Detection of QTc Effects in Small Studies—Implications for Replacing the Thorough QT Study

Georg Ferber, Ph.D.,* Meijian Zhou, Ph.D.,† and Borje Darpo, M.D., Ph.D.†‡

From *Statistik Georg Ferber GmbH, Riehen, Switzerland; †iCardiac Technologies, Rochester, NY; and ‡Karolinska Institute, Division of Cardiovascular Medicine, Department of Clinical Sciences, Danderyd's Hospital, Stockholm, Sweden

Background: ECG assessment with exposure response analysis applied to data from First-in-Man studies has been proposed to replace the thorough QT study for the detection of small QT effects.

Methods: Data from five thorough QT studies, three with moxifloxacin, one study with a drug with a large QTc effect (~25 ms) and one with ketoconazole with a smaller QT effect (~8 ms) were used. By subsampling, studies with 6–18 subjects on drug and six on placebo were simulated 1000 times per sample size to assess whether small QTc effects using ICH E14 criteria could be excluded and the impact of sample size on the estimate and variability of the slope of the concentration/QTc relation.

Results: With a sample size of nine or more on drug and six on placebo, the fraction of “false negative studies” was at or below 5% with data from the studies with moxifloxacin and from the drug with a large QTc effect. With the same sample size and no underlying QTc effect (placebo), the fraction of studies in which an effect above 10 ms could be excluded was above 85%. A treatment effect in the linear concentration-effect model resulted in a lower proportion of “false negatives.” Sample size had little influence on the average slope estimate of the concentration/QTc relationship.

Conclusions: For drugs with a QTc effect of around 12–14 ms, exposure response analysis applied to First-in-Man studies with careful ECG assessment can be used to replace the thorough QT study.

Ann Noninvasive Electrocardiol 2015;20(4):368–377

QT; QTc; thorough QT study; clinical pharmacology; First-in-Man; ICH E14

The ICH E14 clinical guidance on QT assessment requires the majority of new chemical entities with systemic availability to be tested in a study designed to exclude small drug-induced QTc effects, the so-called “thorough QT/QTc” (TQT) study.^{1,2} The TQT is typically performed in healthy volunteers and should include evaluation of the QT effect of supratherapeutic plasma levels of the drug to ensure that levels that can be reached in patients with impaired clearance of the drug are covered.^{3,4} To claim that a drug is devoid of proarrhythmic risk, a QTc effect above 10 ms must be excluded at the peak plasma level (C_{max}) and all other time point postdosing (“by time point” analysis). This criterion is met when the upper bound of the two-sided 90% confidence interval (CI) of the placebo-corrected, change-from-baseline QTc

($\Delta\Delta QTc$) is below 10 ms, a so-called “negative TQT study.” For drugs for which this “threshold” effect cannot be excluded (i.e., a “positive TQT study”), the QTc effect and its potential consequences (adverse events of ventricular arrhythmias or signs/symptoms suggestive thereof) will have to be studied in the targeted patient population, resulting in additional ECG monitoring in phase 2 and 3 trials. The TQT study is a resource intensive study⁵ with the sole purpose of evaluating ECG effects; if the same data can be generated from a routine clinical pharmacology study performed as a standard component of early clinical development, such as the First-in-Human (FIH) study, this would clearly represent a more efficient approach from a resource perspective. In the single-ascending-dose (SAD) part of a FIH study, escalating doses of the

Address for correspondence: Borje Darpo, M.D., Ph.D., Associate Professor Karolinska Institutet, Division of Cardiovascular Medicine, Department of Clinical Sciences, Danderyd's Hospital, Stockholm, Sweden. Fax: +46 8 766 16 32; E-mail: borje.darpo@telia.com

© 2014 Wiley Periodicals, Inc.
DOI:10.1111/anec.12227

drug are given to small cohorts of subjects, typically six subjects on active +2 on placebo/cohort, often up to the maximum tolerated dose. Since several doses are administered to small cohorts, the ability (i.e., the power) of an analysis "by time point" for each of these cohorts will be low and will not allow the exclusion of small QT effects.⁶ Exposure response (ER) analysis, on the other hand, uses all QT and plasma concentration data from all cohorts and enters this into one statistical model; this results in a substantially improved precision of the estimated QTc effect. ER analysis has been extensively used for extrapolation of QT effects from the TQT study to the targeted patient population⁷ and the experience with this approach has increased considerably over the last years, both among regulators and pharmaceutical industry. "Early ECG assessment" with robust ECG monitoring and ER analysis on data from, e.g., the SAD study could represent an opportunity to generate ECG data with the same high level of confidence as from the TQT study.^{8–11}

Even though ER analysis results in a more precise estimation of the QT effect of a drug as compared to the "by time point" analysis, the question remains of whether exclusion of such a small effect as 10 ms with cohorts of only six to nine subjects is possible. Published experience is still relatively scarce and anecdotal,^{9,10} and a comprehensive comparison between "Early ECG assessment" and results from the TQT study applied prospectively to a larger series of drugs is lacking. The IQ-CSRC prospective study was therefore undertaken in collaboration with the FDA to evaluate whether "early ECG assessment" could detect the effect of five QT-prolonging drugs and exclude a QT effect for a negative drug; the design of the study has been published¹² and results are expected during the fall of 2014. As a preparation for the IQ-CSRC study, we investigated the power of small studies by the use of simulation (see Table 3 in Ref. 12) and have now expanded this work to five datasets from TQT studies. Simulation is a valuable tool to assess the performance of analysis methods that are too complex to allow the derivation of the probability distribution of the variables of interest analytically.^{13,14} In principle, there are two ways to perform such simulations: Observations can be created based on a statistical model with known properties, or experimental data collected in similar settings can be rearranged by subsampling according to the setting of interest.

The strength of the former method (parametric simulation) is that results can be compared to the parameters used for simulation, thereby allowing estimation of any bias of a method. The limitation of parametric simulation is that it hinges on the assumptions inherent in the model used to create the data. Simulations based on subsampling from real data, as was done for this study, overcome this problem, but is only reasonable if there is some established standard to compare the results against. For QT assessment, the TQT study serves as such a standard and we therefore used data from five TQT studies: Two studies with moxifloxacin with an expected peak QT effect of approximately 12 ms, two studies with a smaller QT effect (one with moxifloxacin and one with ketoconazole) and one study with a drug with a larger QT effect of around 25 ms.

METHODS

ECG and PK data from five TQT studies in healthy volunteers were used, three studies with moxifloxacin as a positive control, one with ketoconazole and one with an investigational drug with a large QTc effect. For these studies, data from subjects with available ECG and PK data from the placebo and drug treatment (moxifloxacin, ketoconazole, and "drug with a larger QT effect") were used:

- (1) Study 1 (moxifloxacin): Moxifloxacin and placebo data from 50 subjects from the TQT study with lenvatinib.¹⁵ Subjects received a single oral 400 mg dose of moxifloxacin, lenvatinib, and placebo in a crossover design with a predose baseline. The moxifloxacin peak plasma level (C_{max}) and QTc effect for included subjects were consistent with previous experience¹⁶: 2.8 $\mu\text{g/mL}$ (90% CI: 2.6–3.0) and a largest placebo-controlled, change-from-baseline QTcF ($\Delta\Delta\text{QTcF}$) of 12.5 ms (90% CI: 10.8–14.2; Fig. 1A). The slope of the relationship between moxifloxacin plasma levels and $\Delta\Delta\text{QTcF}$ in the used dataset from this study was 2.6 ms/ μg per mL (90% CI: 2.00–3.19), which is in the low expected range based on a pooled analysis of 20 TQT studies (3.1 ms/ μg per mL (90% CI: 2.8–3.3)¹⁶).
- (2) Study 2 (moxifloxacin): Moxifloxacin and placebo data from 29 subjects from a

Figure 1. Placebo-corrected, change-from-baseline QTcF ($\Delta\Delta\text{QTcF}$) and drug plasma levels in the full included dataset from the five TQT studies: (A) Moxifloxacin Study 1; (B) Moxifloxacin Study 2; (C) Moxifloxacin Study 3; (D) Ketoconazole study; and (E) Study with drug with a large QT effect.

methodology study.¹⁷ Subjects received a single oral 400 mg dose of moxifloxacin in a cross-over design with a predose baseline. The moxifloxacin peak plasma level (C_{max}) for

included subjects was 3.1 $\mu\text{g/mL}$ (90% CI: 2.9–3.4) and the largest $\Delta\Delta\text{QTcF}$ was 14.0 ms (90% CI: 11.3–16.7 ms; Fig. 1B). The slope of the concentration/QTc effect in this dataset was

within the expected range, 3.1 ms/ μ g per mL (90% CI: 1.80–4.37).

- (3) Study 3 (moxifloxacin): Moxifloxacin and placebo data from 65 subjects from a QT study with an undisclosed drug in a parallel design study. Subjects received placebo or a single oral dose of 400 mg moxifloxacin on either Day 2 or 14 for a crossover comparison and predose values on the day of dosing were used as baseline. The moxifloxacin peak plasma level ($C_{\max}$) for included subjects was lower in this study (1.7 μ g/mL [90% CI: 1.6–1.8]) as compared to other studies with the same dose of moxifloxacin.¹⁶ The peak QT effect after dosing with moxifloxacin was also lower as compared to studies 1 and 2 with $\Delta\Delta QTcF$ was 8.0 ms (90% CI: 5.3–10.7; Fig. 1C). The slope of the concentration/QTc effect was clearly higher in this dataset as compared to the expected range, 4.6 ms/ μ g per mL (90% CI: 3.88–5.37).
- (4) Study 4 (ketoconazole): Ketoconazole and placebo data from 50 subjects from the QT study with lomitapide, in which ketoconazole was used as a metabolic inhibitor.¹⁸ Subjects received placebo or ketoconazole 200 mg BID for 3 days in a crossover design with a predose baseline. Data from Day 3 were used for this study. The ketoconazole peak plasma level ($C_{\max}$) for included subjects was 5.7 μ g/mL (90% CI: 5.3–6.1) and the largest $\Delta\Delta QTcF$ was 7.6 ms (90% CI: 5.5–9.7; Fig. 1D); the slope of the concentration/QTc relationship was 1.0 ms/ μ g per mL (90% CI: 0.64–1.30), which is similar to the ketoconazole reported in a case study by Dr. Zhu et al.¹⁹
- (5) Study 5 ("drug with large QTc effect"): In a three-way crossover design, placebo, moxifloxacin and an anonymized drug was given to healthy volunteers. Among included subjects ($n = 51$), the drug peak plasma level ($C_{\max}$) for included subjects was 15.8 μ g/mL (90% CI: 144–171) and the largest $\Delta\Delta QTcF$ 25.9 ms (90% CI: 23.9–28.0; Fig. 1E). The concentration/QTc slope in the included dataset was 1.46 ms/ μ g per mL (90% CI: 1.38–1.55).

ECG and Pharmacokinetic Methodology

In all studies, 12-lead ECGs were continuously recorded and ECGs were extracted at predefined time points. In studies 1, 2, 4, and 5, QT measurements were made with the High

Precision QT technique¹⁷; in Study 3, intervals were measured by a central ECG laboratory which used a semiautomated technique with three replicates per time point. Determination of plasma concentrations were performed using validated methodologies.

Data Analysis

For each QT study and each sample size, 1000 studies were simulated by sampling subjects from the study without replacement. To simulate small studies with a drug that has a QTc effect, moxifloxacin/ketoconazole/drug plasma concentration and corresponding change-from-baseline QTcF ($\Delta QTcF$) data from 6–18 subjects on active and six subjects on placebo in each study were selected. Likewise, to simulate small studies with a drug without QTc effect, moxifloxacin/ketoconazole/drug plasma concentration data were matched with $\Delta QTcF$ data from the placebo arm for 6–18 subjects on active (no effect) and six on placebo. Sample sizes (six to nine on active, six on placebo), which are commonly used in First-in-Man studies were selected, but also larger (12, 15, and 18 on active), to evaluate the impact of sample size on the performance of the concentration/QTc model.

Two concentrations-response models were fitted to each of the datasets obtained this way. The first one was a linear mixed effects model with $\Delta QTcF$ as dependent variable, plasma concentrations as covariate and discrete time as a factor.²⁰ The second model in addition included a treatment effect (moxifloxacin/ketoconazole/drug with large effect or placebo), as described in the publication on the design of the IQ-CSRC prospective study.¹² Only an intercept per subject was included as a random effect. Based on the models fitted, the effect on the placebo-corrected $\Delta QTcF$ at the observed geometric mean $C_{\max}$ was predicted together with a two-sided 90% CI. The null hypothesis that the slope of the model was zero was also tested, using the two-sided *t*-test supplied by the function `lme` in the R package `nlme`.²¹ The slope was considered significant if this *P* value was <0.05 . For each scenario, descriptive statistics were calculated across the simulations. The fractions of negative studies (as defined below), the fraction of studies with a significantly positive slope, the mean and variability of the slope of the concentration/QTc relation as estimated across

Figure 2. An example from a simulated study with nine subjects on drug with no QT effect and six on placebo, in which a QTc effect above 10 ms can be excluded ("negative QT assessment"). The upper bound of the 90% confidence interval of the predicted QTc effect (grey shaded area) is below 10 ms (dotted line) within the range of observed plasma levels.

the simulated studies and the impact of adding a treatment effect into the linear model were evaluated. The slopes of the individual simulated studies were also compared to the slope obtained from the full data of each study and its 90% CI.

Criteria for QT Assessment

Previously suggested criteria for a negative QT assessment were used,^{8,9} which are consistent with current regulatory practice for reviews of TQT studies and to project QTc effects in the targeted patient population (see e.g.,^{22,23}):

- The upper bound of the two-sided 90% CI of the predicted placebo-corrected ΔQTcF at the observed geometric mean C_{max} level of the studied drug is below 10 ms (Fig. 2).

When this criterion is met for a study, it means that a QTcF effect above 10 ms can be excluded at the observed peak plasma level of the drug.

Figure 3. (A) Fraction of negative studies across increasing sample sizes for subjects on drug and 6 subjects on placebo. Levels below 5% are highlighted in grey. The fraction of negative studies with active (QT prolonging) drugs corresponds to the rate of false negatives, i.e., when the study incorrectly concludes that a QTc effect above 10 ms can be excluded for a QT-prolonging drug. (B) Fraction of negative studies with the no-effect simulation using placebo QTc and drug plasma concentration data. With a sample size of nine on drug and six on placebo, the fraction is above 90% for three of the studies (studies 1, 2, and 5) and above 85% for studies 3 and 4. The fraction of nonnegative studies (1–fraction of negatives) in the no-effect scenario corresponds to the rate of false positives, i.e., when the study cannot exclude a QTc effect above 10 ms even though the drug does not cause QTc prolongation.

RESULTS

Fitting linear mixed effects models to the subsets of subjects was possible in all subsets generated and all parameters foreseen could be calculated.

Using a model with a treatment effect, the fraction of 1000 simulated studies that were negative was close to zero with data from moxifloxacin studies 1 and 2 and around 5% with data from Study 3 (moxifloxacin) with a sample size of nine or more subjects on drug and six on placebo (Fig. 3A). With data from Study 5 (drug with large QT effect), no negative studies were observed, whereas the fraction was considerably

Figure 4. Fraction of simulated studies with a significantly, positive slope of the plasma concentration/ $\Delta\Delta\text{QTcF}$ relationship. With a sample size of 12 subjects on drug and six on placebo, the fraction of studies with a positive slope was above 80% for three of the studies. A sample size of 15 subjects was required to achieve a fraction at or above 90% for three of the five studies (studies 1, 3, and 5).

higher (25-30%) with data from the ketoconazole study (Study 4). For the no-effect scenario using placebo data, the fraction of negative studies was above 90% on data from studies 1, 2, and 5 and above 85% for studies 3 and 4 with a sample size of nine on drug and six on placebo (Fig. 3B).

Using the same model with a treatment effect, the slope of the relationship between the QTc effect and drug plasma levels was significantly positive in all simulated cases across all sample sizes on data from Study 5 (drug with large QT effect; Fig. 4). For simulations based on moxifloxacin studies 1 and 3, the fraction of studies with a significantly positive slope was above 80% for samples above 12 subjects on drug (with six on placebo). With data from studies 2 (moxifloxacin) and 4 (ketoconazole), the fraction of studies with positive slope was substantially lower, between 45% and 60% with a sample size of 9-12 on drug.

The mean slope of the relationship between the QTc effect ($\Delta\Delta\text{QTcF}$) and drug plasma levels was not much affected by the sample size in the simulated studies with a model that included a treatment effect (Fig. 5A; Table 1). For Study

Figure 5. (A) The mean slope of the plasma concentration/ $\Delta\Delta\text{QTcF}$ relationship as a function of sample size (with six subjects on placebo for all scenarios). The mean slope for all simulated studies was only to a small extent affected by the sample size. Even though not directly comparable, slopes for the simulated studies of parallel design were similar to the mean slope of the full crossover dataset:

Study 1: 2.6 $\text{ms}/\mu\text{g}$ per mL (90% CI: 2.00–3.19);
Study 2: 3.1 $\text{ms}/\mu\text{g}$ per mL (90% CI: 1.80–4.37);
Study 3: 4.6 $\text{ms}/\mu\text{g}$ per mL (90% CI: 3.88–5.37).
Study 4: 1.0 $\text{ms}/\mu\text{g}$ per mL (90% CI: 0.64–1.30)
Study 5: 1.46 $\text{ms}/\mu\text{g}$ per mL (90% CI: 1.38–1.55).

(B) Proportion of simulated studies for which the mean concentration/effect slope fell inside the 90% CI of the corresponding slope based on the full dataset.

3 with moxifloxacin, in which both the peak $\Delta\Delta\text{QTcF}$ effect and the moxifloxacin mean C_{max} level were smaller than in the other two studies, the mean slope was clearly larger than in the two other studies with moxifloxacin (studies 1 and 2), around 4.5 $\text{ms}/\mu\text{g}$ per mL versus 2.5–3 $\text{ms}/\mu\text{g}$ per mL. For all five studies, the mean slope with 6-18 subjects on drug and six on placebo were comparable to the slope in the full dataset. The proportion of simulated studies for which the mean slope fell inside the 90% CI of the corresponding slope based on the full dataset. 30-60% for all studies except Study 2 (moxifloxacin) with 15 subjects (71%) and 18 subjects (78%) on

Figure 6. Fraction of negative studies as a function of number of subjects on active (with six on placebo) with a linear concentration /QTc effect model without treatment effect. The result should be compared with Figure 3, in which a model with a treatment effect was used. The model without treatment effect resulted in an unacceptably high proportion of false negative studies on data from drugs with a small QTc effect from studies 1, 3, and 5 (A). In the no-effect scenario, using placebo data, the fraction of negative studies was larger for all studies, i.e., the proportion of false positives was lower (B).

active. The proportion consistently increased as the sample size grew with data from all studies except Study 5 (drug with large QTc effect; Fig. 5, panel B).

Using a concentration/effect model without a treatment effect had a clear effect on the fraction of negative studies. Figure 6 shows the fraction of negative studies using a model without a treatment effect and should be compared with Figure 3, which shows the results when a model with treatment effect is employed. For drugs with a QT effect, a model with a treatment effect resulted in a lower fraction of negative studies (Fig. 3A and Fig. 6A). With a model without a treatment effect, the fraction of negative studies was never below 5% for studies 1 and 3 (moxifloxacin) and substantially higher for Study 4 (ketoconazole); 50–60% (Fig.

6A) as compared to 25–30% (Fig. 3A). In the “no-effect” scenario with placebo, the fraction of negative studies was also higher with a model without treatment effect; above 90% with six subjects on active with placebo data from all studies except Study 4 and above 95% for all drugs with a sample size of nine subjects or more on active (Fig. 3B and Fig. 6B).

DISCUSSION

A negative QT assessment has important implications for subsequent ECG monitoring in late stage clinical trials. Consequently, the rate of false negatives, i.e., when the test fails to demonstrate an existent safety signal, must be very low. Based on the accumulated experience since the adoption of the ICH E14 clinical guidance in May 2005 and the understanding that no drugs with an unknown QT liability has since been approved, this criterion seems to have been met with the TQT study. It is therefore appropriate to request that a new methodology proposed to replace or to serve as an alternative to the TQT study also results in a very low incidence of false negatives. We therefore tested the ability of small studies to detect the QT effect of two drugs (from four separate studies) with a relatively mild prolongation (6–12 ms) and one drug with a pronounced effect (~25 ms). To detect an effect in this context was defined by using the ICH E14 criterion adapted to ER analysis: an effect on the placebo-corrected, change-from-baseline QTc above 10 ms must be excluded at the observed mean C_{max} of the drug, i.e., the upper bound of the 90% CI of the model-projected effect should be below 10 ms. Drug and placebo data from TQT studies were used and small studies with 6–18 subjects on active and six on placebo were simulated by resampling of the data 1000 times per sample size and drug for a total of 25,000 studies (five drugs with five sample sizes for the active). By using the simulation approach, the likelihood of a given result can be estimated across different sample sizes and the predictive value of the test can be studied.

The rate of false negatives was extremely low for the drug with a large QT effect, which is expected, but also for both moxifloxacin studies in which a mean QTc effect between 12 and 14 ms was demonstrated in the “by time point” analysis of the full dataset (studies 1 and 2; Fig. 3A). With only

six subjects on active, the fraction of studies with a negative result with these drugs was zero or close to zero and well below the 5%, which was set as an acceptable level. With data from moxifloxacin Study 3, in which a lower peak QTc effect (8 ms) was demonstrated in the TQT study, the rate of false negatives was somewhat higher but just above 5% (6.7%) with six subjects on active and at or slightly below 5% for nine subjects and above (5.0%, 5.2%, 4.1%, and 4.8% for 9, 12, 15, and 18 subjects). With ketoconazole, with a smaller QTc effect, the rate of false negatives was between 25% and 30% regardless of the sample size for active, which can be deemed unacceptably high. These results are encouraging and suggest that for a drug with a QTc effect above 10 ms, ER analysis of data from serial ECG monitoring in early clinical pharmacology studies can be used with a similarly low incidence of studies that incorrectly exclude a QT effect of concern. To strengthen this point further, it should be pointed out that the doses of moxifloxacin (400 mg) and ketoconazole (200 mg BID) in these studies were not supratherapeutic; in practice, doses up to the maximum tolerated dose are often tested in the SAD study and achieved plasma levels many times exceed those seen in the targeted patient population. As previously suggested,⁹ a criterion for a negative QT assessment based on ER analysis of data from a SAD study could be to exclude a QTc effect above 10 ms at plasma concentrations that cover levels seen in patients with impaired clearance of the drug, due to intrinsic (e.g., hepatic or renal impairment or P450 CYP polymorphism) or extrinsic (e.g., drug interactions) factors. If a QTc effect above 10 ms can be excluded at plasma levels substantially exceeding those later observed in patients, it appears reasonable to expect that the likelihood that the drug causes QT prolongation of concern in patients is even lower than 5%.

It is also important that the rate of false positives, i.e., when a safety signal cannot be excluded for a negative drug, is low. In the context of QT assessment, this would correspond to the scenario where a QTc effect above 10 ms cannot be excluded for a drug without underlying effect. If this criterion cannot be met in a sufficiently large proportion of FIM studies, the QT assessment would need to be repeated in a TQT study for many projects, which would be inefficient from a resource perspective. We tested this by simulating small studies on data from the placebo treatment in the TQT studies and set the acceptable level to

90%, similar to the power of a statistical test. Small studies simulated on placebo data from studies 2 and 5 exhibited an acceptable fraction of negative studies, above 90%, with only six subjects on active and this criterion was also met with placebo data from Study 1 with nine subjects (Fig. 3B). The fraction of negative studies were somewhat lower for studies 3 and 4, but still above 80% (89% and 85%, respectively) with nine subjects or more on active. Since placebo data were used for this analysis some differences between the studies are worth noticing: Even though all studies were of crossover design, studies 1, 2, and 5 were single dose studies with the QT assessment performed on Day 1; in Study 3 the measurements were made after 14 days of dosing and in Study 4 after 3 days. The precision of Δ QTcF during placebo treatment, measured as the mean standard deviation across time points, was better in the single-dose studies: 6.8, 6.7, and 6.8 ms for studies 1, 2, and 5, respectively, as compared to 8.3 and 8.2 ms for Study 3 and 4. The measurement precision is determined not only by the QT measurement technique,¹⁷ which was a high precision technique for all studies except Study 3 (in which a semiautomated technique was used), but also the design and the experimental conditions at the site, i.e., how subjects are handled during the ECG extraction time windows. Higher precision results in more narrow CIs of the projected QTc effect and is therefore an important determinant for the likelihood of a false positives result. It therefore seems highly relevant to pay attention to all aspects of a QT assessment study that may impact the precision of the measurement, including design, with preference for single-dose, crossover studies, experimental conditions, with preference for sites with a demonstrated track record and measurement technique, with preference for those with high precision. Which rate of false positives a sponsor can accept will obviously vary: our findings, however, support that QT assessment with ER analysis on data from FIM studies can be performed with a rate of false positives below 20% and even below 10% in SAD studies with stringent control of experimental conditions, optimal design and high precision of the QT measurement.

Besides the drug with a large QT effect, for which the slope of the concentration/QT effect relation was significant across all sample sizes, this metric did not perform well with data from drugs that caused small QT prolongation. Data from Study 3, with the largest slope among the

moxifloxacin studies (4.6 vs. 2.6 and 3.1 ms/ μ g per mL), also resulted in a significant slope in more than 90% of simulated studies with nine or more subjects on active, but Study 2 with moxifloxacin and the ketoconazole study resulted in a significant slope in only about 50% of simulated studies with the same sample size (Fig. 4). It has been proposed that the absence of a positive slope, combined with an upper bound of the 90% CI below 10 ms, could constitute a negative QT assessment.²⁴ When such a combination of criteria was tested, the slope parameter did not further reduce the rate of false negatives, which is understandable based on the poor performance of the slope parameter. It should, however, again be pointed out that the doses of moxifloxacin and ketoconazole were not supratherapeutic. As shown by parametric simulation on moxifloxacin data performed by Dr Garnett as a preparation for the IQ-CSRC prospective study, adding a higher dose (800 mg) substantially improves the precision of the slope estimate in small studies, which would increase the likelihood of demonstrating a significant slope in such studies (see Table 4 Darpo et al.¹²).

It is interesting but not surprising that the average slope across 1000 simulated studies was not much affected by the sample size on active and was comparable to the mean slope in the full dataset (Fig. 5A). The proportion of simulated studies for which the mean slope fell inside the 90% CI of the projected QTc effect in the full dataset did, however, not exceed 60%, with a few exceptions, which can be regarded as unsatisfactory. It should be emphasized that the slope parameter derived from a parallel designed study is similar but not identical to the slope from a crossover TQT study, in which each subject's $\Delta\Delta\text{QTcF}$ is used in the concentration/effect model. Even so, based on these data it appears that for drugs with an observed QTc effect, larger studies may be needed to obtain an estimate of the projected QT effect in patients. This would also be consistent with the ICH E14 requirement that drugs for which a QTc effect cannot be excluded, should undergo further evaluation of the QT effect and its' potential proarrhythmic consequences.² Estimation of the QT effects in different patient populations is then also preferably performed using exposure/response analysis.⁷

In addition to the fixed time effects, a treatment effect (active or placebo) can be included in the linear concentration/QTc effect model. When

the two approaches (with and without treatment effect) were compared for the simulated studies, it disclosed important differences in terms of the rate of false negatives and false positives, which need to be accounted for. Roughly speaking, a treatment effect corresponds to the intercept in an analysis based on $\Delta\Delta\text{QTc}$, i.e., where the projected slope would cross the y-axis. A treatment effect that is statistically different from zero can be taken as an indication that the model fit is not very good. As an example, if there is nonlinearity with high plasma concentrations producing a less than proportional QTc effect, a positive treatment effect would be expected. The presence of hysteresis, when the peak QTc effect is observed later than the observed peak drug plasma levels, could also result in a treatment effect. Based on our data, it may be tempting to use a model without a treatment effect, since such a model seems to have a lower likelihood of false positive results (Fig. 6B). Importantly, however, the same model seems to have a higher likelihood of false negative results (Fig. 6A), which is unacceptable from a safety perspective. With a treatment effect, an additional indication of how well the model fits the data is also obtained. Evidently, more experience from small studies is needed and should be made publicly available. Until then, it seems prudent to err on the safe side and include a treatment effect in the concentration/QTc effect model for small QT assessment studies.

With a changing regulatory environment and potential changes in ICH clinical guidance for QT assessment, a large prospective series that will allow a comparison of QT data from FIM studies and TQT studies may realistically never be generated. We have tried to fill this knowledge gap by simulating a large number of small studies from TQT study data. This approach has inherent limitations, primarily in respect to the limited extent of utilized raw data and our findings obviously need confirmation from FIM studies. The IQ-CSRC prospective study¹² can be seen as a first step in this process and will, regardless of successful outcome or not, generate data of general interest. It will also allow us to further test a wider applicability of some of the findings from this simulation study, such as the role of a treatment effect in the concentration/QTc effect model. In addition, it is important that drug developers with access to large amount of data publish their experience in the same spirit of collaborative

research, which has enabled the IQ consortium²⁵ and the Cardiac Safety Research Consortium²⁶ to design and conduct the prospective study in collaboration with FDA.

Limitations of this simulation study include the use of only one dose of each drug; in practice, many doses are evaluated in a typical SAD or multiple ascending dose study, which will result in more subjects than six or nine on active and often more than six in the pooled placebo group. This is likely to result in a greater power of the exposure/response analysis, which will affect both the rate of false negatives and false positives.

In conclusion, our study supports that small FIM studies with serial ECG assessment and stringent experimental control can be used to exclude a QTc effect of clinical and regulatory concern with an incidence of false negatives below 5%.

REFERENCES

1. ICH E14 Questions & Answers. Available at: http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E14/E14_Guideline.pdf. 2014. Accessed October 16, 2014.
2. ICH Harmonized Tripartite Guideline E14. The Clinical Evaluation of QT/QTc Interval Prolongation and Proarrhythmic Potential for Non-Antiarrhythmic Drugs. Available at: http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E14/E14_Guideline.pdf. Accessed October 16, 2014.
3. Darpo B, Sager P. Design of the "Thorough QT Study". *Clin Pharmacol Ther* 2008;83:529.
4. Darpo B. The thorough QT/QTc study 4 years after the implementation of the ICH E14 guidance. *Br J Pharmacol* 2010;159:49–57.
5. Bouvy JC, Koopmanschap MA, Shah RR, et al. The cost-effectiveness of drug regulation: The example of thorough QT/QTc studies. *Clin Pharmacol Ther* 2012;91:281–288.
6. Zhang J, Machado SG. Statistical issues including design and sample size calculation in thorough QT/QTc studies. *J Biopharm Stat* 2008;18:451–467.
7. Garnett CE, Beasley N, Bhattaram VA, et al. Concentration-QT relationships play a key role in the evaluation of proarrhythmic risk during regulatory review. *J Clin Pharmacol* 2008;48:13–18.
8. Darpo B, Garnett C, Benson C, et al. CSRC White paper: Can the thorough QT/QTc study be replaced by 'early QT assessment' in routine clinical pharmacology studies?—Scientific update and a research proposal for a path forward. *Am Heart J* 2014;168:262–272.
9. Darpo B, Garnett C. Early QT assessment How can our confidence in the data be improved? *Br J Clin Pharmacol* 2012;76:642–648.
10. Rohatagi S, Carrothers TJ, Kuwabara-Wagg J, et al. Is a thorough QTc study necessary? The role of modeling and simulation in evaluating the QTc prolongation potential of drugs. *J Clin Pharmacol* 2009;49:1284–1296.
11. Shah RR, Morganroth J. Early investigation of QTc liability: The role of multiple ascending dose (MAD) study. *Drug Saf* 2012;35:695–709.
12. Darpo B, Sarapa N, Garnett C, et al. The IQ-CSRC prospective clinical Phase 1 study: "Can early QT assessment using exposure response analysis replace the thorough QT study?". *Ann Noninvasive Electrocardiol* 2014;19:70–81.
13. Fardipour P, Littman G, Burns DD, et al. Planning and executing response-adaptive learn-phase clinical trials: 1. The process. *Drug Infor J* 2009;43:713–723.
14. Hosmane B, Locke C. A simulation study of power in thorough QT/QTc studies and a normal approximation for planning purposes. *Drug Infor J* 2005;39:447–455.
15. Shumaker RC, Zhou M, Ren M, et al. Effect of lenvatinib (E7080) on the QTc interval: Results from a thorough QT study in healthy volunteers. *Cancer Chemother Pharmacol* 2014; Volume 7: page 1109–1117.
16. Florian JA, Tornøe CW, Brundage R, et al. Population pharmacokinetic and concentration-QTc models for moxifloxacin: Pooled analysis of 20 thorough QT studies. *J Clin Pharmacol* 2011;51:1152–1162.
17. Darpo B, Fossa AA, Couderc JP, et al. Improving the precision of QT measurements. *Cardiol J* 2011;18:401–410.
18. Darpo B, Ferber G, Zhou M, et al. Lomitapide at supratherapeutic plasma levels does not prolong the QTc interval—results from a TQT study with moxifloxacin and ketoconazole. *Ann Noninvasive Electrocardiol* 2013;18:577–589.
19. Zhu H, Wang Y, Gobburu JV, et al. Considerations for clinical trial design and data analyses of thorough QT studies using drug-drug interaction. *J Clin Pharmacol* 2010;50:1106–1111.
20. Ferber G, Wang D, Taubel J. Concentration-effect modelling based on change from baseline to assess the prolonging effect of drugs on QTc together with an estimate of the circadian time course. *J Clin Pharmacol* 2014 Jun 19 [Epub ahead of print].
21. Pinheiro J, Bates D, DebRoy S, Sarkar D, and the R Core team. *Nlme: Linear and Nonlinear Mixed Effects Models*. R package version 3.1–117. Available at: <http://CRAN.R-project.org/package=nlme>. Accessed October 16, 2014.
22. Bedaquiline. Anti-infective Drugs Advisory committee meeting November 28, 2012. FDA Briefing document. Available at: <http://www.fda.gov/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/Anti-Infective-DrugsAdvisoryCommittee/ucm329257.htm>. Accessed October 16, 2014.
23. Vandetanib. (Caprelsa) Clinical Pharmacology review April 2011. Available at: http://www.accessdata.fda.gov/drugsatfda_docs/nda/2011/022405Orig1s000ClinPharmR.pdf. Accessed October 16, 2014.
24. Zhang J. FDA experience with novel QT study designs. DIA Cardiovascular safety meeting Washington DC, April 2012.
25. International Consortium for Innovation and Quality in Pharmaceutical Development. Available at: <http://www.iqconsortium.com/>.
26. Cardiac Safety Research Consortium. Available at: <https://www.cardiac-safety.org/>. Accessed October 2014.