The handle http://hdl.handle.net/1887/19939 holds various files of this Leiden University dissertation.

Author: Chain, Anne S.Y.
Title: Mind the gap: predicting cardiovascular risk during drug development
Date: 2012-10-09
Mind The Gap:
Predicting Cardiovascular Risk During Drug Development-
Summary, Conclusion and Perspectives

CHAPTER 12
12.1 REGULATORY FRAMEWORK AND CHALLENGES IN THE TRANSLATION OF PRODROMIC EFFECTS INTO RISK.

Different intrinsic and extrinsic factors, including genetic susceptibility, are known to prolong QT/QTc-interval in humans. Although QT-interval prolongation is not necessarily harmful, it is currently deemed a biomarker for ventricular tachyarrhythmias, such as torsades de pointes (TdP) and a risk or prognostic factor for sudden cardiac death [1, 2]. Of particular concern to regulatory authorities is the evidence of QTc-interval prolongation for non-antiarrhythmic drugs [3, 4]. Despite continuous efforts in characterising the potential dromotropic effects of new molecules, QTc-interval prolongation has become the second most common reason for post-marketing drug withdrawal. Furthermore, the number of non-antiarrhythmic drugs that carry what has been termed a “QT-liability” continues to increase.

Given the implications of what has become the most recent “pharmacoepidemic”, it is essential to predict the potential for QTc-interval prolongation and to accurately assess its clinical relevance as early as possible to prevent regulatory actions such as denial of approval, delays and severe restrictions on the use of the product [5].

In addition to the development of experimental protocols for the assessment of \textit{in vitro} and \textit{in vivo} cardiovascular safety in preclinical species, a guideline (topic E14) by the International Conference on Harmonisation (ICH) has been used as basis for the clinical evaluation of the proarrhythmic propensity in new compounds [6]. Similar guidelines (S7A and S7B) were also developed for the preclinical studies [7, 8]. In the E14 document, a dedicated trial is recommended to formally dismiss the evidence of a safety signal. The so-called thorough-QT (TQT) study is designed and powered to detect a 5 msec increase and includes a one-sided confidence interval that is less than 10 msec to be deemed negative. While the introduction of appropriate monitoring and evaluation is well intended, the recommendations made are far from perfect (Figure 1).

In this thesis we have attempted to highlight three major flaws in the current approach, which can undermine the evaluation and development of suitable molecules for the treatment of important unmet medical needs. The first aspect refers to \textit{assumptions underlying the translation of pharmacological effects from experimental protocols or clinically
controlled trials to real-life conditions. The impact of varying experimental settings and procedures seems to be overlooked, leading to important gaps in the understanding of how pharmacological properties translate into risk in a strictly quantitative manner. The second aspect pertains to the conceptual definition and biological implications of the differences between association and causation. The choice for the use of QT-interval as a biomarker or proxy for the risk of pro-arrhythmias without accounting for the underlying pharmacokinetic-pharmacodynamic relationships blurs the assessment of causality. Basic signal detection principles appear to be missing in research protocols, including the lack of relative operating curves, consistency of the data and quantitative strength of the association, taking into account the construct validity. The third and possibly most important aspect from a drug development perspective is the lack of a framework that enables effective evidence synthesis for the assessment of risk and benefit-risk ratio. The focus on evidence generation without data integration, including systematic incorporation of prior knowledge leads to less than optimal experimental protocols and inaccurate decision criteria. Consequently, wrong recommendations are likely to arise with regard to the termination of a compound in development. Most importantly, current approaches prevent any comprehensive assessment of what the expected risk is in the target population and how mitigation measures might perform in real life conditions.

In drug development, the ideal situation for the prediction of drug-induced QTc prolongation would be if in vitro experiments showed strong predictive value of the clinical outcome both from a qualitative (QTc prolongator or not) and quantitative (the amount of prolongation per concentration unit) perspective. A qualitative correlation (trying to avoid false positive and false negative results) has been proposed by different authors [9, 10], but has been based on an empirical classification of findings. However, one of the main issues with these qualitative approaches is that the compounds with a small QTc prolonging effect are hard to place and that quantitative approaches that facilitate the integration of evidence obtained during the different stages of development and therapeutic use of drugs are lacking. Although our work has focused on the translation of in vivo data, the availability of predictive measures which take in vitro properties into account remains highly desirable. The possibility to make inferences from in vitro measures would imply the generation of less complex experimental data and decision making at an earlier stage of development. Consequently,
this could represent an opportunity for less attrition post candidate selection.

Figure 1: Clinical trials and in particular the TQT trial design yields estimates of drug effects on a restricted subset of the population. The expectation that such a study will generate accurate evidence of the true safety profile of a new chemical entity in the overall population (with different demographic characteristics, comorbidities and concomitant medication usage) is questionable, especially if one takes into account the lack of assessment of concentration-effect relationship in the analysis of a TQT study.

Instead, we have focused on the subsequent phase of drug development, for which no hurdles should exist with regard to the translational value of data being generated in vivo, in particular in early clinical development, during which pharmacokinetics and intensive cardiovascular monitoring are already common practice. Most importantly, we attempted to identify the translational gaps caused by the differences associated with the controlled conditions used in randomised clinical trials and in particular with regard to the effects of inclusion and exclusion criteria. The premise for this work was that risks are inherent to all drug interventions and as such, decisions about developability and therapeutic use should rely on quantitative assessment of the risk it poses to the target population. Rather than relying on assumptions about the construct validity and generalisability of experimental protocols, we deem more appropriately to rely on inferences about risk by taking into
account the impact of potential differences between experimental protocols and real-life conditions.

Another important point arising in the work proposed here is that **evidence synthesis can be by far more powerful than evidence generation**, as it encompasses prior and existing knowledge. In fact, in conjunction with appropriate clinical and scientific rationale and suitable statistical methodology, evidence synthesis by inferential methods can provide a much more reliable basis for decision making, taking into account practical and ethical constraints associated with the generation of safety data in humans. This issue is particularly relevant for low frequency events as well as when risk results from long term exposure.

Given the aforementioned systematic flaws, the high costs associated with conducting a TQT trial and the need to accurately determine the implications of prodromic effects (if any), it is paramount to explore novel strategies and consider implementation of new processes for the evaluation of cardiovascular safety of non-antiarrhythmic drugs. Most importantly, regulation should reflect scientific understanding. This is clearly not the case for ICH E14, which still relies on scientific and regulatory context that prevailed more than a decade ago, when quantitative pharmacology concepts and pharmacometrics as clinical sciences were evolving. It is imperative that novel, robust methodologies are considered in regulatory decision making and that guidance documents and policies reflect such advancements. By and large, inferential methods by modelling and simulation are under-utilised in safety assessment. Despite the wealth of data generated throughout clinical development and the evidence from historical data of compounds showing similar pharmacological properties, little has been done to ensure that prior knowledge is used in a systematic manner all the way from early clinical trials to post-market safety surveillance.

As shown throughout this thesis, a new paradigm is proposed for safety assessment in which evidence synthesis is favoured, making evidence generation a confirmatory step in the continuum between hypothesis and scientific understanding or knowledge. In this context, inferential methods are central to evidence synthesis and translation of safety findings into hazard or risk. The essence and contribution of our research is the possibility to introduce and incorporate pharmacological concepts into pharmacoepidemiology, bridging the gap between drug development, and
health care and therapeutics. An immediate implication of our approach is the shift from association to causation as the basis for the evaluation of safety signals (Figure 2).

![Figure 2. Inferential methods are essential to establish causality during signal detection. The inclusion of concentration-effect relationships into pharmaco-epidemiological research can provide more accurate, quantitative estimates of drug-induced effects. Different experimental and non-experimental conditions can be evaluated by inference. An important advantage of evaluating risk by inferential methods using modelling and simulation is the possibility to discriminate between drug and systems or patient specific properties.](image)

Based on the premises outlined above, this thesis has provided the basis for reducing attrition during the screening of compounds and improving the efficiency of experimental protocols for the evaluation of drug-induced prodromic effects. In Chapter 1, we have presented an overview of the approaches and techniques currently used for candidate selection and progression of a compound into late clinical development. A critical review is made of the predictive value of the experimental protocols, highlighting the requirements for translation of findings both in terms of experimental procedures as well as in terms of data analysis methodology. In particular it addresses the choices for model parameterisation when making use of inferential methods. This underpins the objectives and the intent of the investigation presented in Chapter 2. The use of hierarchical modelling forms the basis for effective translation of drug-induced effects into risk in real-life conditions. Our work reveals not only the importance of alternative methodology to quantify QT/QTc-interval prolongation as a maker of
cardiovascular safety, but also provides insight into the technical requirements for model parameterisation. A benchmark for the analysis and interpretation of risk during the evaluation of cardiovascular safety was then evaluated which accounts for the underlying concentration-effect relationship. Subsequently, focus is given to the importance of a bridging strategy, which we have named simply as Not-In-Trial Simulations (NITS). The use of NITS is proposed as basis for decision making and risk management. Lastly, different protocol optimisation scenarios are considered for prospective evaluation of novel compounds where one or more of the aforementioned issues are dealt with. Conceptually, our work also included a comprehensive assessment of the specificity and sensitivity of the currently accepted threshold value (i.e., QT-interval prolongation ≥ 10 msec) for defining the clinical relevance of prodromic effects. The predictive value of efforts in early pre-clinical or clinical development relies on the receiver operating characteristics (ROC) of the experimental variables or parameters of interest [11]. Even though we do not present a formal ROC analysis of QTc-interval prolongation, our efforts were aimed at demonstrating that flaws exist in assumptions underlying current policies and guidelines.

12.2 BAYESIAN HIERARCHICAL MODELLING

After a general introduction, we have explored the technical requirements for model parameterisation and defined a new benchmark for the analysis and interpretation of risk during the evaluation of cardiovascular safety. In Chapter 3, it was shown how a Bayesian hierarchical model can be used to describe the relationship between drug concentration and QTc-interval prolongation for \( d,l \)-sotalol, moxifloxacin and grepafloxacin. The use of an integrated modelling approach had also allowed explicit dissection of the drug-induced QTc prolongation, discriminating it from other factors such as variation due to heart rate and circadian rhythm. The drug-specific parameter in the model, slope, did not independently indicate whether risks associated with the use of the compound existed, nor did it indicate if the development of the drug should be stopped. Rather, it provided an unambiguous description of the PKPD relationship, taking into account various sources of variability which enabled a clear interpretation of risk.
From a drug development perspective, this approach is highly relevant, as it allows prospective evaluation of compounds. Given the distinction between drug- and system-specific parameters, drug effects can be characterised even if hysteresis occurs, as in the case of indirect mechanisms or metabolite-induced QTc-interval prolongation. The proposed Bayesian analysis provides clearly interpretable clinical measurements which may help the decision process throughout the development of new compounds. In fact, the posterior (predictive) distribution can be used directly to translate liability to QTc-interval prolongation across varying doses or concentration ranges.

The possibility of extrapolating the effect size across different dose ranges and estimating the associated risk can greatly enhance the decision making process for the regulators, clinicians and drug developers. Furthermore in the investigation the probability associated with an increase in QTc-interval $\geq 10$ msec was provided, given a range of concentrations. This threshold can easily be changed to higher or lower values to accommodate other types of assessments, rendering the model useful for different types of analysis or scenarios, including the comparison of drug effects across species. Most importantly, the results show that accurate judgement of the risk and liability associated with a novel chemical or biological entity cannot be performed without distinguishing drug-specific parameters from system-specific parameters.

12.3 TRANSLATIONAL PHARMACOLOGY

12.3.1 FROM PRECLINICAL EXPERIMENTAL PROTOCOLS TO CLINICAL STUDIES

The ability of the in vitro models to quantitatively predict drug effects at therapeutic concentrations remains often limited based on available methodology. On the other hand, the developability criteria applied during the preclinical evaluation of novel molecules supposedly reflects torsadogenicity without further evidence of the concentration or exposure range at which drug effects occur. The danger is under-estimation of the rate of false positive results, which lead to termination of perfectly efficacious compounds. In fact, the ICH S7B suggests that in vivo assessment of the drug effects on QT-interval duration, performed in addition to a hERG channel
Conclusions and Perspectives

assay, should form the minimal basis for the evaluation of pro-arrhythmic risk. However, the utility of non-clinical assays remains questionable. In order to establish *in vitro*-in *vivo* correlations, non-clinical studies must be calibrated against clinical data. The use of concentration-effect relationships as benchmark for translational research across species has not been applied before, as proposed in this thesis. In *Chapter 4* the same Bayesian hierarchical model developed previously to describe QT-interval prolongation in clinical studies was applied to preclinical data to describe the relationship between drug concentration and QTc-interval prolongation for cisapride, *d,l*-sotalol and moxifloxacin in conscious dogs, establishing their correlation with the pharmacological effects in healthy subjects at comparable exposure ranges. Based to the explicit distinction between drug-specific and system-specific parameters, a model-based analysis of pharmacokinetic and QT-interval data in dogs allowed inferences about the probability of drug-induced QTc prolongation in humans. The drug-specific parameter in the model (slope) did not independently indicate whether there was risk associated with the use of the compounds, nor did it indicate if their development should be stopped. Rather, it provided an unambiguous description of the PKPD relationship, taking into account various sources of variability.

Within the drug development framework, our approach appears to overcome one of the main limitations of earlier PKPD modelling efforts in preclinical research, i.e., it offers the opportunity to distinguish drug-induced effects in a generic manner, enabling the use of the model for prospective evaluation of novel compounds. Models should be used as a tool for predictive purposes rather than only focusing on model parameter estimation. In contrast to recent examples of model-based analysis of QTc-interval prolongation [12-15], we were able to correlate preclinical and human effects in a direct, quantitative manner. Based on the evidence obtained from the evaluation of paradigm compounds, it appears that one of the most important factors for accurate translation of preclinical findings to the effect in humans is the ability to evaluate drug effects across a clinically relevant exposure range. Inferences from extremely high dose levels, often used in toxicology experiments may lead to biased conclusions about the magnitude of the effect on QT-interval and its clinical implication in humans. In addition, our investigation showed some critical shortcomings in current experimental protocols which prevent effective use of model-based approaches in early drug development.
12. 3. 2 FROM CONTROLLED CLINICAL TRIALS TO TREATMENT OUTCOME IN REAL-LIFE CONDITIONS

We have shown that translational efforts are required not only during the progression of a compound from preclinical to clinical phases of development. It is also critical to understand the safety profile of a drug in the overall target population, as during Phase IIb, III and IV. In Chapters 5, 6 and 7 focus is given to the characterisation of the differences between the expected magnitude of drug-induced QTc prolongation and the overall increase in QTc-interval in real-life conditions. In addition, we highlight some important practical and ethical hurdles associated with the generation of safety data, particularly when safety concerns regard low frequency or rare events. Such information can only be derived from epidemiological studies, which are planned and performed only after drug approval. In Chapter 5, we have shown that population selection and inclusion/exclusion criteria applied to clinical protocols in early drug development may lead to significant differences between drug-induced and overall treatment effect size in the target population. Thus far, the evidence of pro-arrhythmia associated with QTc prolongation in real life patients has been primarily linked to drug-induced effects, disregarding the contribution of other important intrinsic and extrinsic determinants of response. Thanks to the use of a model-based approach it was possible to make inferences about the effects of \textit{d,l}-sotalol in patients and evaluate in an integrated manner how different covariates and sources of variability contribute to the observed QTc values in real life patients. Our findings revealed that the distribution of observed QTc values in the real life cohort cannot be explained by the effects of \textit{d,l}-sotalol alone. Other causal factors are present, which significantly affect the observed QTc values. In fact, the binary logistic regressions performed on our data revealed that diabetes, heart failure, arrhythmia, hypertension and myocardial infarction are indeed risk factors for QTc prolongation in this population. In addition to comorbidities, we have also identified statistically significant effects of concomitant medication, both of which contribute to upper tail of the distribution of observed QTc values.

This chapter illustrated how both clinical trial and epidemiological data can be used in an integrated manner for the purpose of signal detection and improved risk management. By applying pharmacokinetic-pharmacodynamic modelling concepts to epidemiological data, it was demonstrated that assessment of the concentration-effect relationships in
Conclusions and Perspectives

healthy subjects has limited value unless one can account for the effect of influential covariates in the target population. By extension, these findings clearly raised questions about the validity of TQT trials as the state-of-the-art approach for determining the risk of pro-arrhythmia or TdP. Indirectly, it also revealed the shortcomings of the double-delta method currently suggested for the TQT studies. The use of time-matched differences between active and placebo arms do not reflect the potential QTc prolongation in real-life conditions. Another advantage associated with the use of a PKPD model was that even if the pharmacokinetics are evaluated at supra-therapeutic levels, drug exposure can be extrapolated to reflect drug concentrations across therapeutic doses, so that accurate inferences can be made of the clinical relevance of the effects on QT-interval.

Our approach also revealed a flaw in the assumption that findings about drug effects in a clinical trial may be generalisable to the target population under real-life conditions. The requirements for the evaluation of drug-related QTc-interval prolongation are presented in Chapter 6, where we have reconciled the differences found between clinical trial population and real patients and assessed the contribution of different factors by simulation methods. The concept of Not-In-Trial Simulation is introduced under the assumption that, QTc-interval (real life population) = age-dependent baseline QTc values + drug-induced fluctuation in QTc + effects of concomitant medications and co-morbidity conditions. We also anticipate the importance of such an approach to assess the characteristics of high-risk population subgroups by introducing such patients into simulation scenarios. This approach can provide more accurate estimates of cardiovascular risk and enable mitigation measures, preventing potential withdrawals due to unexpected cardiovascular safety findings.

This investigation also shed light onto important methodological aspects associated with the assessment of causality, enabling further understanding of the potential differences or mismatch in the safety profile of drugs pre and post-launch. Firstly, present safety trial designs are highly efficacy-oriented. Secondly, the number of patients exposed to the drug in the preapproval phase is not large enough to detect a relatively rare but potentially fatal risk. Thirdly, many subgroups of patients most of whom show higher risk of TdP are usually excluded from clinical trials. These subgroups included: females; the elderly, those with predisposing cardiac or non-cardiac diseases associated with diminished repolarisation reserve (and therefore greater
susceptibility to prolongation of the QT-interval); those with pharmacogenetic defects of drug metabolising enzymes or pharmacological targets such as potassium channel polymorphism; those susceptible to bradycardia or electrolyte imbalance; or those receiving drugs with a potential for pharmacokinetic or pharmacodynamic interactions. Furthermore, our results revealed the effect of age on baseline QTc-interval, a correlation which plays an important role in the observed QTc values in the target patient population, irrespective of treatment.

To assess the robustness of the inferences derived from Not-In-Trial Simulations, a new paradigm compound, cisapride, was evaluated in Chapter 7. Rather than limiting pharmacovigilance research to “retrospective” meta-analytical approaches on empirical data from long term safety (i.e., from clinical practice as well as from randomised clinical trial), we showed once more that model-based approaches enable “prospective” evidence synthesis. Using hierarchical models, it is possible to integrate fixed effects (drug-related and disease related factors) as well random effects (including parameter and model uncertainties). Model parameters can subsequently be used to explore relevant scenarios. QTc measurements in cisapride users were accurately described, especially the extremely high values.

The aforementioned results were considered as part of an external validation procedure for the generalisation of the concept of Not-In-Trial Simulation as a framework for signal detection and risk management in clinical development. Predictions of QTc values in real life were accurate for two compounds with very different mechanisms of action and therapeutic use. Two major determinants of response were proven to be drug-independent, namely the effect of age and comorbidities on baseline values of QTc-interval. Although the work was limited to QTc-interval prolongation, we anticipate that the concept could be applied for any adverse event. It can also be used to discriminate confounding factors, in particular when drug-drug interactions or drug-disease interactions occur which are compound-specific and may not have been identified yet.

From a methodological point of view, this framework offers advantages compared to meta-analysis where the resulting model is primarily descriptive of a large, pooled dataset. Based on Not-In-Trial Simulation scenarios, investigators will have the opportunity to explore hypothetical populations and test different doses which are not confined to the observed
data. The main limitation to our approach at this point of development is the absence of random effects in drug exposure, which is inferred from the prescribed dosing regimens. The inclusion of random effects (between-subject variability) would further improve the assessment and predictability of the exposure-effect relationship in real-life conditions. Further work and different experimental protocols in pharmacovigilance studies will be required to better explain the role of pharmacokinetic variability.

12.3.3 Prognostic value of QTc-interval prolongation as biomarker of risk

The primary interest and intention in investigating the QT/QTc prolongation of new compounds were to reduce the prevalence of drug-induced arrhythmias or torsade de pointes (TdP) by non anti-arrhythmic drugs. By extension, the rationale was aimed at minimising the occurrence of sudden cardiac death (SCD) due to the putative association between prolonged QTc-interval, arrhythmias (TdP) and SCD. Yet, there were no consistent data demonstrating the link between QTc-interval prolongation and SCD. Therefore in Chapter 8, we undertook comprehensive evaluation of the prognostic value of QTc-interval prolongation using 182 cases of SCD in the Rotterdam Study cohort. In this study, we showed that the incidence rate of SCD is better correlated with high absolute baseline QTc-interval values. A prolongation by $\geq 10$ msec did not significantly affect the incidence rate of SCD. These findings were in agreement with data showing that the relative risk for arrhythmic death is higher when resting heart rate is greater than or equal to 84 bpm, and that TdP is rarely associated with QTc-intervals less than 500 msec.

These results raise further questions about the implications of the use of QT-interval prolongation as a surrogate for cardiovascular safety risk. In fact, from the sensitivity analysis performed in our study, SCD incidence was found to be independent from the changes from baseline or time elapsed since last ECG measured and death. This implied that using the absolute QTc values to determine the risk of SCD may be relatively more reliable. Most importantly, our findings contradicted the hypothesis underlying the current requirements by ICH topic E14 guideline on the evaluation of cardiovascular risk for non antiarrhythmic drugs, in which safety is defined by a threshold for an upper interval delta prolongation value of 10 msec. The association between delta 10 msec change in QTc and SCD was not found to be
statistically significant. We envisage therefore that two important concepts will need to be addressed in future revisions of the ICH E14 guideline: the rationale for a threshold of 10 msec and the use of changes from baseline (delta QTc values) to define risk or liability. Based on our results, it can be concluded that drug-induced changes in QTc may not increase the cardiovascular risk unless the underlying baseline QTc values are high.

Whilst it might not be feasible to establish a cut-off range for QTc-intervals without taking into account additional factors contributing to an increase in the risk of SCD, it seemed evident that clinically effective risk management policies should consider baseline QT-interval as a strong prognostic factor for pro-arrhythmias. The assumption of a comparable risk of TdP across the population for an increase in QTc-interval of greater than or equal to 10 msec could not be sustained.

12. 4 THE ROLE OF CLINICAL TRIAL SIMULATIONS IN PROTOCOL OPTIMISATION

12. 4.1 GENERATING RR VALUES AND QT-INTERVALS

The assessment of a drug's effect on the QT-interval must be corrected for changes in heart rate, as it also affects the QT-interval. The shortening effect of heart rate on the QT-interval must be taken into consideration when evaluating possible drug-induced effects. In fact, the relationship between the QT-interval and heart rate can be modelled mathematically, yielding a corrected QT-interval (QTc) which enables discrimination between drug-induced from heart rate-dependent effects.

Effective implementation of clinical trial simulations (CTS) for the assessment of drug-induced QT-interval prolongation requires therefore individual heart rate profiles. In contrast to re-sampling techniques from existing data, in Chapter 9, we developed a model describing the relationship between QT and RR over 24 hours. Our approach enabled generation of realistic RR profiles according to typical protocol sampling schemes, which can be used as input into a model describing the QT-RR relationship. Using the newly proposed method to generate QT/RR values, measurements were taken at discrete intervals according a study protocol, yielding between-
measurement variability which would not be described by beat-to-beat variation or by other previous approaches available in the published literature. By contrast, we have chosen to simulate data in which correlations and variance structure reflect the noise due to measurement and sampling procedures. This novel approach was essential for the evaluation and optimisation of the clinical protocols discussed in Chapters 10 and 11, where hypothetical populations were simulated to explore type I and II error rates using different study designs and methods of analysis.

12. 4. 2 CONCENTRATION-EFFECT RELATIONSHIPS VERSUS DOUBLE-DELTA METHOD

In Chapter 10, TQT study protocols were simulated with varying number of subjects and putative drug-induced QT effects. Different statistical designs, including crossover with full day time-matched baseline, crossover with three pre-dose baseline measurements and parallel designs were considered for the purpose of this investigation. It was demonstrated that the use of a Bayesian hierarchical model has higher or equal specificity and sensitivity compared to the double-delta method. Results strongly suggested that concentration-effect relationships should be considered when analysing the propensity of drug-induced QTc-interval prolongation as opposed to the currently recommended double-delta method. Most importantly, the proposed approach does not only enable accurate estimation of the effect size across the therapeutic exposure range, but it is also more sensitive to detect a safety signal. Conclusive results can be obtained about the liability for QT prolonging effects with a significantly smaller number of subjects. Keeping in mind that the high cost of TQT studies have mainly been attributed to the trial size required, this represents a major opportunity to address ethical, statistical and financial considerations associated with the assessment of a QT-interval prolongation of 5 msec. Another advantage of the use of a model-based approach is the possibility to handle multiplicity issues arising from having to establish non-inferiority at each time point. Instead, drug induced effects are established as a continuous function across the range of observable concentration, rather than at each discrete sampling time. Thus, these encouraging results gave rise to an alternative to the current methodology used in TQT studies as well as a strategy to minimise the cost of a such a trial.
Given the considerable increase in the sensitivity of a model-based approach to detect drug-induced QTc-interval prolongation, in Chapter 11 we have also explored the possibility of using the first-time-in-human (FTIH) study as basis for the assessment of QT liability. In conjunction with evidence of the concentration-effect relationship and the corresponding probability of an increase in QTc-interval in pre-clinical species, it can be envisaged that accurate decisions can be made about the relevance of QT prolonging effects. In addition, such an integrated approach overcomes the flaws in the current execution of TQT studies, which does not necessarily guarantee the accuracy and precision of the effect size. Another advantage to using the proposed Bayesian modelling approach to analyse the proarrhythmic properties of a compound is the possibility to describe its QT-prolonging effect over the therapeutic range. Since FTIH studies represent a mandatory step in clinical development, minor adaptations are required to ensure appropriate sampling schemes are used to collect ECG signal. This could in turn support the strategic decision to proceed or not with development. Furthermore, the results of the analysis demonstrated that accurate estimates of effect size of QT prolongation can be obtained without the inclusion of a positive-control arm. Yet, we do not exclude the possibility of including historical studies with moxifloxacin as parameter priors or benchmark during data analysis.

12.5 RECOMMENDATIONS

Clear shortcomings from various aspects of the ICH E14 guideline have been identified and alternative methods and recommendations have been proposed in this thesis. First, a Bayesian model describing the underlying concentration-QTc prolongation relationship was introduced to describe the actual drug-induced effects over a relevant concentration range. This method should be used instead of the statistically-based double-delta method. This implies that a supra-therapeutic dose is not necessary but even if used, the therapeutic effect can easily be derived using the hierarchical model. Furthermore, the use of a Bayesian approach also showed superior accuracy and precision in predicting actual QTc prolongation over the entire time course of observations, not just at specific time points. Another advantage to using the model-based approach is its ability to describe data across compounds as well as between different species. Beyond predictions, the availability of posterior distributions enables the assessment of the
probability for any pre-defined increase in QTc-interval. By utilising the same model in the analysis of a FTIH trial with enriched ECG monitoring, it is possible to determine the proarrhythmic propensity of compounds without embarking on a TQT study, which implies go/no-go decisions can be made earlier in the drug development cycle.

12.6 PERSPECTIVES

The results presented in this thesis highlight the need for a new paradigm for the assessment of cardiovascular safety, in which evidence generation does not prevail over the role of evidence synthesis, which is by far more encompassing and informative. We also reiterate the importance of revisiting current guidelines, in particular the ICH E14 document, which shows inconsistency with oncoming evidence from scientific research.

Clinical researchers, health care providers, regulatory authorities and policy makers have to realise the implications of experimental design to the generation of evidence under controlled conditions, as in the case of randomised clinical trials. Differences in the patient population during drug evaluation (pre-launch) and clinical use (post-launch) cannot be overlooked. There is a clear contribution of intrinsic and extrinsic factors or covariates determining treatment outcome, which alter, counteract or mitigate drug-related effects. Hence, the crucial question from regulatory perspective is “How efficient and reliable are the pre-approval clinical trials in identifying the clinical risk of TdP, given differences in patient population enrolled, background noise arising from spontaneous intra-individual variability in QTc-interval and the relatively low frequency of the clinically significant drug-induced effect?” [16].

As suggested by the title of this thesis, there is a gap in the translation of safety findings from clinical trials to real-life conditions which must be accounted for in risk management. During clinical development one needs therefore to focus on evidence synthesis and ensure accurate assessment of the underlying pharmacokinetic-pharmacodynamic relationships. Modelling and simulation are essential tools to close this gap. We envisage that further advancements in the prediction of cardiovascular safety can be obtained by expanding the concepts of Not-In-Trial Simulation to multiple endpoints as well as by incorporating mechanistic models to the pharmacometric
framework proposed in this thesis. Given the increasing availability of pharmacovigilance cohorts, access to data will not be the limiting step any longer. In fact, the translational concepts developed here may be applied to a wider range of outcomes, enabling drug developers, clinical researchers and regulators to bridge the gap between controlled clinical trial and real life conditions. In any case, the ability to discriminate between drug- and system-specific parameters will remain critical for the success of a model-based approach.

An additional concept has emerged from our research, which can provide the basis for further integration of clinical and pharmaco-epidemiological data. The concept of Not-In-Trial Simulations (NITS) has been introduced as inferential tool to evaluate the implications of strict inclusion / exclusion criteria imposed on the study population, and the relatively short duration of trials. These constraints make it impossible to detect infrequent or rare safety signals until the drug has been on the market for a longer period of time. The ability to anticipate such signals prior to launch offers the opportunity to explore and implement mitigation measures.

The idea behind NITS is to include the physiologically-relevant factors along with drug-induced effects and explore the contribution from real-life scenarios such as comorbidities and concomitant medications. This work is only possible with combining pharmacological and epidemiological concepts and data. In his publication, Black highlighted “the false conflict between those who advocate randomised trials in all situations and those who believe observational data provide sufficient evidence needs to be replaced with mutual recognition of the complementary roles of the two approaches.” [17] Others have also advocated the synergistic potential for using both kinds of data to aid decision making.

Regulatory and clinical efforts regarding the effective use of drugs presently rely on the concept of positive benefit-risk ratio and as such depend primarily on the availability of oncoming evidence from clinical or observational trials [18-21]. These are often defined by Phase IV long term safety trials or more recently by progressive licensing proposal, which ultimately imposes the burden on patients and drug developers, with the caveat of the having the evidence after the facts. Thus far, alternative methods to formalise prospective trials have been limited to “retrospective” meta-analytical approaches on data arising from empirical data, i.e., from
clinical practice as well as from randomised clinical trials. It is imperative that more efforts are put into how to devise forward-looking strategies.

The Bayesian approach is limited to improving the analytical aspect to detect QTc-interval prolongation of a compound and thus is only used to assess the propensity of proarrhythmic properties. It does not, however, solve other issues in TQT studies of practical or statistical nature with ECG measurements. Figure 3 summarises the potential areas of concern where significant impact can be made to the observed QT-interval data. These include the algorithm used for ECG recordings, the technique selected to correct measurement bias, the nature of the disease that influence ECG values, and various other covariates that can impact the recordings.

**Figure 3. Various factors that affect QT-interval measurements. A combination of contributors should be taken into considerations when assessing QT/QTc-interval prolongation including intrinsic and extrinsic characteristics [22].**

Much research is currently performed to address the precision of QT measurements including the use of automated readings. Efforts are also being made to address statistical issues, identifying sources of variability, new algorithms as well as the issues of reproducibility. In 2001, FDA established a digital ECG Warehouse to store ECG recordings received from
submission. Beyond its official use, the Warehouse also provides a great opportunity for pharmaceutical companies and academia to conduct research. It is an asset that can be used to study new biomarkers of proarrhythmia, correlate them with clinical and laboratory cardiac safety outcomes, improving the diagnosis and treatment of cardiovascular diseases, determine the underlying physiological variation, and use as prior distributions for future Bayesian PKPD analysis for QTc-interval prolongation.

Ultimately, we must "mind the gap" between controlled clinical trials and real-life conditions. Time has come to address these issues by quantifying, predicting, interpreting and eventually mitigating these inherent differences. It cannot be stressed enough that a modelling and simulation-based strategy for risk assessment and management can now be developed.
REFERENCES


