JTpc measures across algorithms

Because the study that generated the data was designed as a thorough QT study, we opted to present the results between the various technologies in terms of double-delta changes caused by each drug. We report the double-delta JTpc values (ΔΔJTpc) which measure the difference between time-matched drug-placebo measurements computed after baseline adjustment. Hence, ΔΔJTpc includes corrections for potential baseline and placebo effects. The double-delta correction is accepted as the most efficient means to estimate a treatment difference between groups [3,16]. We assessed the consistency of JTpc measures across algorithms using the following methods:

1) Method 1 tests whether each measurement technology shows a dose-dependent effect on the JTpc interval for dofetilide and quinidine, but not for the other drugs (ranolazine and verapamil), confirming that predominant hERG block (dofetilide, quinidine) prolongs JTpc and that balanced ion channel block (ranolazine, verapamil) does not prolong JTpc. R code implementing the linear mixed effects models used to perform this analysis and the plots reported in this study is available on GitHub [17].
Both of the following Method 2 and 3 (below) were based exclusively on ECG measurements at the time (Tmax) of maximum drug concentration (Cmax) for each subject and treatment arm of the study. The time of the measurements prior to dosing was used as the baseline time point.

2) Method 2 tests the stability of the RR, RJ, RTpc and JTpc interval measurements. It is expected that smaller variations at each time point indicate better measurement precision. Therefore, we report the width of the arm-specific 95% confidence interval at Tmax across all subjects for each ECG interval and technology.

3) Method 3 evaluates the dispersion of a measurement technique in reference to the average measurements across all techniques. It is a metric to measure the divergence ($Div_{x_m}$) of the RJ, RTpc, and JTpc measurements for each algorithm in reference to the average of all algorithms. $\bar{x}_{m,k}$ denotes the population-based average value of an interval $x$ (RJ, RTpc, or JTpc) using method $m$ (AMPS, Philips, Mortara, FDA) for a study arm $k$ (dofetilide, ranolazine, verapamil and quinidine) at Tmax. $Div_{x_m}$ is computed as the average difference (across study arms) between $\bar{x}_{m,k}$ and the average value of $\bar{x}_{m,k}$ for all methods:

$$Div_{x_m} = \frac{1}{K} \sum_{k=1}^{K} \{ \bar{x}_{m,k} - \bar{y}_k \}$$

Eq. 1

where

$$\bar{y}_k = \frac{1}{M} \sum_{m=1}^{M} \bar{x}_{m,k}$$

Eq. 2
$K$ is the number of study arms, and $M$ is the number of measurement methods. We report the mean and standard deviation of this difference across arms to deliver a measure of divergence for each measurement technique.

The analysis and comparison of the results were computed using R 3.4.0 (R Foundation for Statistical Computing, Vienna, Austria). We also used a set of procedures from the FDA designed specifically to report the drug concentration effect on the JTpc intervals. It is important to note that the participants of the initiative were blinded to these criteria when assessing the ECGs. Criteria were disclosed to participants after receiving their submissions, i.e., after the ISCE 2017 J-Tpeak initiative session.

Results

The JTtp initiative was opened on October 17th, 2016, and deadline for submission of the results was set to February 28th, 2017. Five organizations contributed to the initiative: AMPS LLC., the U.S.A. Food and Drug Administration (FDA), Mortara Instrument Inc., and Philips Healthcare, as well as the University of Rochester (UR).

The University of Rochester ran the FDA’s open-source algorithm (available on GitHub) on the data. This algorithm provides only T-apex and T-offset locations from median beat using lead VM. The results obtained by UR were the same as for FDA-1, therefore we did not include them in the following sections. Hence, we report two versions of the FDA measurements: measurements submitted by FDA to the initiative (FDA-1), and measurements from the FDA study that were manually adjudicated and published in Johannesen et al. [3] (FDA-2) [17].
Completeness of the submitted measurements

A fair comparison of the technologies must include an evaluation of data completeness. Therefore, we report the missing values for all the participants focusing exclusively on the RR and JTpc interval measurements, and acknowledging that, if the RR value was missing, then JTpc would be missing as well. As an important note, there were 9 recordings in the dataset that did not include a median beat. The 10-second ECGs of these recordings with no median beats had either noise or a missing lead which made it impossible/inappropriate to derive the median beat and/or compute the VM lead. Therefore, participants who based their measurements on the existing median beat (rather than computing their own median beat) had at least 0.17% of the data missing (n=9).

Mortara processed and reported results for all files. AMPS was missing 0.17% of RR values (due to 9 missing median beats from the dataset) and 0.21% of JTpc intervals (9 missing, plus 2 additional files not analyzed). Philips was missing 0.06% of the results for both JTpc and RR. The FDA-1 and FDA-2 methods were missing 0.17% of values for both RR and JTpc. All technologies reported measurements for >99.75% of the overall dataset.

Drug-induced changes in JTpc across study arms

Figure 1 provides the mean JTpc values and upper bound of the 95% confidence intervals across the subject cohort for all study arms and for the five measurement methods. The mean JTpc was computed by averaging all values in a subject (excluding baseline) and then averaging these values across subjects.

The range of variation between the methods is below 10 msec for all arms of the study, except for the quinidine arm where the spread of mean values across the tested technologies was close
to 20 ms. The largest difference between drug-induced JTpc changes was found in the quinidine arm between the Mortara and FDA-2 technologies: 257.6±2.1 vs. 239.4±2.3 msec (mean±SD, p<0.0001), respectively.

Method 1: Drug-induced double-delta changes of the JTpc intervals and Drug concentration-dependent changes the JTpc intervals

The results describing the double-delta changes for the four arms of the study at Tmax are shown in Figure 2. First, all measurement technologies consistently discriminate between predominant hERG blockers (dofetilide/quinidine) versus multi-channel blockers (verapamil/ranolazine). Specifically, dofetilide and quinidine are strongly associated with JTpc prolongation with a lower 95% confidence interval bound above 10 msec for quinidine and above 25 msec for dofetilide. Second, the drugs with so-called balanced multichannel blockade are associated with small effect on the JTpc interval with a higher 95% confidence bound below or equal to 10 msec for all technologies. This figure also highlights the level of stability of the drug-induced changes corrected for baseline and placebo in each arm of the study. It reveals that the predominantly hERG blocking drugs are associated with larger variability.

The dependency between drug concentration and interval prolongation is one of the primary outcomes of any drug evaluation study. In this work, we applied the exposure-response technique used by the FDA to study this dependency in all measurement techniques [3]. The results are summarized in Figure 3. Importantly, we did not identify visual differences between FDA-1 and FDA-2, and we therefore opted to present FDA-1 results only. The graphs in Figure 3 show good consistency in the exposure-response relationship of the drugs between all
measurement techniques. However, one can highlight a set of subtle differences. The variability around the linear relationship is varying the most between technologies for quinidine. The mean predicted value of ∆ΔJTpc at quinidine’s Cmax ranges from 25 (FDA) to 45 msec (Mortara). The widest confidence interval at this same quinidine concentration value was obtained for the FDA measurements, and the smallest for the Mortara measurements. Finally, the slope of the ranolazine concentration dependency is slightly higher in Mortara and AMPS measurements than in the rest of the evaluated technologies.

Method 2: Stability of the RR, JTpc, RTPc and RJ intervals

Table 1 reports the width of the 95% confidence interval (CI) for the RR, JTpc, RJ, and RTpc interval measurements at baseline and Tmax of the quinidine arm. We limited our report to these two groups of measurements because, in general, baseline had the highest stability while quinidine ECG measures at Tmax had the lowest stability. The RJ intervals are the most stable, with a CI range below 2 msec, while JTpc and RTpc were similar but less stable compared to RJ. Hence, the J point had only a small contribution to the variability of the JTpc interval; rather, the identification of the T-apex is the main source of variability for all measurement methods.

Method 3: Divergence of the JTpc, RTPc and RJ intervals across measurement methods

Table 2 reports the divergence of each method from the mean estimate for all measurement methods, for all three intervals (JTpc, RTpc and RJ). The most divergence was found for the FDA results (4.1±0.5 msec) for the RJ interval, and Mortara for both the RTpc and
JTPc intervals (2.5±2.5 and 5.8±3.1 msec, respectively). However, these divergences are on average below 6 msec, which reveals consistency across methods.

**When the JTPc measurement methods do not match**

The largest variability of JTPc measurements was found in ECGs recorded in subjects exposed to either dofetilide or quinidine. These drugs are known to affect the morphology of the T-wave profoundly, and a review of the median beats under these drugs, especially for quinidine, shows that the T-waves in the VM lead are either flat, have a plateau phase, or have multiple bumps (notched). Examples of such T-waves are displayed in Figure 4 for two ECG tracings from subjects exposed to quinidine.

**Discussion**

The JTP initiative was designed in order to assess whether various existing automatic ECG delineation technologies, from the public domain or proprietary, could confirm the results published by the FDA on the propensity of this interval to respond differently to single versus multi-channel blocking drugs and specifically for predominant hERG blockers (dofetilide and quinidine) versus drugs with balanced hERG and late sodium and/or calcium blockade (verapamil and ranolazine). The results of this study demonstrate that, regardless of the technology adopted for the detection of the T-apex, the JTPc interval consistently differentiates between these two classes of drugs.

However, the analysis of the stability of the JTPc across measurement technologies has highlighted an increased variability of measurements for the drugs associated with strong hERG
blockage, namely, dofetilide and quinidine. As shown in Figure 4, ECG tracings with abnormal T-wave morphology (notched, biphasic, or plateaued T-wave) led to strong discordance between algorithm outputs. This lack of consistency in the reported location of the T-apex in abnormal T-wave configurations reflects the lack of common definition of the location of the T-apex in such configurations. As shown in Figure 2, AMPS and Mortara measurements in the quinidine arm are higher than for the other participants. Visual review of the tracings showed that these two technologies primarily selected the second peak of the notched T-wave as shown in Panel B of Figure 4. The need to measure the JTpc interval in such an abnormal T-wave may be questionable, since such drug-induced T-wave configurations would be considered suspicious. Therefore, ensuring measurement consistency could be addressed by establishing a specific set of rules for algorithms to handle this type of signal morphology.

The FDA scientists designed and publicly released an automatic algorithm for T-apex and T-end detection (only) [4]. Therefore, it is worth noting that the algorithm was not evaluated for detection of the J point, which is an input data for the FDA JTp algorithm. In this assessment, we used the J point or end of QRS included in the reference annotation set (FDA-2), although the J point could be detected using other procedures.

Finally, it is worth noting that this type of exercise has been conducted in the past, enabling a comparison of technologies from various manufacturers [18,19]. These exercises are important because they highlight the quality of the technologies currently available to clinicians and researchers, and in some cases provide an incentive for improving the current technologies by highlighting their weaknesses and limitations. In this work, the differences described across the various technologies are very small despite the lack of a gold standard around the T-apex
definition. Moreover, differences between double-delta changes at Tmax across all methodologies are on average below 6 ms.

**Limitations**

We propose a scientific endeavor to evaluate multiple technologies for automatic measurement of the JTpeak interval. The significance of this effort is related to the CiPA initiative, which integrates a clinical assessment of the effect of tested drugs on the JTpeak interval. This interval is not clinically used, and therefore its evaluation has never been reported. Importantly, there are many factors independent from the T-apex and J point detection which play a significant role in automatic JTpeak measurements. Amongst these factors are lead selection, beat selection, and RR interval measurement for heart rate correction. In the initiative, we purposely did not prescribe any method to the participants for processing these components so participants could use their existing methods. However, we did require each group to deliver at least the JTpeak measurement for a representative beat on the VM lead for each ECG recording. One should note that the computation of the VM lead, the representative beat, and the RR interval (number of beats selected for averaging RR, or inclusion hysteresis correction) may have influenced the measurements. The impact of these parameters is not delivered in our report.

**Conclusions**

We conducted an evaluation of a set of technologies designed to measure the heart rate-corrected JTp interval from standard 12-lead ECGs. The outcome of the study revealed that the technologies submitted to this initiative deliver similar measurements of drug-specific effects on
the JTpc interval. Therefore, the utility of JTpc prolongation as a marker of strong hERG blocking drugs vs. drugs with balanced ion channel effects is technology-independent. Discrepancies between measurement methods which give rise to concern were observed only during strong drug-induced changes of T-wave morphology.

**Disclaimer**

This article reflects the views of the authors and should not be construed as representing FDA’s views or policies.
Figure 1: Mean and 95% confidence interval (upper bound) of the JTpc interval measurements for all measurement technologies investigated in the initiative (see text for more details).
Figure 2: Double-delta changes of the JTpc (DDJTpc) intervals induced by dofetilide, quinidine, ranolazine, and verapamil.
Figure 3: JTpC interval drug concentration dependency for FDA-1. Mortara, Philips, and AMPS technologies. FDA-2 results were not included because they were visually not different from the FDA-1 results.
Figure 4: Two examples of median beats extracted from the quinidine arm where vertical lines indicate the location of the J point and T-apex across the submitted technologies. These examples illustrate the behavior of the various technologies for abnormal T-wave morphology and specifically a T-wave with a plateau phase (Panel A), and a flat, double bumped T-wave (Panel B).
# TABLES

## Table 1: 95% confidence interval for RR, JTpc, RJ and RTpc intervals at Tmax using Method 2

<table>
<thead>
<tr>
<th>Method 2</th>
<th>95% CI of RR (msec.)</th>
<th>95% CI of JTpc (msec.)</th>
<th>95% CI of RJ (msec)</th>
<th>95% CI of RTpc (msec)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>Q</td>
<td>B</td>
<td>Q</td>
</tr>
<tr>
<td>AMPS</td>
<td>28.5</td>
<td>15.6</td>
<td>4.2</td>
<td>4.5</td>
</tr>
<tr>
<td>Mortara</td>
<td>28.5</td>
<td>15.6</td>
<td>4.2</td>
<td>4.1</td>
</tr>
<tr>
<td>Philips</td>
<td>28.5</td>
<td>15.6</td>
<td>4.5</td>
<td>4.5</td>
</tr>
<tr>
<td>FDA-1</td>
<td>28.4</td>
<td>15.6</td>
<td>4.4</td>
<td>4.7</td>
</tr>
<tr>
<td>FDA-2</td>
<td>28.5</td>
<td>15.6</td>
<td>4.5</td>
<td>4.4</td>
</tr>
</tbody>
</table>

Table 1: 95% confidence interval for RR, JTpc, RJ and RTpc intervals at Tmax using Method 2: B: baseline, and Q: quinidine arm.

## Table 2: Summary of results for Method 3 (measurement divergence vs. mean)

<table>
<thead>
<tr>
<th>Study arms</th>
<th>Divergence from mean (msec.)</th>
<th>RJ</th>
<th>RTpc</th>
<th>JTpc</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMPS</td>
<td>-1.3±0.4</td>
<td>-0.3±0.6</td>
<td>2.9±1.0</td>
<td></td>
</tr>
<tr>
<td>Mortara</td>
<td>-0.3±0.3</td>
<td>2.5±2.5</td>
<td>5.8±3.1</td>
<td></td>
</tr>
<tr>
<td>Philips</td>
<td>-2.5±0.3</td>
<td>-0.8±1.2</td>
<td>3.7±0.9</td>
<td></td>
</tr>
<tr>
<td>FDA-1</td>
<td>4.1±0.5</td>
<td>-1.4±2.0</td>
<td>-3.7±1.3</td>
<td></td>
</tr>
<tr>
<td>FDA-2</td>
<td>4.1±0.5</td>
<td>-2.3±1.2</td>
<td>-5.1±0.8</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Summary of results for Method 3 (measurement divergence vs. mean).
Reference List


