Detection of QT prolongation using a novel electrocardiographic analysis algorithm applying intelligent automation: Prospective blinded evaluation using the Cardiac Safety Research Consortium electrocardiographic database

Cynthia L. Green, PhD,[a] Paul Kligfield, MD,[b] Samuel George, MD,[c] Ihor Gussak, MD, PhD,[c] Branislav Vajdic, PhD,[c] Philip Sager, MD,[d] and Mitchell W. Krucoff, MD[a] *Durham, NC; New York, NY; Princeton, NJ; and Palo Alto, CA*

Background The Cardiac Safety Research Consortium (CSRC) provides both "learning" and blinded "testing" digital electrocardiographic (ECG) data sets from thorough QT (TQT) studies annotated for submission to the US Food and Drug Administration (FDA) to developers of ECG analysis technologies. This article reports the first results from a blinded testing data set that examines developer reanalysis of original sponsor-reported core laboratory data.

Methods A total of 11,925 anonymized ECGs including both moxifloxacin and placebo arms of a parallel-group TQT in 181 subjects were blindly analyzed using a novel ECG analysis algorithm applying intelligent automation. Developer-measured ECG intervals were submitted to CSRC for unblinding, temporal reconstruction of the TQT exposures, and statistical comparison to core laboratory findings previously submitted to FDA by the pharmaceutical sponsor. Primary comparisons included baseline-adjusted interval measurements, baseline- and placebo-adjusted moxifloxacin QTcF changes (ddQTcF), and associated variability measures.

Results Developer and sponsor-reported baseline-adjusted data were similar with average differences <1 ms for all intervals. Both developer- and sponsor-reported data demonstrated assay sensitivity with similar ddQTcF changes. Average within-subject SD for triplicate QTcF measurements was significantly lower for developer- than sponsor-reported data (5.4 and 7.2 ms, respectively; $P < .001$).

Conclusion The virtually automated ECG algorithm used for this analysis produced similar yet less variable TQT results compared with the sponsor-reported study, without the use of a manual core laboratory. These findings indicate that CSRC ECG data sets can be useful for evaluating novel methods and algorithms for determining drug-induced QT/QTc prolongation. Although the results should not constitute endorsement of specific algorithms by either CSRC or FDA, the value of a public domain digital ECG warehouse to provide prospective, blinded comparisons of ECG technologies applied for QT/QTc measurement is illustrated. (Am Heart J 2012;163:365-71.)

The QT interval as recorded by a surface electrocardiogram (ECG) is a measure of the time from the onset of ventricular depolarization to the end of ventricular repolarization, with QT prolongation generally interpreted as a surrogate marker for possible delayed myocardial repolarization. Drug-induced QT prolongation is associated with an increased risk of potentially fatal ventricular arrhythmias such as torsades de pointes. To minimize the incidence of drug-induced torsades de pointes, the US Food and Drug Administration (FDA), through the International Committee on Harmonization E14 guidance document, recommends that all novel drugs with systemic exposure submitted for marketing approval should be first evaluated to assess cardiac risk in target patient populations.[1] The thorough QT (TQT) study is a rigorous evaluation of the effect of a novel drug

From the [a]Duke University Medical Center and Duke Clinical Research Institute, Durham, NC, [b]Division of Cardiology, Department of Medicine, Weill Cornell Medical College, New York, NY, [c]NewCardio, Inc, Princeton, NJ, and [d]Pharmaceutical/Device Consultant, Palo Alto, CA.

Submitted November 10, 2011; accepted November 10, 2011.

Reprint requests: Cynthia L. Green, PhD, Duke University Medical Center, Durham, NC 27705. E-mail: cindy.green@duke.edu

0002-8703/$ - see front matter

doi:10.1016/j.ahj.2011.11.009

on cardiac repolarization as measured by the QT interval and heart rate–corrected QT interval (QTc) duration, with mean QTc prolongation of 5 ms as the threshold for regulatory concern.[2,3]

A TQT study requires precise interval duration measurements (IDMs) on as many as 30,000 manually or semiautomated measured ECG measurements per study, each of which is commonly overread for confirmation by an experienced cardiologist.[4-8] Availability of a fully or virtually automated algorithm to measure IDMs accurately may accelerate timelines and reduce variability of ECG measurements with consequent reduction in the costs of ECG analyses for TQT studies, making cardiac safety considerations in new drug development less cumbersome. To use data generated by any proposed ECG algorithm, a quantitative and unbiased method to evaluate the reliability of such new algorithms compared with contemporary core laboratory processes is needed. Accordingly, any ECG measurement method claiming to be accurate and reliable must properly detect the positive control (moxifloxacin) effect in a TQT study.

In this document, we present results from the first use of the Cardiac Safety Research Consortium (CSRC) ECG warehouse QT/QTc testing data set, comparing IDMs obtained from a commercially available ECG algorithm applying intelligent automation (QTinno, NewCardio, Inc, Santa Clara, CA; referred to herein as developer reanalysis) blinded QT/QTc analysis to results of a contemporary core laboratory as originally submitted to the FDA on behalf of a pharmaceutical sponsor (sponsor reported). Only the moxifloxacin and placebo arms of the parallel TQT study were released by the sponsor for use by the CSRC.

Background of the CSRC ECG warehouse data

The CSRC is a public-private partnership formed within the Critical Path Initiative to advance scientific knowledge on cardiac safety issues pertinent to new drug and device development.[9] In collaboration with the FDA and the Duke Clinical Research Institute, the CSRC maintains a centralized ECG repository (the ECG warehouse) for digital ECG recordings from TQT and related studies submitted to the FDA during drug development and postmarket surveillance by industry sponsors and released to CSRC. Pharmaceutical company sponsors, who retain ownership of the recordings, have made a number of ECG warehouse data sets available to investigators, equipment manufacturers, and algorithm developers, for CSRC-approved research and development studies.

Most ECG warehouse data are waveforms collected from TQT studies performed in healthy volunteers during early phases of new drug evaluation. Cardiac Safety Research Consortium ECG data sets are composed of anonymized digital XML waveforms and key descriptive data from baseline, placebo, and moxifloxacin periods of TQT studies. The CSRC ECG warehouse supports the public domain use of these data sets for CSRC-approved research. Such research includes the testing ECG algorithms applicable to FDA cardiac safety interests such as detection of QT changes associated with moxifloxacin compared with placebo.

To meet the needs for reliable and unbiased assessment of novel ECG analysis algorithms, the CSRC has partitioned sets of ECG waveforms released with descriptor data such as treatment assignment ("unblended" or "training" data) and sets of ECG waveforms without such descriptors ("blinded" or "testing" data) for public-domain research use. This partitioning allows developers to refine algorithms using CSRC training data, while reserving testing data for more rigorous, reportable performance testing. The testing ECG database (XML waveforms only) is provided to algorithm developers under strict predefined conditions to assure unbiased, accurate, and reliable assessment of the algorithm. These conditions include full blinding with respect to the administration of pharmaceutical agents, an independent analysis of the algorithm developer's IDMs by a CSRC-affiliated statistician, and an agreement by the developer that the results will be made public regardless of the outcome.[10]

Methods

Study design

The testing data set used for validating algorithm performance was a positive- and placebo- controlled parallel design TQT study (referred to as CSRC Study 5). Data included 11,925 anonymized digital ECG waveforms from 181 subjects, 90 of which were randomized to the positive-control arm and 91 to the placebo arm. Positive-control subjects received 400 mg oral moxifloxacin, which has been shown to prolong the QTc interval by approximately 10 ms while being adequately safe to use a single dose in studies of normal volunteers.[11]

All subjects had three 12-lead digital ECGs obtained within 1 to 3 minutes (providing 3 ECGs for each time point) on day 1 (baseline) and then on day 1 (treatment) at the following time points relative to placebo or moxifloxacin administration: predose (15 min) and 1, 2, 3, 4, 5, 6, 8, 10, 12, and 24 hours postdose. The time-matched ECGs obtained on day 1 provided the baseline values. A total of 11,925 ECGs were originally submitted by the sponsor to the ECG warehouse and then subsequently reanalyzed by the developer. Of these ECGs, 241 (2.1%) were incomplete (2.5 seconds of recorded ECG data per lead compared with 10 seconds expected) and were not evaluable by the ECG algorithm (QTinno, NewCardio, Inc, Santa Clara, CA); 11 ECG files were rejected by the intelligent automation of the ECG algorithm because of high noise content; and no data were reported by the sponsor for one XML file. Thus, a total of 11,672 ECGs were available for the comparative analysis.

The virtually automated ECG reanalysis of the XML waveforms was made using QTinno, an ECG analysis algorithm that

Table I. Bland-Altman data for baseline-adjusted cardiac intervals, comparing developer reanalysis to sponsor-reported mean values at each time point

	Sponsor reported		Developer reanalysis		Difference	
Variable	Mean (ms)	SD (ms)	Mean (ms)	SD (ms)	Mean (ms)	SD (ms)
dQT	3.12	15.93	2.67	16.02	–0.44	8.28
dRR	2.60	88.58	1.90	86.62	–0.70	27.85
dQTcB	2.79	13.62	2.47	11.77	–0.32	10.39
dQTcF	2.89	11.29	2.54	9.90	–0.35	9.25
dPR	0.09	7.71	0.11	7.02	0.02	5.72
dQRS	–0.28	5.52	0.03	3.81	0.31	6.49

Average change from baseline and SD, as well as average difference and SD between analysis methods are presented for each interval in milliseconds.

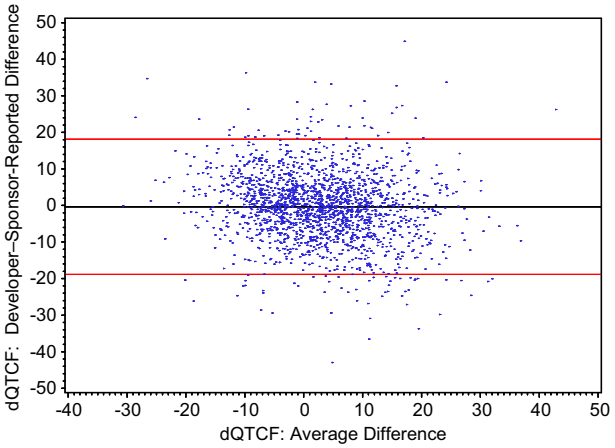
applies intelligent automation, which has been previously described.[12] Briefly, the QTinno algorithm uses the full 12-lead ECG input to generate a vector magnitude lead and then uses the vector magnitude lead for all IDMs. Fiducial point placements are made primarily by curve fitting to a third-order polynomial, an approach that was chosen because it is relatively noise resistant and does not require location of an isoelectric baseline. The intelligent automation of the QTinno algorithm separates those ECG readings that do not require adjudication by a human from those that are suggested for adjudication. Sixty-seven (0.57%) of the available 11,672 ECG files were flagged by the algorithm for expert cardiologist visual adjudication because of high noise content such as cardiac intervals outside the reference range, marked sinus arrhythmia, and other outlier factors. Of the 67 flagged files, 24 required adjustments by the cardiologist reader, 0.21% of the overall number of files, whereas 43 of the flagged files (0.37%) were accepted by the reader without change.

Statistical methods

The time-matched change from baseline of the following continuous ECG interval parameters was summarized at each time point on treatment day 1 during both the placebo and moxifloxacin treatment periods using central tendency analyses: QT, QTcB (Bazett's heart rate QT correction), QTcF (Fridericia's heart rate QT correction), RR, PR, and QRS. The actual QT values at each time point between sponsor-reported core laboratory and developer reanalysis measurements were not directly compared because no absolute independent gold standard exists for individual waveform measurements. Rather, using analytical comparison standards set by CSRC, differences in test performance were assessed by the ability of each method to identify a moxifloxacin effect compared with placebo by intrinsic variability within the methods and by the ability of the methods to identify a drug effect in randomly selected simulations of smaller subsets within the overall population, as described below.

Three replicate measurements were available for each interval at each time point and were averaged for each subject, denoted as the mean interval value. For each subject and time point, the change from baseline value was derived using the mean interval value on day 1 subtracted from the time-matched mean value on

Figure 1



Bland-Altman plot for baseline-adjusted QTcF values (dQTcF) for sponsor-reported core laboratory and developer reanalysis data. Horizontal lines represent the average difference between measurement methods (–0.35 ms) and ± 2 SDs from the mean difference (–18.5 ms, 18.5 ms).

Table II. Average within-subject SDs for QTcF based on SD of replicates at each time point in milliseconds, averaged over the entire time course for each subject

Group	Method	n	Mean SD (ms)	LCL (ms)	UCL (ms)	*P*
All	Sponsor reported	181	7.17	6.91	7.43	
	Developer	181	5.43	5.19	5.67	
	Difference	181	1.74	1.53	1.95	<.0001
Moxifloxacin	Sponsor reported	90	7.14	6.79	7.48	
	Developer	90	5.41	5.05	5.76	
	Difference	90	1.73	1.42	2.04	<.0001
Placebo	Sponsor reported	91	7.20	6.81	7.59	
	Developer	91	5.45	5.11	5.79	
	Difference	91	1.75	1.45	2.05	<.0001

Lower (LCL) and upper (UCL) 90% CI limits for the mean SD are given along corresponding *P* value for testing the null hypothesis of no difference between analysis methods.

treatment day 1. All data analyses were generated using SAS software, Version 9 (SAS Institute Inc, Cary, NC).

Bland-Altman plots of developer reanalysis and sponsor-reported core laboratory changes from baseline were reviewed for each interval (dQT, dQTcB, dQTcF, dRR, dPR, and dQRS) to evaluate unexpected differences or patterns, whereas time trend plots of the 2-sided 90% confidence intervals (CIs) of the mean value at each time point were used to review similarity/overlap of each analysis method. The average difference between developer and sponsor-reported core laboratory change from baseline values, as well as SD of the differences, was computed to observe expected CIs approximately 0 ms.

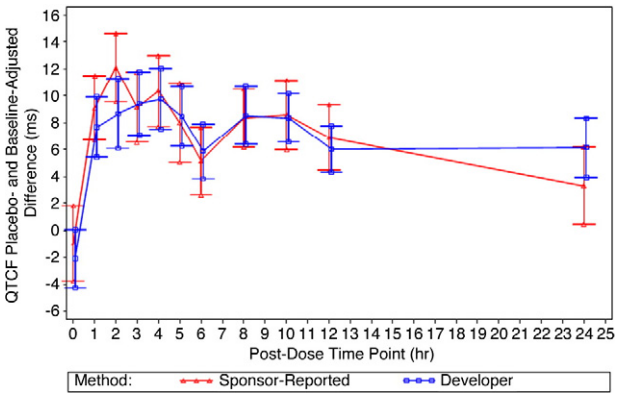
The within-subject variability of each interval was compared using the SD of the triplicate measurements at each time point

Table III. Baseline- and placebo-adjusted QTcF (ddQTcF) at prespecified time points (TP) for sponsor-reported and developer reanalysis average differences

	Sponsor reported					Developer reanalysis				
TP (h)	Mean (ms)	SE (ms)	90% LCL* (ms)	90% LCL (ms)	90% UCL (ms)	Mean (ms)	SE (ms)	90% LCL* (ms)	90% LCL (ms)	90% UCL (ms)
0	−0.98	1.68		−3.75	1.79	−2.08	1.30		−4.23	0.07
1	9.10	1.44	5.85	6.72	11.48	7.67	1.35	4.63	5.44	9.90
2	12.08	1.52	8.63	9.56	14.60	8.67	1.57	5.12	6.08	11.27
3	9.13	1.57	5.58	6.53	11.74	9.41	1.41	6.21	7.07	11.74
4	10.34	1.58	6.76	7.72	12.96	9.75	1.36	6.67	7.50	12.00
5	7.99	1.77		5.06	10.92	8.45	1.34		6.24	10.66
6	5.14	1.54		2.60	7.68	5.86	1.22		3.84	7.87
8	8.34	1.30		6.18	10.49	8.52	1.29		6.38	10.66
10	8.54	1.55		5.97	11.12	8.35	1.08		6.56	10.14
12	6.90	1.47		4.47	9.34	6.02	1.03		4.32	7.73
24	3.33	1.72		0.49	6.18	6.14	1.33		3.95	8.34
Average	7.27	1.56				6.98	1.30			

Values are given in milliseconds for the mean ddQTcF, standard error (SE), and 90% upper (UCL) and lower (LCL) CIs.
*Bonferroni multiple testing adjustment by 4 time points.

Figure 2



Raw interval results for baseline- and placebo-adjusted QTcF values (ddQTcF) for sponsor-reported and developer reanalysis data, presented as mean values with 2-sided 90% CIs for each postdose time point and analysis method.

averaged across the whole time course for each subject. A paired *t* test was used to test the null hypothesis that the mean developer SD was the same as the mean sponsor-reported core laboratory SD for each interval. $P \leq .05$ was considered statistically significant.

Detection of the expected baseline- and placebo-adjusted moxifloxacin prolongation of the QTcF interval (ddQTcF) was evaluated using time trend plots of the raw mean differences and 90% CIs at each local time point for each analysis method. To determine assay sensitivity while maintaining the overall 2-sided significance level ($\alpha = .10$), Bonferroni multiple testing adjustment of the first 4 postdose time points ($\alpha = .025$) was used. If one of the adjusted lower bounds of the 90% CIs during the first 4 hours postdose was greater than 5 ms, it would be concluded that the analysis method is sensitive to detecting small drug effects on cardiac repolarization.

Table IV. Covariance estimates in milliseconds based on the mixed ANCOVA model for sponsor-reported and developer reanalysis data

	Between-subject SD (ms)		Within-subject SD (ms)	
Interval	Sponsor reported	Developer reanalysis	Sponsor reported	Developer reanalysis
dQT	17.24	17.56	9.52	8.99
dRR	87.60	88.79	54.31	51.94
dQTcB	14.62	13.91	8.94	7.63
dQTcF	12.94	12.07	7.25	6.07
dPR	11.65	10.79	5.27	5.01
dQRS	5.07	5.05	3.71	2.54

The between- and within-subject SDs are given for each interval and measurement method.

The variability in the change from baseline for QTcF (dQTcF), as described by the between-subject and within-subject SDs, was estimated by fitting a mixed analysis of covariance (ANCOVA) model for each analysis method.[13] The normal mixed model for each dQTcF-dependent variable is given by

$$y = X\alpha + Z\beta + \varepsilon,$$

where $\boldsymbol{X}$ is the matrix of fixed effects and $\boldsymbol{Z}$ is the matrix of random effects. Subject is the random effect, whereas baseline QTcF, time point, treatment group, and time by treatment interaction are the fixed effects. We assumed a compound symmetry variance structure with equal variances across both treatment groups, because this is the most common variance structure for these data type, and tested this assumption using likelihood ratio tests comparing models using other possible covariance structures. Although not presented in this article, the 2-sided 90% CIs for ddQTcF based on the mixed ANCOVA model

Table V. Theoretical statistical power calculations, using 5,000 random selections of smaller sample sizes than the original study population

Sample size	Method	Average peak moxifloxacin effect (ms)	Average time of peak moxifloxacin effect (h)	Between-subject average SD (ms)	Within-subject average SD (ms)	Power
10	Sponsor reported	15.49	4.65	12.81	7.23	0.168
10	Developer	14.52	4.87	11.88	6.05	0.165
20	Sponsor reported	14.33	3.74	12.90	7.24	0.301
20	Developer	13.54	4.15	11.98	6.06	0.296
30	Sponsor reported	13.86	3.41	12.91	7.24	0.418
30	Developer	13.12	3.86	12.01	6.06	0.413
40	Sponsor reported	13.62	3.25	12.93	7.25	0.514
40	Developer	12.92	3.77	12.04	6.07	0.524
50	Sponsor reported	13.46	3.15	12.92	7.25	0.602
50	Developer	12.78	3.75	12.05	6.07	0.616
60	Sponsor reported	13.34	3.08	12.92	7.25	0.675
60	Developer	12.68	3.71	12.04	6.07	0.689
70	Sponsor reported	13.24	2.98	12.93	7.24	0.745
70	Developer	12.60	3.73	12.05	6.07	0.761
80	Sponsor reported	13.18	2.95	12.93	7.25	0.805
80	Developer	12.56	3.74	12.06	6.07	0.822

Average peak moxifloxacin effect in milliseconds, average hour of peak effect, between- and within-subject average SDs, and power are given for each sample size.

were constructed using the residual error to confirm results similar to raw mean results.

The theoretical statistical power was estimated by bootstrap simulations on randomly selected subject groups ranging in size from 10 to 50 to observe the average maximum effect, the average within- and between-subject variances, and the number of times that assay sensitivity was detected (power) in smaller sample sizes. Assay sensitivity was determined using the Bonferroni-adjusted lower 90% CI of first 4 postdose time points. The power curves show the minimum theoretical number of subjects that would be required to achieve at least 80% power for each analysis method. Statistical power was also computed for each method considering only the between-subject SDs of the original study design and assuming a 2-sided significance level of .10 (adjusted for 4 time points) with a true mean difference of 12 ms.[14]

Results

Table I presents summary data from Bland-Altman comparisons of sponsor-reported and developer reanalysis for baseline-adjusted changes in each interval including dQT, dRR, dQTcB, dQTcF, dPR, and dQRS intervals. The measurement methods were similar with regard to change from baseline, with an absolute mean difference <1 ms for each interval. Figure 1 displays the Bland-Altman plot for dQTcF and shows expected scatter around the mean difference (−0.35 ms) with most points lying within 2 SDs of the mean.

The average within-subject SD across all subjects and time points and by treatment group is presented in Table II for QTcF. A paired t test rejected the null hypothesis that the sponsor-reported core laboratory and developer reanalysis SDs are the same for QTcF ($P <$ .0001). Similar comparisons were done for each IDM and yielded similar results.

Hourly changes in baseline- and placebo-adjusted QTcF intervals (ddQTcF) in sponsor-reported core laboratory and developer reanalysis data are presented in Table III and Figure 2. Average values for mean ddQTcF were similar (7.27 and 6.98 ms, respectively), as were hourly point estimates throughout the time course of the study. Seven of the 11 hourly point estimates for mean ddQTcF were within 1 ms of each other, and CIs overlapped at all time points. In both methods, at least 3 Bonferroni-adjusted lower 90% CI limits excluded 5 ms demonstrating assay sensitivity. Average standard errors were 16.7% larger for core laboratory data than for developer data (1.56 and 1.30 ms, respectively).

Table IV shows the relative variability of the sponsor-reported core laboratory and developer reanalysis data based on the ANCOVA model for baseline-adjusted change in each interval. The between-subject SDs for dQTcF were similar (12.94 ms for core laboratory and 12.07 ms for developer data), whereas within-subject SDs were approximately 17% larger for core laboratory relative to developer data (7.25 and 6.07 ms, respectively). Although not shown, similar differences were observed when the moxifloxacin and placebo groups were evaluated separately. For other intervals, the between-subject variances were similar, whereas the within-subject variances were modestly lower for developer than for core laboratory data.

Sponsor-reported core laboratory and developer reanalysis data were also compared for theoretical statistical power using bootstrap simulations in Table V using sample sizes varying from 10 to 80. In these simulations, both the average between- and within-subject SDs were modestly lower for developer data than for core laboratory data; however, the average peak moxifloxacin effect was higher in the core laboratory data. Nonetheless,

average between-subject SD in the developer data translated into a 2.1% increase in conditional power for assay sensitivity. Considering the between-subject SDs of the original study design in Figure 3, 65 subjects per treatment group would be required based on the core laboratory analyses versus 57 per treatment group using the developer analytic method to detect assay sensitivity with at least 80% power in a similar parallel study design providing a 12% reduction in sample size.

Discussion

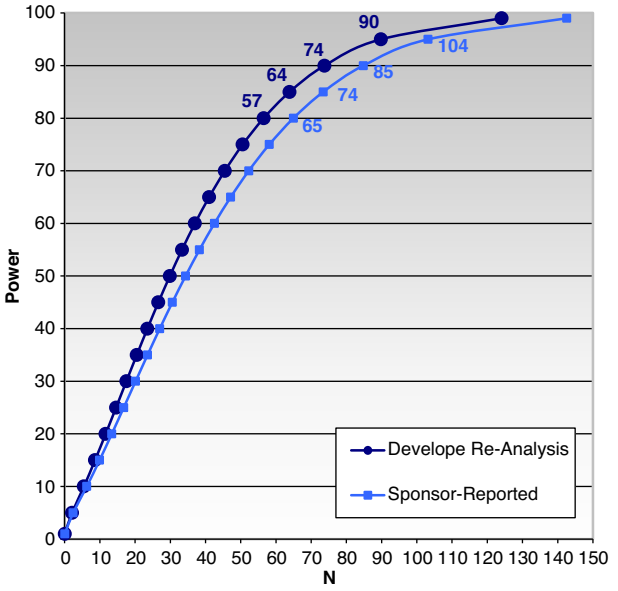
This study was undertaken by the developer using the CSRC digital ECG warehouse QT data sets to perform a blinded, prospective evaluation of a novel ECG measurement technology compared with core laboratory data submitted to the US FDA by the pharmaceutical sponsor. All developer interval measurements were made blinded to treatment or other descriptors including temporal sequence. All processes for matching developer interval measurements to clinical descriptors and statistical methods for comparing assay sensitivity and variability between methods were developed and executed by CSRC independently of the developer.

Results obtained with the virtually automated developer ECG algorithm in the present study were similar to the sponsor-reported core laboratory data, both by Bland-Altman analysis of baseline-adjusted change in QT, QTcB, QTcF, PR, and QRS intervals and for hourly changes in ddQTcF. These findings indicate the use of blinded testing data sets provided in the public domain through the agencies of the CSRC digital ECG warehouse. Such use includes the unbiased evaluation and comparison of new ECG technologies promoting better or more efficient quantification of QT safety concerns in early-phase drug development.

Moreover, hourly ddQTcF results from both measurement methods correlated well with average ddQTcF values recently published by Yan et al.[15] Using a mixed-model statistical approach, Yan et al reported that the average peak ddQTcF from 6 parallel-group studies was 11.3 ms (90% CI, 9.3-13.2 ms), occurring 4 hours after a single oral dose of moxifloxacin 400 mg. In comparison, using a mixed-model, approach the peak ddQTcF in our study was 12.79 ms (90% CI, 9.13-16.45 ms) occurring at 2 hours for sponsor-reported data, and 12.48 ms (90% CI, 9.15-15.81 ms) occurring at 4 hours for developer data.

Although mean values for measured cardiac intervals and ddQTcF were similar between sponsor-reported core laboratory and developer reanalysis data, the developer had significantly lower measurement variability than the corresponding core laboratory data. In early-phase clinical drug studies, reducing measurement variability is important for several reasons. For example, a drug that prolongs the QT interval is identified as having a QT liability if at any study time point, the 90% upper confidence limit of ddQTcF is ≥10 ms. This raises the possibility of a false-positive result if the drug does not, in fact, significantly prolong QT, but the CIs are wide because of measurement variability. Similarly, substantial measurement variability may result in a failure to detect the expected moxifloxacin effect (assay sensitivity) with at least 1 time point at which the 90% lower confidence limit excludes 5 ms.

Figure 3



Power and sample size estimates for determining assay sensitivity. Assuming 4 time points postdose evaluated with overall 1-sided significance level of .05 and true mean effect of 12 ms. Sponsor-reported and developer reanalysis data were presented using estimated variance from original study.

In multiple ascending dose and other early-phase trials, lower measurement variability may also improve pharmacokinetic/pharmacodynamic correlations and improve study design for TQT and subsequent studies by providing better estimates of a drug's likely QT effect in smaller sample sizes.

In bootstrap simulations, the lower between-subject variability observed with developer data did not translate into a significantly higher theoretical study power. Although the developer data had 7% lower average between-subject SDs (which would favor higher power in bootstrap simulations), the sponsor-reported core laboratory data had approximately a 5% higher peak ddQTcF value. Because theoretical study power represents the proportion of simulations where the 90% lower confidence limit of ddQTcF excluded 5 ms at least once during the first 4 hours postdose, the higher core laboratory peak ddQTcF value made this event more likely and largely offset the difference in between-subject

SDs. Sample size considerations of TQT studies are primarily driven by the baseline- and placebo-adjusted experimental drug comparisons. However, if we consider only the between-subject SDs of the original study design, approximately 12% fewer subjects per treatment group would be required to detect assay sensitivity based on the developer reanalysis data.

Summary

There is not a criterion standard for direct comparisons of QT IDMs, and hence, the CSRC does not report such comparison between core laboratory and new algorithms under testing. In this study, both core laboratory and developer algorithms identify similar time- and magnitude-related changes in baseline- and placebo-adjusted QT/QTc response to moxifloxacin. Thus, QTinno appears to be a reasonable and efficient alternative to the semiautomated methods used by core laboratories traditionally selected by pharmaceutical sponsors for TQT studies.

Interpretation of the implications of these findings may vary depending on the specifics of actual practical or regulatory applications. Thus, the methodologies used in this study may be useful for such considerations with these and other future ECG technology evaluations; however, the findings from these evaluations should not be construed as specific endorsement of the algorithm used in this study by either the CSRC or the FDA.

Acknowledgements

CLG was supported in part by a Clinical and Translational Sciences Award Grant to Duke University (UL1 RR024128).

Disclosures

Cynthia Green and Paul Kligfield have no conflicts of interest to report.

Samuel George, Ihor Gussak and Branislav Vajdic are all employees of NewCardio,Inc., and hold equity interests in NewCardio.

Philip Sager was once a consultant for NewCardio but no longer serves in this role.

Mitchell Krucoff is currently a consultant for NewCardio (less than $10K per year).

References

1. Food and Drug Administration. International Conference on Harmonisation; guidance on E14 clinical evaluation of QT/QTc interval prolongation and proarrhythmic potential for non- antiarrhythmic drugs; availability. Notice Fed Regist 2005;70:61134-5.
2. Shah RR. Drugs, QTc interval prolongation and final ICH E14 guideline: an important milestone with challenges ahead. Drug Saf 2005;28:1009-28.
3. Darpo B. The thorough QT/QTc study 4 years after the implementation of the ICH E14 guidance. Br J Pharmacol 2010; 159:49-57.
4. Malhotra BK, Glue P, Sweeney K, et al. Thorough QT study with recommended and supratherapeutic doses of tolterodine. Clin Pharmacol Ther 2007;81:377-85.
5. Serra DB, Affrime MB, Bedigian MP, et al. QT and QTc interval with standard and supratherapeutic doses of darifenacin, a muscarinic M3 selective receptor antagonist for the treatment of overactive bladder. J Clin Pharmacol 2005;45:1038-47.
6. Zhang L, Chappell J, Gonzales C, et al. QT effects of duloxetine at supratherapeutic doses: a placebo and positive controlled study. J Cardio Pharmacol 2007;49:146-53.
7. Hulhoven R, Rossillon D, Bridson WE, et al. Effect of levetiracetam on cardiac repolarization in healthy subjects: a single-dose, randomized, placebo- and active-controlled, four-way crossover study. Clin Ther 2008;30:260-70.
8. Morganroth J, Lepor H, Hill LA, et al. Effects of the selective α1A-adrenoceptor antagonist silodosin on ECGs of healthy men in a randomized, double-blind, placebo- and moxifloxacin-controlled study. Clin Pharmacol Ther 2010;87:609-13.
9. CSRC scientific rationale and mission statement. Avalable at: https://www.cardiac-safety.org/about-us/. Last accessed May 19, 2011.
10. Kligfield PW, Green CL, Mortara J, et al. The Cardiac Safety Research Consortium electrocardiogram warehouse: thorough QT database specifications and principles of use for algorithm development and testing. Am Heart J 2010;160:1023-8.
11. Bloomfield DM, Kost JT, Ghosh K, et al. The effect of moxifloxacin on QTc and implications for the design of thorough QT studies. Clin Pharmacol Ther 2008;84:475-80.
12. Sarapa N, Gussak I, Vajdic B, et al. Comparison of QTinno, a fully automated electrocardiographic analysis program, to semiautomated electrocardiographic analysis methods in a drug safety study in healthy subjects. J Electrocardiol 2009;42:358-66.
13. Verbeke G, Molenberghs G. Linear mixed models for longitudinal data, New York: Springer Science + Business Media, Inc.; 2000.
14. Zhang J. Testing for a positive control in a thorough QTc study. J Biopharm Stat 2008;18:517-28.
15. Yan YK, Zhang J, Ng MJ, et al. Statistical characteristics of moxifloxacin-induced QTc effect. J Biopharm Stat 2010;20: 497-507.