Results From the IQ-CSRC Prospective Study Support Replacement of the Thorough QT Study by QT Assessment in the Early Clinical Phase

B Darpo1,2*, C Benson3†, C Dota4*, G Ferber5, C Garnett6*, CL Green7, V Jarugula8†, L Johannesen9, J Keirns10†, K Kruďys11, J Liu11, C Ortemann-Renon12*, S Riley13, N Sarapa14†, B Smith2, RR Stoltz15, M Zhou2 and N Stockbridge16

The QT effects of five “QT-positive” and one negative drug were tested to evaluate whether exposure–response analysis can detect QT effects in a small study with healthy subjects. Each drug was given to nine subjects (six for placebo) in two dose levels; positive drugs were chosen to cause 10 to 12 ms and 15 to 20 ms QTcF prolongation. The slope of the concentration/\(\Delta QTc\) effect was significantly positive for ondansetron, quinine, dolasetron, moxifloxacin, and dofetilide. For the lower dose, an effect above 10 ms could not be excluded, i.e., the upper bound of the confidence interval for the predicted mean \(\Delta QTcF\) effect was above 10 ms. For the negative drug, levocetirizine, a \(\Delta QTcF\) effect above 10 ms was excluded at 6-fold the therapeutic dose. The study provides evidence that robust QT assessment in early-phase clinical studies can replace the thorough QT study.

Safety concerns arising from observations of QT prolongation and potentially lethal proarrhythmias caused by non-antiarrhythmic drugs during the 1990s1,2 led to the regulatory request to improve the characterization of potential ECG effects of new drugs3,4. In 2005, the ICH E14 clinical guidance for QT assessment5 was implemented, which mandated that all new drugs with systemic availability should undergo systematic evaluation of the potential to cause QT prolongation, typically in a so-called thorough QT (TQT) study in healthy subjects. The study is designed to exclude a QT effect above the threshold of concern, i.e., an effect exceeding 10 ms must be excluded at all post-dose timepoints.6,7 In case such a “threshold” effect cannot be excluded, the implications in terms of ECG monitoring in late-stage trials are substantial and if the drug is approved, the labeling will include appropriate recommendations.7 The TQT study is resource-intensive8 and if an alternative way of QT assessment could be incorporated into a routinely performed early-phase clinical pharmacology study, this would present not only a more efficient approach but also come with other advantages, such as improved understanding of any QT concerns early in clinical development. The first-in-human (FIH) single ascending dose (SAD) and multiple ascending dose (MAD) studies seem best suited for this purpose; often, achieved plasma levels substantially exceed therapeutic levels later observed in patients. Provided that an intense ECG assessment schedule coinciding with pharmacokinetic sampling is incorporated into the design, SAD and MAD studies represent an opportunity to generate ECG data with the same high level of confidence as the TQT study.9–11 However, because doses are distributed across several small cohorts with often fewer than 10 subjects on active drug and two on placebo,
the power to exclude small effects in a “by timepoint” analysis, as mandated for the TQT study, is low. In contrast, when an exposure–response (ER) analysis is utilized, all data across the often wide range of plasma concentrations of the drug are used in the same model, which substantially improves the precision of the estimated QTc effect. ER analysis is mentioned in the E14 Q&A document from March 2014 as “promising in terms of enhancing our confidence to characterize QTc prolongation” and the methodology has been used in the evaluation of TQT studies to characterize QTc effects. The analysis has been used to predict QT effects with doses and formulations not directly evaluated in the TQT study, predict effects in specific populations, and under certain conditions (e.g., drug interactions) that increase exposure to a drug due to intrinsic and extrinsic factors and clarify ambiguous results from the TQT study.

The objective of this study was to evaluate whether QT assessment performed in early-phase studies using an intense ECG schedule and ER analysis can detect a small QT effect with the same confidence as a thorough QT study. The study was undertaken as a collaborative research effort between the Clinical Pharmacology Leadership Group of the Consortium for Innovation and Quality in Pharmaceutical Development (IQ) and the Cardiac Safety Research Consortium (CSRC). Six drugs with a well-characterized QT effect, of which five have been shown to be positive, were evaluated in healthy volunteers. Drugs and doses were identified in collaboration with the US Food and Drug Administration (FDA) based on a shared understanding that if the study successfully detects the QT effect of the positive drugs, a similar approach (i.e., QT assessment in early-phase clinical studies) could potentially serve as an alternative to the TQT study.

Based on the results of this study and on the extensive experience that has been gained over recent years with ER analysis of QT data, we propose that a clinically relevant QT effect can be excluded with ER analysis applied to early-phase clinical data.
The targeted peak plasma levels were generally achieved on Day 1 and Day 2, with the exception of ondansetron, for which levels were lower than anticipated on Day 2 (Table 1, Table 2).

### Table 1: Study drugs and doses administered

<table>
<thead>
<tr>
<th>Drug</th>
<th>Day 1</th>
<th>Day 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ondansetron</td>
<td>52 mg p.o.</td>
<td>32 mg i.v. by 15-minute infusion</td>
</tr>
<tr>
<td>Quinine</td>
<td>648 mg p.o. Q8 hours (3 doses on Day 1 and one in the morning of Day 2)</td>
<td>648 mg q8h x 4</td>
</tr>
<tr>
<td>Dolasetron</td>
<td>100 mg p.o.</td>
<td>150 mg i.v. by 15-minute infusion</td>
</tr>
<tr>
<td>Moxifloxacin</td>
<td>400 mg p.o.</td>
<td>800 mg i.v. by 60-minute infusion</td>
</tr>
<tr>
<td>Dofetilide</td>
<td>0.125 mg p.o.</td>
<td>0.25 mg p.o.</td>
</tr>
<tr>
<td>Levocetirizine</td>
<td>5 mg p.o.</td>
<td>30 mg p.o.</td>
</tr>
</tbody>
</table>

i.v.: intravenous; p.o.: per oral.

Table 2 Largest mean ΔΔQTcF by timepoint and observed arithmetic mean C\textsubscript{max} across treatments and study days

<table>
<thead>
<tr>
<th>Drug</th>
<th>Day 1</th>
<th>Day 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ΔΔQTcF mean (90% CI) ms &amp; C\textsubscript{max} mean (SD) ng/mL &amp; T\textsubscript{max} mean (SD) hour</td>
<td>ΔΔQTcF mean (90% CI) ms &amp; C\textsubscript{max} mean (SD) ng/mL &amp; T\textsubscript{max} mean (SD) hour</td>
</tr>
<tr>
<td>Ondansetron</td>
<td>12.2 (7.1 to 17.4) &amp; 295 (94.6) &amp; 1.9 (1.1)</td>
<td>10.2 (4.0 to 16.5) &amp; 236 (56.7) &amp; 0.6 (0.2)</td>
</tr>
<tr>
<td>Quinine\textsuperscript{a}</td>
<td>13.3 (4.2 to 22.4) &amp; 3,819 (1,296) &amp; 2.4 (0.53)</td>
<td>22.1 (13.3 to 30.9) &amp; 5,827 (2,107) &amp; 1.5 (0.8)</td>
</tr>
<tr>
<td>Dolasetron\textsuperscript{b}</td>
<td>6.5 (1.5 to 11.5) &amp; 217 (50) &amp; 0.9 (0.2)</td>
<td>12.2 (3.8 to 20.5) &amp; 403 (88) &amp; 0.5 (0)</td>
</tr>
<tr>
<td>Moxifloxacin</td>
<td>11.9 (6.3 to 17.5) &amp; 1,929 (562) &amp; 1.3 (0.9)</td>
<td>33.4 (28.2 to 38.6) &amp; 4,663 (948) &amp; 0.9 (0.2)</td>
</tr>
<tr>
<td>Dofetilide</td>
<td>14.2 (9.0 to 19.4) &amp; 0.43 (0.1) &amp; 2.9 (1.45)</td>
<td>24.5 (15.7 to 33.3) &amp; 0.92 (0.27) &amp; 3.7 (1.8)</td>
</tr>
<tr>
<td>Levocetirizine</td>
<td>1.8 (–4.1 to 7.6) &amp; 160 (35) &amp; 0.8 (0.5)</td>
<td>3.1 (–4.6 to 10.7) &amp; 1,024 (203) &amp; 0.9 (0.5)</td>
</tr>
</tbody>
</table>

CI: confidence interval; ΔΔQTcF: placebo adjusted change from baseline QTcF.
\textsuperscript{a}After the 1\textsuperscript{st} dose on Day 1, i.e., within 8 hours.
\textsuperscript{b}Hydrodolasetron pharmacokinetic parameters reported.
with the FDA, which identified five QT-positive drugs and the doses at which they should be evaluated. Criteria for negative and positive QT assessment were agreed on and the FDA endorsed the concept that if this pilot study met the criteria for positive QT assessment for all five QT-positive drugs, this would provide evidence to support that a similar approach, applied to early-phase clinical pharmacology studies, can be used as an alternative to the TQT study.

Substantial experience has been gained with ER analysis and the methodology is now routinely used to predict QT effects in the targeted patient population.14,18–24 On the basis of this experience,15 we believe it is now time to consider this methodology in early-phase clinical studies as the primary viable alternative to the TQT study to exclude a small QT effect. Extensive experience with QT-prolonging drugs shows that the QT effect is driven by plasma levels of the drug or main metabolites, with few exceptions, as discussed below. It therefore seems sensible to focus on QT effects in relation to plasma concentration, rather than by time-point without consideration of the pharmacology of the drug. The results of this study lend further support to this paradigm shift: The 5 QT-positive drugs were given at a dose that was expected to cause QT prolongation identified by the FDA as corresponding to the level of regulatory concern. All five drugs met the prospectively agreed criteria of 1) a statistically significant, positive slope of the concentration/QTc relationship and 2) a QTc effect above 10 ms could not be excluded for the lower dose, i.e., the upper bound of the 90% CI of the model-predicted DD QTcF effect at the observed Cmax exceeded 10 ms (Table 3). The criterion for negative QT assessment used in this study corresponds to the one defining a negative TQT study, adapted to the use of ER analysis: A QTc effect above 10 ms must be excluded with the supratherapeutic dose, i.e., the upper bound of the 90% CI of the model-predicted ΔΔQTcF

### Table 3: Slope of the concentration/QTc relationship, geometric mean plasma levels, and projected ΔΔQTcF effect

<table>
<thead>
<tr>
<th>Drug</th>
<th>Slope, mean ms per ng/mL</th>
<th>LB 90% CI</th>
<th>UB 90% CI</th>
<th>Width of  90% CI</th>
<th>Treatment effect ms</th>
<th>Geometric Cmax Day 1, ng/mL</th>
<th>Predicted ΔΔQTcF effect mean, ms</th>
<th>LB 90% CI</th>
<th>UB 90% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive drugs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ondansetron</td>
<td>0.033</td>
<td>0.025</td>
<td>0.042</td>
<td>0.017</td>
<td>0.2</td>
<td>284</td>
<td>9.7</td>
<td>6.2</td>
<td>12.8</td>
</tr>
<tr>
<td>Day 1 only</td>
<td>0.032</td>
<td>0.022</td>
<td>0.043</td>
<td>0.021</td>
<td>0.3</td>
<td>9.5</td>
<td>7.2</td>
<td>13.5</td>
<td></td>
</tr>
<tr>
<td>Parallel design n = 7</td>
<td>0.042</td>
<td>0.031</td>
<td>0.052</td>
<td>0.021</td>
<td>−0.6</td>
<td>259</td>
<td>10.2</td>
<td>6.8</td>
<td>13.5</td>
</tr>
<tr>
<td>Quinine</td>
<td>0.004</td>
<td>0.0034</td>
<td>0.0047</td>
<td>0.0013</td>
<td>−3.0</td>
<td>3623</td>
<td>11.6</td>
<td>6.8</td>
<td>17.1</td>
</tr>
<tr>
<td>Day 1 only</td>
<td>0.004</td>
<td>0.0031</td>
<td>0.0051</td>
<td>0.0020</td>
<td>−4.9</td>
<td>9.8</td>
<td>6.7</td>
<td>17.3</td>
<td></td>
</tr>
<tr>
<td>Parallel design n = 7</td>
<td>0.0034</td>
<td>0.0027</td>
<td>0.0041</td>
<td>0.0014</td>
<td>−2.8</td>
<td>3643</td>
<td>9.5</td>
<td>4.8</td>
<td>14.5</td>
</tr>
<tr>
<td>Dolasetron</td>
<td>0.021</td>
<td>0.013</td>
<td>0.028</td>
<td>0.015</td>
<td>3.1</td>
<td>211</td>
<td>7.4</td>
<td>3.0</td>
<td>11.0</td>
</tr>
<tr>
<td>Day 1 only</td>
<td>0.016</td>
<td>0.0008</td>
<td>0.032</td>
<td>0.031</td>
<td>3.3</td>
<td>6.8</td>
<td>3.4</td>
<td>11.6</td>
<td></td>
</tr>
<tr>
<td>Parallel design n = 7</td>
<td>0.020</td>
<td>0.012</td>
<td>0.029</td>
<td>0.017</td>
<td>3.2</td>
<td>205</td>
<td>7.3</td>
<td>2.7</td>
<td>11.5</td>
</tr>
<tr>
<td>Moxifloxacin</td>
<td>0.0065</td>
<td>0.0059</td>
<td>0.0072</td>
<td>0.0013</td>
<td>2.3</td>
<td>1862</td>
<td>14.5</td>
<td>10.5</td>
<td>17.7</td>
</tr>
<tr>
<td>Day 1 only</td>
<td>0.0045</td>
<td>0.0025</td>
<td>0.0065</td>
<td>0.0041</td>
<td>3.4</td>
<td>11.7</td>
<td>10.6</td>
<td>17.9</td>
<td></td>
</tr>
<tr>
<td>Parallel design n = 7</td>
<td>0.0065</td>
<td>0.0058</td>
<td>0.0072</td>
<td>0.0013</td>
<td>2.2</td>
<td>1708</td>
<td>13.3</td>
<td>9.6</td>
<td>17.0</td>
</tr>
<tr>
<td>Dofetilidea</td>
<td>22.2</td>
<td>18.9</td>
<td>25.6</td>
<td>6.7</td>
<td>1.1</td>
<td>0.40</td>
<td>10.5</td>
<td>6.3</td>
<td>14.9</td>
</tr>
<tr>
<td>Day 1 only</td>
<td>28.7</td>
<td>20.6</td>
<td>36.7</td>
<td>16.1</td>
<td>−0.9</td>
<td>11.3</td>
<td>6.1</td>
<td>14.6</td>
<td></td>
</tr>
<tr>
<td>Parallel design n = 7</td>
<td>25.0</td>
<td>20.9</td>
<td>29.0</td>
<td>8.1</td>
<td>−1.1</td>
<td>0.40</td>
<td>8.9</td>
<td>5.1</td>
<td>13.9</td>
</tr>
<tr>
<td>Negative drug (Day 2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Levocetirazine</td>
<td>0.0014</td>
<td>−0.0013</td>
<td>0.0041</td>
<td>0.0054</td>
<td>0.7</td>
<td>1005b</td>
<td>2.1</td>
<td>−2.3</td>
<td>6.1</td>
</tr>
<tr>
<td>Day 2 only</td>
<td>0.00042</td>
<td>−0.0032</td>
<td>0.0041</td>
<td>0.0073</td>
<td>1.6</td>
<td>2.0</td>
<td>−2.6</td>
<td>6.0</td>
<td></td>
</tr>
<tr>
<td>Parallel design n = 8</td>
<td>−0.0015</td>
<td>−0.0046</td>
<td>0.0017</td>
<td>0.0063</td>
<td>1.8</td>
<td>1014</td>
<td>0.3</td>
<td>−4.7</td>
<td>4.2</td>
</tr>
</tbody>
</table>

CI: confidence interval calculated using a bias-corrected nonparametric bootstrap procedure, which includes variability of Cmax; LB: lower bound; UB: upper bound; Day 1 (Day 2 for levocetirazine): single dose using Day 1 data only for QT-positive drugs and Day 2 data only for levocetirazine.

aFor comparative purposes, parameters and predictions for dofetilide derived from a linear model are shown. bGeometric mean Cmax on Day 2 for levocetirazine; ΔΔQTcF: placebo adjusted change from baseline QTcF. cFor each drug, subjects also dosed with placebo were excluded in this post-hoc analysis.
The predicted effect on $\Delta\Delta QTcF$ using concentration/QTc effect models. The solid black line with gray shaded area denotes the model-predicted mean placebo-adjusted $\Delta QTcF$ with 90% CI as a function of plasma concentration. The horizontal red lines with tick marks show the range of plasma concentrations divided into deciles. Red squares with vertical bars denote the observed arithmetic means and 90% CIs for the placebo-adjusted $\Delta QTcF$ within each plasma concentration decile. The placebo-adjusted $\Delta QTcF$ was derived from the individual $\Delta QTcF$ for the active subtracted by the mean predicted $\Delta QTcF$ for placebo from the model. The blue box denotes the observed geometric mean $C_{max}$ on Day 1 (Day 2 for levocetirizine). (a) Ondansetron; (b) Quinine; (c) Hydrodolasetron; (d) Moxifloxacin; (e) Dofetilide; (f) Levocetirizine.

Effect at $C_{max}$ must be below 10 ms. Levocetirizine at 6-fold the therapeutic dose clearly met this criterion. The study was thereby able to meet its primary objective to demonstrate that a small, early-phase clinical study can detect drugs with a QT effect at the level of regulatory concern and, importantly, the study can also be used to exclude small QT effects for drugs with no underlying clinically relevant effect. In addition, previously described effects on cardiac conduction (PR and QRS intervals) with dolasetron$^{20}$ and quinine$^{25,26}$ were confirmed.

Data from 1-day-only (lower dose of the QT-positive drugs and higher dose of levocetirizine) were also analyzed to evaluate ER analysis when the peak QT effects were lower. This analysis resulted in very similar estimates of the ER slope and the predicted effect (Table 3), but increased the variability of the slope estimate with consequently somewhat wider CIs. The difference was relatively minor and criteria (negative and positive) for all drugs were met also in this analysis (Table 3). It is noteworthy that for dolasetron the mean peak $\Delta\Delta QTcF$ effect by timepoint was only 6.5 ms on Day 1. The finding that this study was able to detect such a small QT effect should provide further reassurance that a drug is truly negative at therapeutic plasma levels, when high doses are evaluated and the QT assessment is negative at clearly supratherapeutic levels.

ER analysis has frequently been used in a descriptive or exploratory way to supplement the primary by timepoint analysis described in the ICH E14 document. This practice has been criticized from a statistical viewpoint$^{27,28}$ when the model was incompletely prespecified. In this study we used this analysis as the primary...
confirmatory analysis and, as a consequence, we prospectively defined criteria for model selection. According to these criteria, hysteresis was excluded and the ER relationship was linear for all drugs except dofetilide; alternative models (Emax, log-linear, and square root) were therefore explored for dofetilide and an Emax model was determined to best fit the data. A nonlinear Emax relationship has been described with other antiarrhythmic drugs and some tyrosine kinase inhibitors with relatively pronounced QT effects, but in the FDA’s review of dofetilide, the ER relation was best described with a linear model. Our finding may well have been a chance finding, to which multiplicity in testing for linearity may have contributed, since substantially lower dofetilide plasma levels were achieved in this study (mean Cmax below 1 ng/mL), as compared to the studies in the dossier, in which levels up to 5 ng/mL were seen. It is important to note that with both a linear and the nonlinear Emax model, criteria for a QT-positive drug were clearly met and the predicted QT effect with both models was similar: 11.6 ms (90% CI 7.0–16.0) with Emax and 10.5 ms (90% CI 6.3–14.9) with linear. The slope of the dofetilide ER relationship using a linear model in this study (22 ms per ng/mL) was within the range described in the Tikosyn label (15–25 ms per ng/mL after single dose). Our finding may well have been a chance finding, to which multiplicity in testing for linearity may have contributed, since substantially lower dofetilide plasma levels were achieved in this study (mean Cmax below 1 ng/mL), as compared to the studies in the dossier, in which levels up to 5 ng/mL were seen. It is important to note that with both a linear and the nonlinear Emax model, criteria for a QT-positive drug were clearly met and the predicted QT effect with both models was similar: 11.6 ms (90% CI 7.0–16.0) with Emax and 10.5 ms (90% CI 6.3–14.9) with linear. The slope of the dofetilide ER relationship using a linear model in this study (22 ms per ng/mL) was within the range described in the Tikosyn label (15–25 ms per ng/mL after single dose). It is, however, worth emphasizing that the objective of early-phase clinical QT assessment is to exclude small QT effects, rather than to provide a robust characterization of the ER relationship. If the QT effect is unambiguous, the consequence would be the same as for a positive TQT study: the QT effect should be further characterized in the targeted patient population. Likewise, a negative assessment would have the same implications as a negative TQT study. There may also be cases in which the sponsor chooses to confirm a less clear signal in a formal TQT study.

For some drugs, the pharmacokinetic profile will require an assessment of ECG effects using multiple dosing over several days, and the same considerations would apply as for the choice between a single- and multiple-dose TQT study. Such drugs include those that demonstrate substantial accumulation, if a sufficiently high single dose cannot be administered. Longer durations of administration will also be required for rare drugs with slowly appearing metabolites causing QT effects, multiple-dose studies in such cases may be suggested by the results from nonclinical assays, e.g., a positive hERG assay with a major metabolite or QT effects in multiple-dose studies in animals. Early-phase clinical studies intended for QT assessment will not incorporate a positive control and it is therefore important to address the concern with respect to this practice. Since the TQT study is used to exclude a small QT effect, the positive

### Table 4 Largest placebo adjusted change from baseline PR

<table>
<thead>
<tr>
<th></th>
<th>Day 1</th>
<th>Day 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (ms)</td>
<td>90% CI (ms)</td>
</tr>
<tr>
<td>Ondansetron</td>
<td>2.5</td>
<td>−5.0 to 10.0</td>
</tr>
<tr>
<td>Quininea</td>
<td>8.9</td>
<td>4.3 to 13.5</td>
</tr>
<tr>
<td>Dolasetron</td>
<td>4.1</td>
<td>−0.1 to 8.3</td>
</tr>
<tr>
<td>Moxifloxacin</td>
<td>0.8</td>
<td>−7.2 to 8.9</td>
</tr>
<tr>
<td>Dofetilde</td>
<td>5.7</td>
<td>−1.3 to 12.6</td>
</tr>
<tr>
<td>Levocetirizine</td>
<td>4.2</td>
<td>−5.2 to 13.6</td>
</tr>
</tbody>
</table>

*a*Including timepoints up to 8 hours post-dose only. Tmax: mean peak plasma level by timepoint; △PR: placebo adjusted change from baseline PR interval.

### Table 5 Largest placebo adjusted change from baseline QRS

<table>
<thead>
<tr>
<th></th>
<th>Day 1</th>
<th>Day 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (ms)</td>
<td>90% CI (ms)</td>
</tr>
<tr>
<td>Ondansetron</td>
<td>0.7</td>
<td>−0.5 to 1.9</td>
</tr>
<tr>
<td>Quininea</td>
<td>4.0</td>
<td>2.4 to 5.7</td>
</tr>
<tr>
<td>Dolasetron</td>
<td>2.1</td>
<td>0.9 to 3.2</td>
</tr>
<tr>
<td>Moxifloxacin</td>
<td>1.0</td>
<td>−0.1 to 2.2</td>
</tr>
<tr>
<td>Dofetilde</td>
<td>0.2</td>
<td>−0.9 to 1.3</td>
</tr>
<tr>
<td>Levocetirizine</td>
<td>0.3</td>
<td>−0.8 to 1.5</td>
</tr>
</tbody>
</table>

*a*Including timepoints up to 8 hours post-dose only. Tmax: mean peak plasma level by timepoint; △QRS: placebo adjusted change from baseline QRS interval.
control serves the purpose of demonstrating that the experimental conditions of the study are sensitive enough to detect a small effect of the investigational compound, should there be one, thereby providing protection against false negatives. The concern over false negatives is the most important from a drug safety perspective, but this risk is, in our view, very low when ER analysis is applied to early-phase QT studies, provided a wide range of drug plasma levels has been achieved and an intense ECG/PK schedule has been implemented using the same experimental conditions and ECG methodologies as in TQT studies. Most of the experience from early-phase QT assessment, however, is still anecdotal, and there are few published examples. To estimate the risk of false negatives and false positives with ER analysis in this setting, simulation of a large number of small studies with 6 to 18 subjects on active treatment and six on placebo was performed in a recently published study by Ferber et al. Data from five TQT studies were used; three studies with moxifloxacin with peak ΔQTcF effect of 12.5, 14.0, and 8.0 ms, one study with ketoconazole with a smaller QT effect (ΔQTcF 7.6 ms), and one with a drug with a larger effect (ΔQTcF 25.9 ms). The criterion for negative QT assessment using ER analysis was the same as in this study. The rate of false negatives was 1% or lower with six subjects on active in two of the three moxifloxacin datasets and around 5% in the third. The simulation provides confidence in QT assessment in small studies, but must obviously be confirmed in real-life clinical trials. The rate of false positives (nonnegatives) was below 20%, with nine subjects on active (six on placebo) and near or below 10% with 12 subjects, which is important from a resource perspective; otherwise, drug developers may have to repeat QT assessments in TQT studies for many drugs.

Based on our results and the extensive experience with ER analysis for evaluation of QT effects that has been gained over the last years, we believe that QT assessment in early-phase clinical studies can be proposed as an alternative or replacement for the TQT study. The following criterion could then be used as a basis for a request for a TQT waiver:

- The upper bound of the two-sided 90% confidence interval of the predicted placebo adjusted ΔQTc should be below 10 ms at the highest clinically relevant plasma concentrations of the drug.

“Clinically relevant plasma concentrations” of the parent compound and abundant metabolites will often not be known at the time of an early-phase clinical study. In case observed plasma levels in the early clinical study do not substantially exceed those later seen in patients with drug concentrations increased by intrinsic or extrinsic factors, it may not be possible to exclude QT effects at these higher concentrations. When ER analysis is used, there will always be plasma concentrations below which the study can be deemed negative. It is therefore important to emphasize that it is the predicted QT effect at the highest clinically relevant concentrations that will define the interpretation of the study. This concern is analogous to the question of whether the selected supratherapeutic dose has been high enough in a TQT study.

The digital, continuous ECG waveforms from the study have been stored and will be made available for public research, under a governance structure similar to the CSRC ECG warehouse.

**Limitations**

This study was designed to provide validation of QT assessment in early-phase clinical studies, but there are important differences as compared to standard FIH studies; only two doses of each drug were studied in a partial crossover design, whereas many dose groups are evaluated in an SAD study. Since precision of the slope of the ER relationship is largely driven by the QT effect at high plasma levels, it does not seem likely that the results would be substantially different by adding groups with lower doses of the drugs.

Unlike many SAD studies, this study was not strictly of parallel design. To evaluate the impact of this difference in design, a post-hoc analysis was performed in which subjects who received placebo were excluded from the active group (“Parallel design” in Table 3); this had little effect on the results and all drugs still met their criteria.

It should also be acknowledged that there may be scenarios that are not easily defined prospectively and with more experience, refined criteria for model selection for ER analysis will be identified to cover exceptional cases.

**METHODS**

**Study design**

The study was a three-period, third-party blinded, randomized, placebo-controlled study in 20 healthy volunteers with the primary objective to evaluate the effect of the drugs on QTcF using ER analysis. Secondary endpoints included pharmacokinetics and effects on other ECG intervals. The design and purpose of the study has been previously published. Each subject underwent three treatment periods, each with 2 consecutive treatment days and with at least 5 days of washout between periods. An incomplete block design resulted in each study drug being administered to nine subjects and placebo to six subjects in separate periods. Six drugs with a well-characterized QT effect were selected for the evaluation, five “QT-positive”: ondansetron, quinine, dolasetron, moxifloxacin, and doxetila, and one “QT-negative” drug, levocetirizine (Table 1). The five QT-positive drugs were chosen in discussions with the FDA and selection criteria included toxicity profile allowing administration to healthy subjects, lack of substantial heart rate effect, and the magnitude of QTc prolongation; the lower dose on Day 1 was recommended by the FDA to achieve a mean QTc effect representing the threshold of regulatory concern: 9 to 12 ms. A higher dose was added on Day 2, which was expected to result in ΔQTcF of about 15 to 20 ms, to mimic an SAD study in which doses that generate plasma levels exceeding therapeutic concentrations are commonly evaluated, with the benefit of increasing the precision of the predicted effect.

Subjects were fasting overnight for 8 hours before dosing and for the first 4 hours postdosing and thereafter received meals in a standardized way. Study treatments were blinded to subjects and the investigating site staff by using third-party dosing and blindfolding of subjects.

The study protocol and Informed Consent Form were reviewed and approved by the Schulman Associates Institutional Review Board, Cincinnati, OH.

**ECG methodology**

Continuous digital 12-lead ECGs were recorded from 1 hour prior to dosing on Day 1 to 24 hours after dosing on Day 2. Subjects were resting quietly in supine position for at least 10 minutes prior to and 5 minutes after timepoints for ECG and PK sampling. The same ECG and PK schedule was used for all treatments, designed to capture the effect near the time of peak plasma levels of all drugs. Twelve-lead ECGs were extracted from the
continuous recording from a 5-minute window preceding the timepoints: −30, −20, and −10 minutes prior to first dose on Day 1 and 0.5, 1, 2, 3, 4, 6, 8, 12, and 24 hours after dosing on both days (i.e., 21 timepoints). ECG intervals were measured at a central ECG laboratory (iCardiac Technologies, Rochester, NY) fully blinded to study treatments, timepoints, and subject identification. At each protocol-specified timepoint, QT and RR intervals were measured from up to 10 second 12-lead ECG recordings (i.e., 10 replicates). High-precision QT analysis was performed on all analyzable beats in the 10 ECG replicates. All “high-confidence” beats and all “low-confidence” beats found acceptable by manual review were included in the analysis. The QTc interval was derived using Fridericia’s formula from the preceding RR interval and the QT interval in each beat and the median QTcF in each replicate was then calculated. The mean across medians from all replicates was used as the subject’s reportable value at that timepoint. Measurement of PR and QRS intervals was performed semi-automatically on three sequential beats from each of the three replicates with the highest confidence score. The mean value was calculated for each replicate and then the mean of these were used as the subject’s reportable value at the timepoint.

Pharmacokinetic methodology
All drug levels were analyzed using liquid chromatography (LC) with tandem mass spectrometric detection (MS/MS).

Ondansetron. Ondansetron and the internal standard, ondansetron-d5, were extracted from human plasma by liquid–liquid extraction. After evaporation under nitrogen, the residue was reconstituted and analyzed. The standard curve range was from 0.1 to 50 ng/mL for ondansetron using a plasma sample volume of 100 μL (Covance Study No. 8231745).

Quinine. Quinine and the internal standard, quinine-d3, were extracted from human plasma by protein precipitation. After evaporation under nitrogen, the residue was reconstituted and analyzed. The standard curve range was from 500 to 15,000 ng/mL for quinine, using a plasma sample volume of 50 μL (Covance Study No. 8295947).

Dolasetron. Hydrodolasetron and the internal standard, hydrodolasetron-d4, were extracted from human plasma by salt-induced phase separation extraction. After evaporation under nitrogen, the residue was reconstituted and analyzed. The standard curve range was from 25 to 1,000 ng/mL for hydrodolasetron, using a plasma sample volume of 50 μL (Covance Study No. 8295945).

Moxifloxacin. Moxifloxacin and the internal standard, moxifloxacin-d4, were extracted from human plasma by protein precipitation. The standard curve range was from 25 to 5,000 ng/mL for moxifloxacin, using a plasma sample volume of 50 μL (Covance Study No. 8225508).

Dofetilide. Dofetilide and the internal standard, dofetilide-d4, were extracted from human plasma by liquid–liquid extraction. After evaporation under nitrogen, the residue was reconstituted and analyzed. The standard curve range was from 0.05 to 25 ng/mL for dofetilide, using a plasma sample volume of 100 μL (Covance Study No. 8295950).

Levocetirizine. While levocetirizine is the R-enantiomer of racemic cetirizine, the method was achiral and designed to determine racemic cetirizine concentrations during sample analysis. Levocetirizine and the internal standard, levocetirizine-d4, were extracted from human plasma by solid-phase extraction. The eluate was diluted and analyzed. The standard curve range was from 20 to 3,000 ng/mL for levocetirizine, using a plasma sample volume of 50 μL (Covance Study No. 8295952).

Calculation of pharmacokinetics
Pharmacokinetic (PK) parameters (Cmax, Tmax, and AUC) were calculated using standard noncompartmental methods for each active treatment, separately for Day 1 and Day 2. Parameters were obtained from time 0 to 24 hours on each day with the exception of Day 1 on quinine, for which the time between the first and second dose (0 to 8 hours) of the drug was used. AUC was derived using the linear-log trapezoidal rule; values below the level of quantification were set to 0 throughout.

Data analysis and interpretation
The primary variable for the ER analysis was the change-from-baseline QTcF (ΔQTcF), where the mean of the three predose ECG readings on Day 1 was used as the baseline. The concentration of the parent compound (for dolasetron, the main metabolite hydrodolasetron) was used as a covariate.

Investigation of hysteresis. Prior to model selection for the ER analysis, the absence of hysteresis was established. To detect hysteresis, individual ΔQTcF was computed as ΔQTcF minus the time-matched mean ΔQTcF of the placebo group. For each day, the time of the largest mean ΔQTcF (Umax) was determined. If the largest mean ΔQTcF exceeded 5 ms at ≥3 timepoints, the time difference between Umax and the Tmax of the drug level exceeded 1 hour, and the one-sided one-sample Wilcoxon test for the difference between ΔQTcF at Tmax and at Umax was formally significant at the 1% level, it was concluded that hysteresis existed. In such a case, a PK model with an additional effect compartment was to replace the model described below.

Model selection. To assess the appropriateness of a linear model, normal QQ-plots for the residuals and plots of weighted residuals vs. concentration and vs. fitted values were produced. A model with a quadratic term in concentration was fitted and the quadratic term was tested on the two-sided 5% alpha level. In case of a significant quadratic term, nonlinear models, such as a log-linear model and an Emax model, were to be investigated and the primary model selected based on the Akaike Information Criterion and plausibility arguments.

ER analysis. In the absence of hysteresis and unless the prespecified test procedure for linearity indicated otherwise, the primary analysis was based on a linear mixed-effects model implemented in R (www.r-project.org), v. 3.1.0 and the lme4 package, v. 1.1-7, with ΔQTcF as the dependent variable, drug plasma concentration as a continuous covariate, treatment (active or placebo), and reduced time (i.e., time with all nonsignificant timepoints combined into one) as categorical factors, and a random intercept per subject within period. In each model, not more than two subjects contributed in two periods (i.e., to both the active drug and placebo). Data from the period under active drug were considered independent of data of the period under placebo within each of these subjects (i.e., different random intercepts were allowed for each period in these two subjects). In other words, each drug was analyzed as if the data came from a parallel group design. The reduced time variable was treated as categorical factor, representing the time elapsed since first drug administration (i.e., time spans both Day 1 and 2). All postdose data from Days 1 and 2 were used. The degrees of freedom for the model estimates were determined by the Kenward-Rogers method (R package lmermeans v. 2.10). From the model, the slope (i.e., the regression parameter for the concentration) and the treatment effect were estimated together with two-sided 90% CIs. A "reduced time” variable was determined from the placebo data using a linear mixed effects model with ΔQTcF as the dependent variable with time as a factor, a random intercept per subject, and the fixed intercept set to zero. Only timepoints with an effect significant at the two-sided 10% alpha level in this model were retained; all other timepoints were assigned to a common level “Time 0.”

The predicted mean ΔQTcF at the observed geometric mean Cmax (i.e., the product with the slope estimate + treatment effect [active – placebo]) was calculated. Two-sided 90% CIs of the estimate were calculated using a bias-corrected nonparametric bootstrap procedure in the boot package, v. 3.1.11 with 3,000 resamples and subject as the unit of
resampling. Resampling was done independently for the active and the placebo subjects/periods. For each resample, the model was fitted and the prediction was made at the geometric mean C_{max} determined from the resampled data. The CI was determined from the distribution of resampled predicted values.

Criteria for QT assessment. Criteria for the "QT-positive" drugs were based on the predicted QTc effect of the lower dose and the criterion for the "QT negative" drug was applied to the predicted effect of the higher dose.

To demonstrate a QT effect of the 5 "QT-positive" drugs:

- The upper bound of the two-sided 90% CI of the predicted mean ΔΔQTcF was to be greater than 10 ms at the observed geometric mean C_{max} on Day 1.
- The slope of the concentration/QTc effect relationship was to be statistically significant.

To exclude a QT effect of concern for the "QT negative" drug (levocetirizine):

- The upper bound of the two-sided 90% CI of the predicted mean ΔΔQTcF was to be less than 10 ms at the observed geometric mean C_{max} on Day 2.

ΔΔQTcF by timepoint. For each timepoint, an analysis of variance model was fitted with ΔΔQTcF as dependent variable and treatment (active or placebo) as factor and baseline QTcF as a covariate. From this model, the difference (active – placebo) was estimated with a two-sided 90% CI. Separate models were fitted for each treatment, all of them using the same placebo data. Change-from-baseline in heart rate, QTcF, PR, and QRS were calculated using descriptive summary statistics.

The Interdisciplinary Review team for QT studies at the FDA received the full analysis dataset and performed an independent analysis compliant with the prospectively agreed statistical analysis plan. The study was conducted between February and June 2014.

Study Highlights

WHAT IS THE CURRENT KNOWLEDGE ON THIS TOPIC?

✓ Definitive assessment of a drug’s effect on ECG intervals is typically performed in a designated stand-alone TQT study.

WHAT QUESTION DID THIS STUDY ADDRESS?

✓ Can definitive ECG assessment be performed as part of a standard early-phase clinical study, such as the first-in-human study?

WHAT THIS STUDY ADDS TO OUR KNOWLEDGE

✓ The study, which was designed in collaboration with the FDA, demonstrated that QT assessment using exposure–response analysis of data from a small study in healthy volunteers was able to detect mild QT prolongation at the level of regulatory concern and that a QT effect above 10 ms could be excluded for a drug with no underlying effect. The study thereby provides validation of the concept of definitive ECG assessment in early-phase clinical studies.

HOW THIS MIGHT CHANGE CLINICAL PHARMACOLOGY AND THERAPEUTICS

✓ The study demonstrates that ECG assessment in early-phase clinical studies can be used as an alternative to a TQT study to exclude small QT effect by a new drug.

ABBREVIATIONS

Δ Change-from-baseline
ΔΔ Placebo-adjusted, change-from-baseline
CSRC Cardiac Safety Research Consortium
ER Exposure–response
FIH First-in-human
IQ International Consortium for Innovation and Quality in Pharmaceutical Development
MAD Multiple ascending dose
SAD Single ascending dose
TQT Thorough QT

ACKNOWLEDGMENTS

The authors thank Donna Kowalski, Astellas, Bela Patel, GSK, Dutta Santanu, Covance and Kuenhi Tsai, Merck for their contribution with regard to discussions on the statistical analysis of the study. The study was funded by iCardiac Technologies, Rochester, NY, USA.

CONFLICT OF INTEREST/DISCLOSURE

B.D. holds stock and stock options in iCardiac Technologies. B.S. holds stock options in iCardiac Technologies.

AUTHOR CONTRIBUTIONS

B. D., C. B., C. D., G. F., C. G., V. J., L. J., J. K., K. K., C. O-R., S. R., N. S., and N. S. designed the research; R. R. S. and B. S. performed the research; G. F., M. Z., and J. L. analyzed the data; all authors wrote the article.

© 2014 American Society for Clinical Pharmacology and Therapeutics


