Journal of Pharmacokinetics and Pharmacodynamics (2019) 46:617–626
https://doi.org/10.1007/s10928-019-09662-3

ORIGINAL PAPER

Cardiac risk assessment based on early Phase I data and PK-QTc analysis is concordant with the outcome of thorough QTc trials: an assessment based on eleven drug candidates

Puneet Gaitonde[1] · Yeamin Huh[2] · Borje Darpo[3] · Georg Ferber[4] · Günter Heimann[5] · James Li[6] · Kaifeng Lu[7] · Bernard Sebastien[8] · Kuenhi Tsai[9] · Steve Riley[1]

Received: 16 May 2018 / Accepted: 17 October 2019 / Published online: 30 October 2019

Abstract

Cardiac safety assessment is a key regulatory requirement for almost all new drugs. Until recently, one evaluation aspect was via a specifically designated, expensive, and resource intensive thorough QTc study, and a by-time-point analysis using an intersection–union test (IUT). ICH E14 Q&A (R3) (http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E14/E14_Q_As_R3__Step4.pdf) allows for analysis of the PK-QTc relationship using early Phase I data to assess QTc liability. In this paper, we compared the cardiac risk assessment based on the early Phase I analysis with that from a thorough QTc study across eleven drug candidate programs, and demonstrate that the conclusions are largely the same. The early Phase I analysis is based upon a linear mixed effect model with known covariance structure (Dosne et al. in Stat Med 36(24):3844–3857, 2017). The treatment effect was evaluated at the supratherapeutic C_{max} as observed in the thorough QTc study using a non-parametric bootstrap analysis to generate 90% confidence intervals for the treatment effect, and implementation of the standardized methodology in R and SAS software yielded consistent results. The risk assessment based on the concentration–response analysis on the early Phase I data was concordant with that based on the standard analysis of the thorough QTc study for nine out of the eleven drug candidates. This retrospective analysis is consistent with and supportive of the conclusion of a previous prospective analysis by Darpo et al. (Clin Pharmacol Ther 97(4):326–335, 2015) to evaluate whether C-QTc analysis can detect QTc effects in a small study with healthy subjects.

Keywords Concentration-QTc · Cardiac risk assessment · Linear mixed effects model · Bootstrap confidence intervals

Introduction

Cardiac risk assessment for most drug development programs has relied heavily on the outcome of a thorough QTc study ever since establishment of the ICH E14 guideline in 2005 which gives guidance on the design and analysis of such trials [2]. The recommended analysis (the

A part of the current work was presented as a poster at the 2016 American Conference on Pharmacometrics [1].

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10928-019-09662-3) contains supplementary material, which is available to authorized users.

✉ Steve Riley
Steve.Riley@Pfizer.com

[1] Clinical Pharmacology, Global Product Development, Pfizer Inc, Groton, CT 06340, USA

[2] Pharmacometrics, Global Product Development, Pfizer Inc, Groton, CT 06340, USA

[3] ERT, Rochester, NY, USA

[4] Statistik Georg Ferber GmbH, Riehen, Switzerland

[5] Biostatistics & Pharmacometrics, Novartis Pharma AG, Basel, Switzerland

[6] Former Quantitative Clinical Pharmacology, AstraZeneca, Cambridge, UK

[7] Statistical Science, Allergan plc, Madison, NJ, USA

[8] Clinical Trial Simulation, Sanofi, Chilly-Mazarin, France

[9] Biometrics, Teclison Limited, Montclair, NJ 07042, USA

intersection–union test; IUT) is conducted by time point. However, the operating characteristics of the recommended analysis have often been challenged [3, 4], because of a high false positive rate, large sample size required, and associated costs [5]. Several authors [6–9] have suggested using concentration–response analysis instead of IUT to increase the efficiency of the analysis. An additional advantage of a concentration–response analysis is that it allows for cardiac risk assessment across pooled data from different early Phase I studies (i.e. single and multiple dose studies) [7, 9, 10] and also estimates risk at doses not studied, whereas the classical intersection–union test cannot be easily extended to these situations.

In order to replace a thorough QTc study (not to be considered as a Phase I study in the context of this manuscript) by an analysis of early Phase I data, one needs to ensure the quality of the ECG measurements [11]. In this context one may ask the question as to whether the two approaches are comparable. A first attempt to answer this question was the collaborative study between the Consortium for Innovation and Quality in Pharmaceutical Development (IQ) and the Cardiac Safety Research Consortium (CSRC) [6]. This prospective study included a range of compounds with known cardiac risk (as defined by a positive thorough QTc study) and mimicked the design of Phase I studies with small sample sizes. Pharmacokinetic (PK) concentrations and QTc intervals were assessed, and concentration–response analyses were conducted for these compounds, confirming the known effects from the historical thorough QTc studies.

The results of these assessments with model-based concentration–response analyses of early Phase I data have been accepted when a positive signal is demonstrated [12]. Following the results reported in the IQ-CSRC study, the ICH E14 Q&A Revision 3 was published [11], broadening the scope of concentration–response analyses based on early Phase I data, and defining its role in regulatory decision-making.

The objective of the current research was to provide further evidence that the cardiac risk assessment based on concentration–response analyses of early Phase I data is concordant to the risk assessment based upon a thorough QTc trial. Different to the approach taken by Darpo et al. [6], our research is purely retrospective in nature. Sponsors from within the Optimizing QT (OQT) consortium were invited to volunteer datasets comprising of a thorough QTc study and associated single and/or multiple ascending dose studies with ECG assessments according to the required standards for thorough QTc studies. All studies completed between years 2002 and 2014, and selected per pre-defined dataset selection criteria, were included in this assessment. In total, eleven drug candidate programs were included. The early Phase I concentration–response analysis was compared to the corresponding concentration–response analysis based on thorough QTc studies, as well as to the original intersection–union test. To provide comparable results and to avoid bias due to different analysis models, all early Phase I and thorough QTc studies were re-analyzed according to a pre-specified "standard" analysis. Criteria to assess concordance or discordance between the early Phase 1 analysis and the thorough QTc analysis were pre-specified as well. The corresponding R and SAS analysis code is provided in the supplementary material to this paper.

Methods

Data

Datasets associated with twenty-eight Pfizer Inc. drug candidate programs were initially reviewed. Data from thorough QTc[1] studies and corresponding placebo-controlled, parallel group or cross-over early Phase I studies in healthy volunteers under fasted conditions were evaluated against pre-defined selection criteria that included, (a) drug concentrations for Phase I studies and for the thorough QTc study had to be measured at the same nominal time points when ECG data were collected, (b) plasma drug concentration range observed in the early Phase I studies covered the observed concentration range in the corresponding thorough QTc study, (c) PK sampling must adequately characterize the respective maximum peak drug concentration (C_{max}) in both early Phase I and thorough QTc studies, and (d) replicate ECGs were measured at baseline and at all post-dose time points to the standards which are applied in thorough QTc studies. In this retrospective analysis, these criteria were assessed for each program such that, in general, the study populations, inclusion/exclusion criteria, and ECG acquisition rules were comparable between the Phase I and the thorough QTc studies for a given drug candidate to enable pooling of the Phase I data and ultimately, comparability of the results from the early Phase I studies with the corresponding thorough QTc studies. Based on these criteria, eleven drug candidates with studies completed between years 2002 and 2014 were identified, including three positive (Candidates # 4, 6, and 7) and eight negative (Candidates # 1, 2, 3, 5, 8, 9, 10 and 11) drug candidates. (The classification 'positive' and 'negative' refers to the intersection–union test as defined

[1] Some of the selected studies were completed prior to the publication of ICH E14 guidance in 2005. Therefore, a formal thorough QTc designation is not applicable to them, even though they were designed and conducted similarly to a thorough QTc study. In the context of the current research, they are referred to as thorough QTc studies.

by ICH E14.) Although only single ascending dose data were available for drug candidate #2, it met the pre-defined selection criteria and was therefore included in the analysis. The drug candidate programs were diverse when compared for their respective sample size, year of study conduct, blood sampling and ECG assessment density (Table 1).

Dataset preparation

Dataset handling and processing was done in R [13] (Version 3.2.3; see the Supplementary Material SM1 and SM2). For the analysis, data associated with the positive control (moxifloxacin treatment) from the thorough QTc studies, and records containing either missing PK or ECG observations at a given nominal time point were excluded. The PK concentrations for the placebo group were set to zero. After data processing, the new dataset contained seven columns namely; baseline adjusted QTc interval, drug concentration (PK), protocol ID, nominal time post-first dose, unique Subject identifier (or unique Subject:Period identifier in case of a cross-over trial), number of baseline measurements by subject, and dose.

Data analysis

Concentration-QTc analysis

The baseline-adjusted QTc interval (ΔQTc) for each subject within each clinical study was calculated by subtracting that subject's pre-dose mean baseline QTc from the post-dose QTc values at each time point and used as the response variable. Fridericia's correction of the QT interval [14] was used for all analyses, and the analysis was performed separately for each of the eleven drug development programs. Mean values of duplicate or triplicate QT interval collections were obtained, so that a single ΔQTc observation was available for analysis at each nominal time point. Baseline QTc values were determined as the mean of ECG measurements at all time points before dosing. In the studies which were part of this investigation, there was either one single pre-dose time point (0 h), or three time points (e.g., − 1.5, − 1, − 0.5 h) pre-dose.

A linear mixed effects model (henceforth referred to as 'C-QTc model') was used to analyze the early Phase I C-QTc data (namely single- and/or multiple-dose studies), as well as, the corresponding thorough QTc data with slight modification [15]. More precisely, the model

$$\Delta QTc_{klt} = p_t + \vartheta C_{klt} + \eta_{kl} + \varepsilon_{klt}$$

was used for the C-QT analysis. This model includes a fixed effect (p_t) for the circadian rhythm, the drug concentration (C_{klt}) as a covariate with a fixed effect slope (ϑ), and a random subject effect (η_{kl}). The additive random noise (ε_{klt}) with common variance σ^2 across all time points and subjects is also included. The indices l, k, and t stand for dose, subject identifier, and nominal time post-first dose, respectively. For pooled analysis of early Phase I studies, the model was modified to

$$\Delta QTc_{sklt} = \alpha_s + p_{st} + \vartheta C_{sklt} + \eta_{skl} + \varepsilon_{sklt}$$

Table 1 Early Phase I (SAD/MAD) and thorough QTc (tQTc) study overview and intersection–union test outcome for the eleven drug candidates

Candidate #	Year	tQTc Age range (years)	SAD Age range (years)	MAD Age range (years)	# of subjects		IUT result
					tQTc (M:F)	Phase I (M:F)	
1	2002–2008	21–51	19–45	23–63	60 (32:28)	154 (129:25)	Negative
2	2002–2008	21–48	19–53	–	36 (24:12)	28 (24:4)	Negative
3	2003–2008	21–55	19–53	20–55	32 (14:18)	44 (48:2)[a]	Negative
4	2003–2004	18–53	21–53	21–53	66 (47:19)	83 (80:3)	Positive
5	2006–2010	24–55	21–42	19–53	48 (48:0)	56 (56:0)	Negative
6	2004–2005	18–55	21–42	20–45	35 (30:5)	54 (54:0)	Positive
7	2006–2011	22–53	20–54	21–40	41 (44:0)[a]	52 (52:0)	Positive
8	2003–2003	19–44	21–45	18–41	61 (30:31)	96 (96:0)	Negative
9	2002–2004	19–55	18–52	20–54	60 (37:23)	95 (92:3)	Negative
10	2010–2014	18–54	20–49	23–54	42 (20:22)	64 (63:1)	Negative
11	2002–2005	18–54	18–42	23–45	70 (0:71)[a]	115 (115:0)	Negative

M male, *F* female

[a]Enrolled/assigned to treatment

and additionally includes a fixed study effect (α_s). The index s stands for the different studies, and the diurnal variation is allowed to vary with study to account for differences in design (for example meal times) which may impact the diurnal variation. In case of a crossover study, the random effect for subject was replaced by a random subject-period interaction. The variance of the subject-specific random effects was modeled to be a fraction of the residual noise (σ^2 in the case of one pre-baseline time-point and $\sigma^2/3$ in the case of three pre-baseline time-points). A justification for this C-QTc model and a detailed description of the corresponding analysis has been provided by Dosne et al. [15] (see also Supplementary Material SM1, SM2 and SM3 for the corresponding R and SAS analysis code).

The model for the intersection–union test as per ICH E14 was also a variation of the C-QTc model described above. The covariate "concentration" was replaced by a "treatment-time" interaction (γ_{lt}) as a fixed factor), and appropriate contrasts for each time point were used to conduct the intersection–union test.

All three models (i.e. C-QTc model for early Phase I and thorough QT, respectively, and ICH E14 model) use the same correlation structure which is governed by the random subject effects. For crossover studies, this was replaced by a random subject-period interaction term. This implies that observations from one subject, but in different periods, are regarded as independent or uncorrelated, and allows one to pool data from parallel group studies with data from crossover studies within a particular drug candidate program. For a justification of the model, we refer to [15] again, and to the supplementary material. Data analysis was pre-specified and documented in an analysis plan. Several checks were done to assess the appropriateness of the primary analysis. The potential for nonlinearity or hysteresis in the exposure–response curve was investigated graphically using QTcF vs drug concentration plots, and mean drug concentration and QTcF vs time plots, respectively. Additionally, standard diagnostic plots were examined to evaluate model fits to the data and ensure no gross violation of assumptions were evident.

Model-predicted change from baseline QTc using bootstrap approach

The primary parameter of interest when evaluating the concentration–response analysis is ϑC_{max}, the expected QTc interval prolongation at the expected maximum drug concentration. The slope, ϑ, was estimated using the linear model from the previous section, and the expected C_{max} (in the context of the current retrospective analysis) was the geometric mean of the individual C_{max} observations of the corresponding thorough QT study. A confidence interval around this parameter was obtained using a non-parametric approach, generating 1000 bootstrap copies (same initial seed number for each drug candidate program) of the initial dataset. To accommodate different study designs and respect correlation within subjects, re-sampling was stratified by study and dose by randomly drawing the entire data (PK and QTc) of a subject with replacement from the observed subjects of a given study. The number of subjects across the studies and dose groups permitted minimal duplication of individuals within any given bootstrap replicate, thereby minimizing bias in the derived confidence intervals. For each of the 1000 bootstrap copies of the original dataset, a bootstrap copy of the parameter estimate for ϑC_{max} was determined and used the 5th and 95th percentile of these bootstrap estimates to define the confidence interval.

Assessment of model predictive performance

The following three approaches were employed to compare the analysis based on early Phase I data with that based on thorough QTc data, although comparison of the C-QTc model predictions between the Phase 1 and thorough QTc studies ("Comparison of C-QTc model predicted effect at the supratherapeutic maximal concentration between early Phase I and thorough QTc data") were of primary interest.

Comparison of C-QTc model predicted effect at the supratherapeutic maximal concentration between early Phase I and thorough QTc data The two-sided 90% bootstrap confidence intervals for the key parameter, ϑC_{max} (slope multiplied by the expected C_{max} at the supratherapeutic dose) were graphically compared. Both the analysis based on the thorough QTc data as well as the analysis based upon the early Phase I data were considered to be positive if the corresponding upper 90% CI limit was above the 10 ms regulatory threshold. The analysis was considered negative if this limit was below 10 ms. The goal of this comparison was to demonstrate concordant conclusions of cardiac risk between analysis of early Phase I data and thorough QTc data.

Comparison of the C-QTc model predicted effect at the supratherapeutic maximal concentration from the early Phase I and the ICH E14 intersection–union test analysis based on the thorough QTc data The two-sided 90% bootstrap confidence interval for the key parameter, ϑC_{max} (slope multiplied by expected C_{max} at the supratherapeutic dose) obtained from the analysis of early Phase I studies was compared with the corresponding drug candidate's intersection–union test result from the thorough QTc study. Analysis of the early Phase I data was performed in the

same manner as described under "Comparison of C-QTc model predicted effect at the supratherapeutic maximal concentration between early Phase I and thorough QTc data" section. The analysis based on the intersection–union test based on thorough QTc data was considered positive if there was at least one time point with upper confidence limit above 10 ms. The analysis based upon the early Phase I data was considered to be positive if the corresponding upper 90% confidence limit for ϑC_{max} was above the 10 ms regulatory threshold. The goal of this comparison was to demonstrate concordant conclusions of cardiac risk between concentration–response analysis and the intersection–union test from thorough QTc studies. Given that the intersection–union test result only considers the largest time-matched ΔΔQTc effect and not necessarily the effect at C_{max}, the concentration at each time point may only be the same or lower than the supratherapeutic C_{max} value that is used in the C-QTc model-based prediction of the early Phase I data.

Comparison of C-QTc model slope estimates between early Phase I and thorough QTc Study The two-sided 95% confidence interval for the slope (ϑ) parameter was obtained from the model described in "Concentration-QTc analysis", and graphically compared. The parameter estimate from the early Phase I dataset was to be contained within the confidence interval from the thorough QTc study in order for the early Phase I study analysis to be regarded as 'predictive' of the thorough QTc study. The variability of the PK concentrations was not taken into consideration in this assessment.

Software

Data analysis was conducted in both; R [13] (Version 3.2.3), and SAS software [16] (Version 9.2) using Restricted Maximum Likelihood (REML) estimation method. Model diagnostics and graphical outputs were generated in R using appropriate R packages.

Results

Concentration-QTc relationship

The graphical hysteresis assessment provided no consistent evidence of delayed drug effect for the eleven drug candidates (representative plot shown in Fig. 1a). The linear model was considered to be adequate in all cases. Model diagnostic plots of standardized residuals (representative QQ plot shown in Fig. 1b; representative histogram plot shown in Fig. 1c) suggested no substantial deviations from normality. Hence the primary analysis is provided for all eleven drugs as the sole analysis.

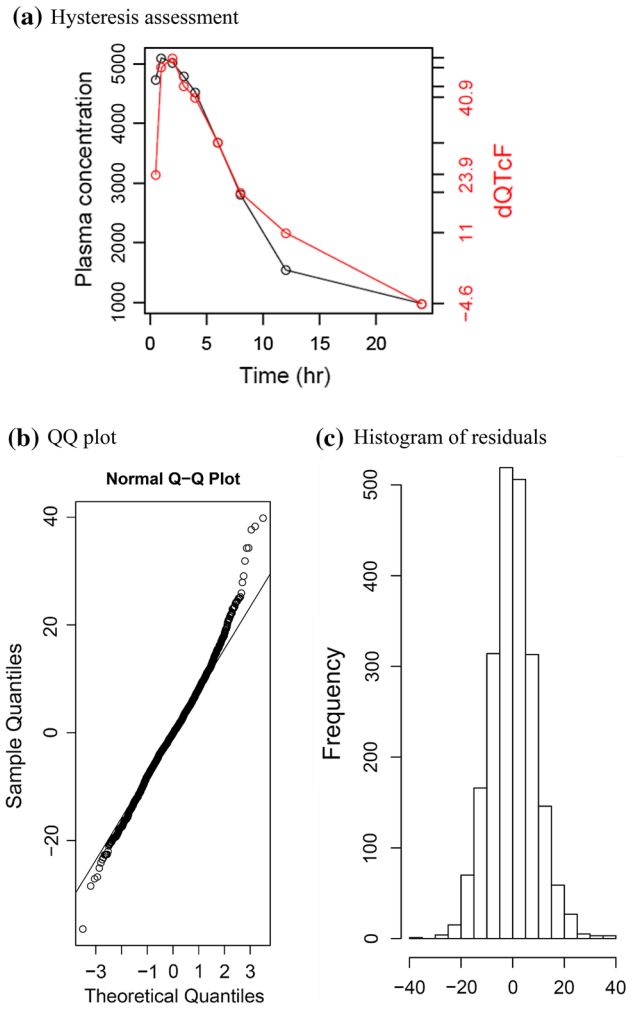

Fig. 1 Representative plots to assess data suitability (**a**) and model fit diagnostics (**b**, **c**) from R analysis

The implementation of the analyses in R and SAS resulted in nearly identical point estimates and confidence intervals based on standard errors, bootstrap estimates for all eleven Phase I drug candidates, as well as for the corresponding thorough QTc studies (see Supplementary Material, Table 1). Minor differences were noted in the confidence intervals between R and SAS software, which were attributed to the respective software program's default degrees of freedom calculation.

Comparison of C-QTc model predicted effect at the supratherapeutic maximal concentration between early Phase I and thorough QTc data

Figure 2 presents the parameter estimate (in ms) and the two-sided 90% bootstrap confidence intervals for the key parameter ϑC_{max}. The confidence intervals based upon

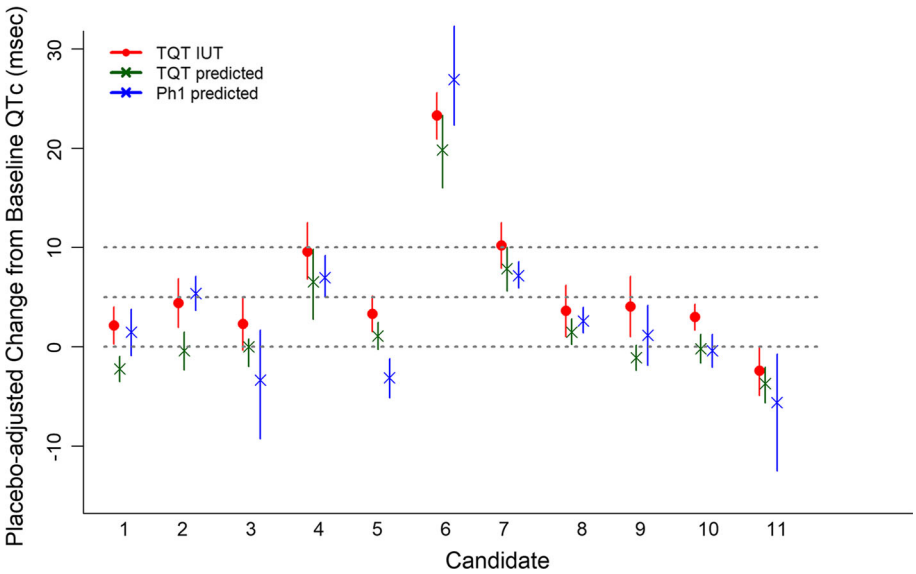

Fig. 2 Comparison of R model results from "Comparison of C-QTc model predicted effect at the supratherapeutic maximal concentration between early Phase I and thorough QTc data" and "Comparison of the C-QTc model predicted effect at the supratherapeutic maximal concentration from the early Phase I and the ICH E14 intersection-union test analysis based on the thorough QTc data". C-QTc model bootstrap estimates for ΔΔQTcF (y-axis) of early Phase I (blue), thorough QTc study (green), and intersection–union test results (red) for each of the eleven drug candidates (x-axis). The red dots, blue and green crosses represent the point estimates, and the ends of the lines indicate the 90% confidence interval limits. The three dashed horizontal lines (bottom to top) indicate 0, 5, and 10 ms (Color figure online)

early Phase I data (blue) and corresponding thorough QTc data (green) provide concordant conclusions for all eleven drug candidates. It is noteworthy that for the positive drug candidates #4 and 7 (as deemed by the respective ICH E14 intersection–union test outcomes), results were negative when based on C-QTc model analyses in both the thorough QTc and early Phase I studies (see Supplementary Material, Table 1).

Comparison of the C-QTc model predicted effect in Phase I studies and the ICH E14 by-time point effect in the thorough QTc data

Figure 2 also presents the parameter estimate (in ms) and the two-sided 90% bootstrap confidence intervals for the key parameter ϑC_{max} based upon the early Phase I data (blue) and the confidence interval for the time point with the maximal upper limit (red) from the corresponding intersection–union test. Concordant results were obtained for nine out of eleven drug candidates: for the two other drug candidates the intersection–union test results were positive with an estimate $\sim$ 10 ms, whereas the concentration–response analyses were negative from both the thorough QTc and the early Phase I studies.

Comparison of C-QTc model slope estimates between early Phase I and thorough QTc Study

Figure 3 presents the point estimate and two-sided 95% confidence interval for the slope parameter (msec per unit concentration) based on the early Phase I (blue) and thorough QTc (green) data for the eleven drug candidates. The point estimate of the early Phase I slope for four out of the eleven drug candidates (# 4, 7, 8 and 10) was contained within the 95% confidence interval obtained from the thorough QTc study. Furthermore, the 95% confidence interval from the early Phase I and the thorough QTc studies overlapped for seven out of eleven drug candidates (# 3, 4, 7, 8, 9, 10 and 11). In some cases (# 2 and 6) the confidence intervals obtained from the two analyses did not overlap, but this can be attributed to very narrow confidence intervals observed for at least one of the analyses. Looking at the trend in the slope estimates, for drug candidates # 4, 6 and 7, the 95% lower confidence limits were greater than zero (for both the early Phase I and the thorough QTc analyses), indicating a statistically significant concentration–response relationship. For drug candidates # 3, 9 and 10, the 95% confidence intervals for both early Phase I and thorough QTc slopes included zero, whereas for drug candidate # 11, the 95% upper confidence limits for both early Phase I and thorough QTc slopes were negative. Fully discordant results (one confidence interval

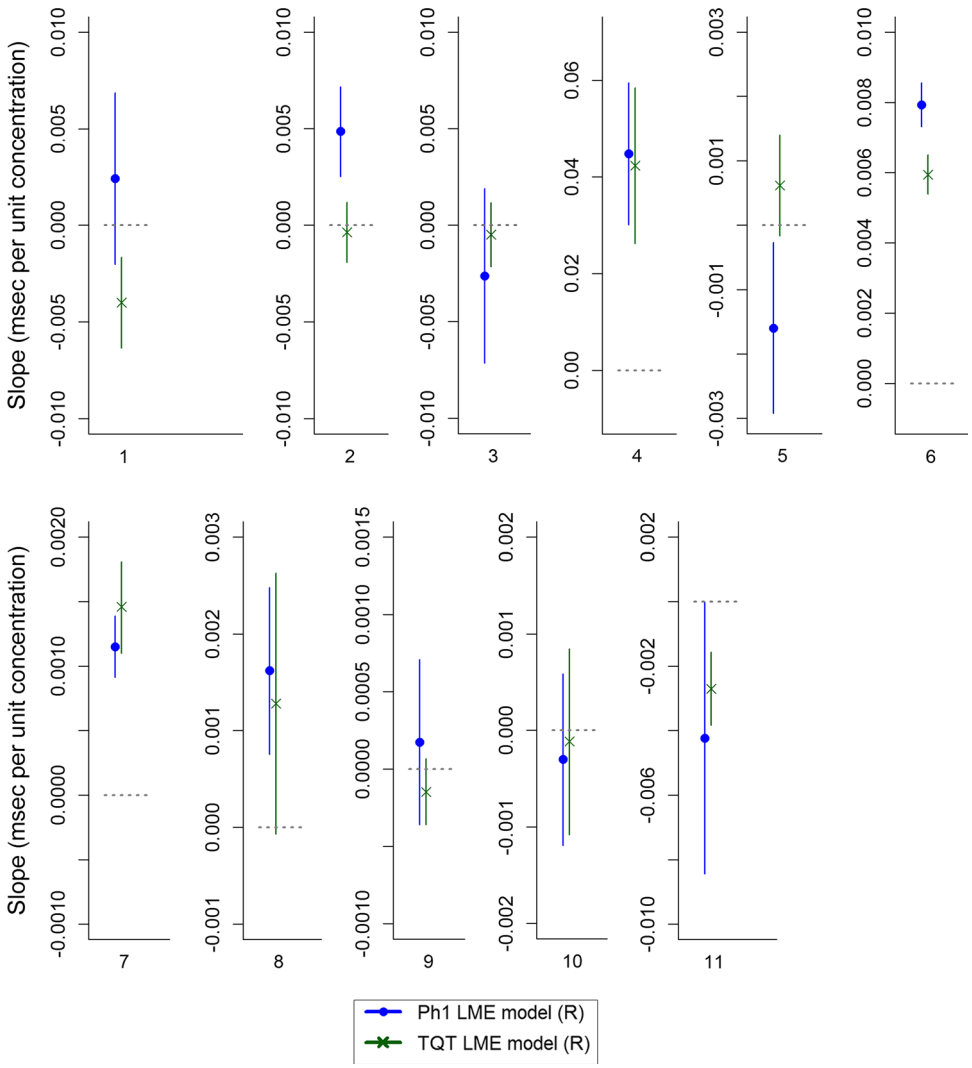

Fig. 3 Comparison of R model results from "Comparison of C-QTc model slope estimates between early Phase I and thorough QTc study". C-QTc model slope estimates (y-axes) of early Phase I (blue) and thorough QTc (green) studies for each of the eleven drug candidates (x-axes). Blue dots and green crosses represent the point estimates, and the ends of the lines indicate the 95% confidence interval limits based on standard error. Dashed horizontal line indicates a zero slope (Color figure online)

completely above zero, the other completely below) were not found for any of the eleven drug candidates.

Research limitations

This was a retrospective analysis of studies conducted between 2002 and 2014. The authors did their due diligence to ensure reasonably similar data quality across drug candidates was maintained during analysis, but some unidentified differences may exist in the data which may or may not be influential in the results.

Discussion

Analyses of the eleven drug candidates demonstrated that positive or negative results of thorough QTc studies as per ICH E14 guidance were adequately predicted by concentration–response analyses based on early Phase I data. In nine out of eleven cases, a concordant cardiac risk assessment was observed. In the two discordant cases, the point estimate of the predicted QTc effect (ϑC_{max}) at the expected C_{max} of the supratherapeutic dose was > 5 ms, i.e. at a level of regulatory concern as per ICH E14. Hence, a potential cardiac risk would not have been overlooked for these two cases. Moreover, a similar discordance would have been observed between the primary analysis (based on the intersection–union test) and the secondary analysis (based on the concentration–response model) within the thorough QTc study itself. It is also to be noted here that the largest upper bound of the intersection–union test in the two discordant cases was > 10 ms, while the upper confidence limit of C-QTc for early Phase I and thorough QTc datasets were marginally below 10 ms. This could be attributed to the reported positive bias inherent in the intersection–union test analysis [17]. Additionally, the potential for a high percentage of false positives with the

intersection–union test has been reported [4]. Even in terms of the conservative ICH E14 criterion, all of the compounds with negative thorough QTc study results were correctly predicted as negative based on the early Phase I data. Therefore, our findings suggest that the lack of a positive cardiac risk based on concentration–response analysis of early Phase I data would be sufficient to waive the need for thorough QTc study. This is also consistent with a previous report [12].

Cases 1, 2, 5, and 6 demonstrate notable differences in predicted ΔΔQTc between the early Phase 1 and TQT studies based on the C-QTc model. Closer examination of these studies identified that the TQT studies all had triplicate ECGs and baseline defined as measurements at time = 0 just before dosing. The SAD/MAD studies in these cases had varying definitions of baseline, e.g., time-matched ECGs vs baseline at time = 0, varying number of replicate ECGs at each time point, different durations of PK-ECG sampling, etc., that may have contributed to the somewhat discrepant results. However, it is worth noting that the inferences drawn in these 4 cases are the same whether considering the early Phase 1 or TQT-based results.

It is also important to make a distinction here, that the study by Darpo et al. [12] was prospective in nature, had clear a priori objectives, and was planned and executed to the standard of a thorough QTc study. In our case examples, the early Phase I studies were not designed for thorough QTc waiver possibility, yet the retrospective analysis using the proposed methodology provides evidence that early Phase 1 analysis could give confidence in predicting thorough QTc results.

At the beginning of this cross-company effort, it was decided to use one standard model across all drug development programs for the purpose of the standardized analysis reported in this paper. This meant that the data were reanalyzed with the pre-specified model, in order to ensure comparability across programs.

The selected model (see "Model-predicted change from baseline QTc using bootstrap approach") is the same model used by Dosne et al. [15] and explained in their supplementary material. Dosne et al. assumed a joint model for pre- and post-baseline QTc observations, from which they derive the model for the primary endpoint change from mean baseline. Many of the original random effects cancel out, resulting in the covariance structure described in "Model-predicted change from baseline QTc using bootstrap approach". In addition, for crossover studies the between-period correlation vanishes, so that the data can be treated as if they came from a parallel group design. This allows one to pool across parallel group and crossover trials within one drug development program. Assessments based on the observed data showed that the model was reasonable.

However, using a standard model for the purpose of this paper should not be construed that we are proposing this model as a cross-company standard. We recognize that other models are equally as justifiable as the one employed here.

One such model was proposed in a recent paper by Garnett et al. [18]. Common to both approaches is that they include a pre-specified model. Implementing a pre-specified model alleviates concerns about data-driven decisions in the modeling process influencing inferences made from the results. Common to both pre-specified models are the fixed factor for time, use of concentration as a covariate with a fixed effect slope, and a random subject effect. Note that the notation for the factor time in the Garnett et al. [18] paper is somewhat misleading, as it appears to be a covariate ($\theta_3 time_j$), This is a notational issue, and time is not meant to be a covariate (personal communication). In both papers, the primary inference (tests or confidence intervals) relates to the fixed effect slope parameter.

However, the model by Garnett et al. [18] is more complicated as compared to the one used in this paper and contains additional parameters. For example, it contains a baseline QTc interval as a covariate, Moreover, it includes a subject-specific random slope. Such a random effect leads to a complex covariance structure of the data points, which depends on the concentrations, and one may question whether this is a plausible assumption. Even if it was, the primary inference for QT analyses is about the fixed slope parameter and while ignoring the corresponding random effect may lead to a slightly less efficient estimator, it is still unbiased.

These differences can be viewed as a matter of preference. A scientifically debatable point is the treatment-specific intercept, which is part of the model by Garnett et al. [18], as it implies a discontinuous concentration–response curve. In any case, the model proposed by Garnett et al. [18] warrants further debate, which is beyond the scope of this paper.

Pooling data across Phase I studies can provide a sufficient basis for establishing the concentration–response relationship between PK and the QTc interval [19] by increasing the available sample size and widening the exposure range [18]. The early Phase I studies provide a unique opportunity to evaluate the effect of compounds on the QTc interval over a wide range of concentrations under a controlled setting [8]. Employing QT data from early Phase I studies to serve as a substitute for the thorough QTc study requires careful design, planning, and data measurement quality and collection [18]. Analysis of early Phase I data can typically be undertaken immediately after the study is completed, or postponed to a time when

sufficient data exist to better define the appropriate supratherapeutic concentration. To enable decision-making on the need for intensive monitoring in Phase III, data must be analyzed well in advance of the confirmatory, late stage trials. At that time, sufficient knowledge about the potential therapeutic dose, as well as, intrinsic and extrinsic factors that may influence the PK of the drug candidate would also have been obtained, which are critical factors in determining if the data are sufficient to waive the request for a thorough QTc study. Early Phase I studies generally cover a wide dose (and exposure) range beyond what is usually explored in later phase studies. Having such an adequately wide range helps to avoid extrapolation outside of the observed concentration range when making a prediction at C_{max} of the supratherapeutic dose, when that dose is identified during development, with sufficient confidence.

Careful consideration of drug- and study-specific characteristics must be evaluated prior to a decision to pool data across studies for concentration–response analysis. Some of such characteristic factors, e.g., fed versus fasted studies, drug candidates with active metabolites, and combination therapies, have been identified in the white paper as complex cases warranting special consideration [18].

In conclusion, we propose that implementation of a model-based concentration–response analysis of early Phase I data can provide a robust assessment of QT prolongation potential and inform the need for intensive ECG monitoring during late-phase studies.

Acknowledgements The authors would like to acknowledge the Optimizing QT (OQT) group members (Anne Chain, Arne Ring, Atsunori Kaibara, Barry Koplowitz, Bela Patel, Bob Kringle, Charles Benson, Christine Garnett, Christoffer Tornoe, Cindy Green, Donna Kowalski, Geraldine Ferron, Hao Zhu, Jaya Natarajan, Jiang Liu, Jim Keirns, Jo Boni, Kathy Reyderman, Lars Johannesen, Luana Pesco Koplowitz, Meijian Zhou, Michelle Green, Pete Bonate, Rik de Greef, TJ Carrothers, Yaning Wang, Ying Cao) who have dutifully participated in the teleconference meetings and provided valuable intellectual and scientific contributions over the years.

Disclosure PG, YH, and SR are employees and stockholders of Pfizer Inc. BS is employee and stock holder of Sanofi. GH is employee and stockholder of Novartis AG. GF is a stockholder of Novartis AG.

References

1. Gaitonde P, Huh Y, Heimann G, Li J, Lu K, Benson C, Darpo B, Ferber G, Keirns J, Sebastien B, Tsai K, Wang Y, Zhou M, Riley S (2016) Abstracts accepted for American Conference on Pharmacometrics 2016 (ACoP7). J Pharmacokinet Pharmacodyn 43(1):76. https://doi.org/10.1007/s10928-016-9485-x
2. Food and Drug Administration (2005) 70 FR 61134—International Conference on Harmonisation; Guidance on E14 Clinical Evaluation of QT/QTc Interval Prolongation and Proarrhythmic Potential for Non-Antiarrhythmic Drugs. Office of the Federal Register, National Archives and Records Administration. https://www.gpo.gov/fdsys/granule/FR-2005-10-20/05-20971
3. Chapel S, Hutmacher MM, Bockbrader H, de Greef R, Lalonde RL (2011) Comparison of QTc data analysis methods recommended by the ICH E14 guidance and exposure-response analysis: case study of a thorough QT study of asenapine. Clin Pharmacol Ther 89(1):75–80. https://doi.org/10.1038/clpt.2010.220
4. Hutmacher MM, Chapel S, Agin MA, Fleishaker JC, Lalonde RL (2008) Performance characteristics for some typical QT study designs under the ICH E-14 guidance. J Clin Pharmacol 48(2):215–224. https://doi.org/10.1177/0091270007311921
5. Bouvy JC, Koopmanschap MA, Shah RR, Schellekens H (2012) The cost-effectiveness of drug regulation: the example of thorough QT/QTc studies. Clin Pharmacol Ther 91(2):281–288. https://doi.org/10.1038/clpt.2011.224
6. Darpo B, Sarapa N, Garnett C, Benson C, Dota C, Ferber G, Jarugula V, Johannesen L, Keirns J, Krudys K, Ortemann-Renon C, Riley S, Rogers-Subramaniam D, Stockbridge N (2014) The IQ-CSRC prospective clinical Phase 1 study: "Can early QT assessment using exposure response analysis replace the thorough QT study?". Annals of noninvasive electrocardiology : the official journal of the International Society for Holter and Noninvasive Electrocardiology, Inc 19 (1):70-81. https://doi.org/10.1111/anec.12128
7. Rohatagi S, Carrothers TJ, Kuwabara-Wagg J, Khariton T (2009) Is a thorough QTc study necessary? The role of modeling and simulation in evaluating the QTc prolongation potential of drugs. J Clin Pharmacol 49(11):1284–1296. https://doi.org/10.1177/0091270009341184
8. Russell T, Riley SP, Cook JA, Lalonde RL (2008) A perspective on the use of concentration-QT modeling in drug development. J Clin Pharmacol 48(1):9–12. https://doi.org/10.1177/0091270007311115
9. Shah RR, Morganroth J (2012) Early investigation of QTc liability: the role of multiple ascending dose (MAD) study. Drug Saf 35(9):695–709. https://doi.org/10.2165/11634810-000000000-00000
10. Ferber G, Zhou M, Darpo B (2015) Detection of QTc effects in small studies–implications for replacing the thorough QT study. Annals of Noninvasive Electrocardiol 20(4):368–377. https://doi.org/10.1111/anec.12227
11. ICH E14 Implementation Working Group (2015) ICH E14 guideline: The clinical evaluation of QT/QTc interval prolongation and proarrhythmic potential for non-antiarrhythmic drugs, questions & answers (R3). http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E14/E14_Q_As_R3__Step4.pdf
12. Darpo B, Benson C, Dota C, Ferber G, Garnett C, Green CL, Jarugula V, Johannesen L, Keirns J, Krudys K, Liu J, Ortemann-Renon C, Riley S, Sarapa N, Smith B, Stoltz RR, Zhou M, Stockbridge N (2015) Results from the IQ-CSRC prospective study support replacement of the thorough QT study by QT assessment in the early clinical phase. Clin Pharmacol Ther 97(4):326–335. https://doi.org/10.1002/cpt.60
13. R Development Core Team (2008) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna
14. Fridericia LS (1920) Die Systolendauer im Elektrokardiogramm bei normalen Menschen und bei Herzkranken. Acta Med Scand 53:469–486
15. Dosne AG, Bergstrand M, Karlsson MO, Renard D, Heimann G (2017) Model averaging for robust assessment of QT prolongation by concentration-response analysis. Stat Med 36(24):3844–3857. https://doi.org/10.1002/sim.7395
16. SAS Institute Inc (2012) SAS software. SAS, Cary

17. Darpo B, Garnett C, Benson CT, Keirns J, Leishman D, Malik M, Mehrotra N, Prasad K, Riley S, Rodriguez I, Sager P, Sarapa N, Wallis R (2014) Cardiac Safety Research Consortium: 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 168(3):262–272. https://doi.org/10.1016/j.ahj.2014.06.003
18. Garnett C, Bonate PL, Dang Q, Ferber G, Huang D, Liu J, Mehrotra D, Riley S, Sager P, Tornoe C, Wang Y (2017) Scientific white paper on concentration-QTc modeling. J Pharmacokinet Pharmacodyn. https://doi.org/10.1007/s10928-017-9558-5
19. Murphy PJ, Yasuda S, Nakai K, Yoshinaga T, Hall N, Zhou M, Aluri J, Rege B, Moline M, Ferry J, Darpo B (2017) Concentration-response modeling of ECG data from early-phase clinical studies as an alternative clinical and regulatory approach to assessing QT risk—experience from the development program of lemborexant. J Clin Pharmacol 57(1):96–104. https://doi.org/10.1002/jcph.785