The handle [http://hdl.handle.net/1887/19939](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
Scope And Intent Of Investigation
2.1 BACKGROUND

The evidence of dromotropic and arrhythmogenic properties of new chemical entities (NCEs) remains a major cause of attrition in drug development. The pro-arrhythmic properties of a new chemical entity (NCE) are investigated throughout the drug development cycle, from the early in vitro screening phase all the way to clinical trials and post-launch surveillance. At the moment, each and every single study conducted pertaining to a compound is performed independently. Experimental evidence is not generated with the objective to characterise concentration-effect relationships, nor to enable integration of the results from preceding stages as the basis for predicting or extrapolating findings in a subsequent stage of the development process. The lack of data integration and the absence of accurate markers and measures of drug effect, which provide direct evidence or support the use of inferential methods for the assessment of cardiovascular risk, represent additional contributing factors to the poor decision making and inadequate go/no go criteria [1].

On the other hand, drug developers and investigators are faced with a mandatory step in clinical development which imposes the implementation of a so-called thorough-QT (TQT) trial. The International Conference on Harmonisation (ICH) E14 guideline was meant to inform industry on how to collect, interpret and analyse QT data in a TQT trial and thereby promote a more systematic process for assessing the proarrhythmic risk of NCEs. The TQT trial is considered negative if the 95% upper interval of the largest time-matched QTc measurements, does not include 10 msec after adjusting for baseline and placebo effects [2]. Among other issues, the approach neglects the underlying pharmacokinetic-pharmacodynamic relationships. Furthermore, a significant association or correlation between 10 msec prolongation and fatal arrhythmias has not been established so far. Instead, it appears that prolonged QT/QTc-intervals, defined as an interval greater than 450 msec in males and 470 msec in females is indeed associated with cardiac mortality. Regulatory agencies and other stakeholders do not seem to have assessed the implications of such findings for drug development, nor their relevance for the evaluation of cardiovascular safety.

In this thesis we will apply a model-based approach to the evaluation of cardiovascular safety of a NCE. More specifically, we will critically assess the requirements for data integration during development and life-cycle
management with the scope of facilitating the evidence synthesis for pro-arrrhythmic effects. To this purpose, focus will be given to the relevance of pharmacokinetic-pharmacodynamic relationships and to the accurate identification of drug- and system-specific parameters as basis for the translation of findings from early development stages into real-life conditions [3]. An important aspect of this translational investigation will be the identification of decision criteria and requirements for the progression of a compound into late-stage clinical development.

Whilst data integration and the development of new technologies have been the subject of extensive discussion among stakeholders [4], little attention has been paid to the need for revised (model-based) criteria supporting the decision making process. Of particular relevance is the ability to derive quantitative information about drug effects in such a way that parameters, rather than experimental variables can be used to assess the implications of such effects for the target patient population. We envisage that if inherent physiological differences across species or experimental conditions can be parameterised independently from drug-specific effects, it will be possible to accurately assess and predict risk in humans.

We aim therefore at evaluating how to optimally integrate data and explore model parameterisations that prove effective for the translation between species and experimental protocol conditions. We will confront the current regulatory perspective that imposes formal evidence of safety based on the results of a thorough-QT study [5] and will demonstrate that accurate assessment of safety should rely on the evidence from pharmacokinetic-pharmacodynamic relationships and on the quantification of other relevant intrinsic and extrinsic determinants of variability in drug response. Moreover, we will attempt to demonstrate the flaws in the assumption that evidence from clinical trials in healthy subjects using supra-therapeutic doses suffices in characterising safety. Clearly, one overlooks the implications of differences in demographics as well as the impact of inclusion and exclusion criteria which are applied to Phases I, II and III trials. We will assess how demographic factors, comorbidities and concomitant medications affect outcome and consequently risk.

From a methodological perspective we will also show how modelling and simulation (M&S) techniques can be used for the evaluation of safety and risk, given the ethical constraints for human experimentation when dealing
with safety issues. As a matter of fact, we will defend the views that M&S should be considered as the technology of choice for risk assessment and risk management. In order to predict how the treatment effects perform in the overall patient population, clinical trial simulations (CTS) will be used to evaluate drug effects in various populations with different characteristics. Furthermore, CTS can be critical for the optimisation of trial protocols, enabling increased signal-to-noise ratio and facilitating decision making at early stages of drug development.

In summary, our hypothesis is that a model-based approach is required for accurate identification of pro-arrhythmic drugs and subsequently for the quantification of the dromotrophic effect and its consequences on cardiovascular safety. We will demonstrate that model parameterisation is a critical component for effective decision making. A model must discriminate between system- and drug-specific properties to allow prediction of drug performance in real-life population. Specifically, this thesis will evaluate existing data from compounds with known cardiovascular effects to address the following issues:

1. How modelling and simulation (M&S) can facilitate the translation of studies across different phases of drug development, improving the predictive value of experiments.

2. How historical data can be used to support model parameter estimation and subsequently applied as priors during the evaluation of drug safety.

3. How concentration-effect (PKPD) relationships can be used as basis for the analysis of QTc-interval prolongation.

4. How PKPD concepts can be used in conjunction with pharmacoepidemiological data to discriminate drug-induced effects from other confounders in the target patient population.

5. How decision criteria can be optimised for the evaluation of novel compounds.
Given the breadth and complexity of the subject, the focus of our investigation is split into three main sections. After a general introduction, we explore the technical requirements for model parameterisation and define the benchmark for the analysis and interpretation of risk during the evaluation of cardiovascular safety. Subsequently, focus is given to a bridging strategy, which 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.

2.2 GENERAL INTRODUCTION

In Chapter 1, an overview of the current technologies available for assessing QT/QTc-interval prolongation is presented, which provides insight into some of challenges for translation and prediction of drug effects in humans. Limitations and flaws are highlighted and used as a foundation for this thesis. Furthermore, an overview is given of the physiological, clinical and epidemiological basis for the characterisation of drug-induced effects on cardiac conductivity, more specifically on QTc-interval, together with the theoretical basis for the measurement of QT/QTc-interval and its association with fatal arrhythmias. This is followed by a discussion on the various methods used to establish correlations between observations in clinical trials and in the target patient population under real-life conditions. We emphasise that not all drugs that prolong QT-interval to the same extent carry the same risk of causing TdP. It is therefore important to appreciate that QTc-interval prolongation may not constitute a direct risk, as it could be inferred from some of the clinical findings. Although QTc-interval prolongation is one of the major precursors of drug-induced TdP, this arrhythmia does not develop invariably in all individuals with equivalent prolongation of the QTc-interval [6]. Neither do all drugs that prolong the QT-interval to an equivalent duration carry the same risk of inducing TdP. Clearly, unwanted drug-induced prolongation presents an important concern for drug development, thus the general progression of the publications of guidance for industry is also ensued.

Lastly, we introduce important methodological concepts pertinent to the evaluation of exposure-response relationships and consequently useful for the prediction of safety. In recent years, the use of Bayesian methods has
increased considerably as a tool for the characterisation of exposure-response relationships and other medical applications [7]. The key difference between classical and Bayesian statistics is the interpretation of probability. In a Bayesian context, probability is considered to be a quantification of ‘degree of belief’. We take advantage of this concept as the basis for the technical implementation of the model-based approach proposed throughout this thesis. In contrast to traditional statistical reasoning, which relies on the likelihood, i.e. the probability of observing the data given a range of parameter values, Bayesian statistics relies on the posterior distribution, which is the distribution of the parameters (i.e., parameter space) given the observed data. The prior distribution can include all relevant information that is not captured by the actual experiment.

Based on a comprehensive review of the scientific rationale and of evidence from clinical and epidemiological data, specific issues have been identified in the current strategy for the assessment of pro-arrhythmic properties and QTc-interval prolongation which will underpin the scope of this thesis, as detailed here in Chapter 2.

2.3 BAYESIAN HIERARCHICAL MODELS IN SAFETY ASSESSMENT

As indicated previously, the lack of data integration and of integrative approaches for decision making contributes to the attrition rate in drug discovery and development. A streamlined approach is required which takes into account the concentration-response relationship across different phases of development, irrespective of the species, populations or experimental conditions. The evidence of risk of hazard or lack thereof must have predictive validity or value for the target population, both in qualitative and quantitative terms. Such considerations are not formally embedded in the requirements for evidence generation, and consequently lead to the implementation of experimental protocols which are not fully informative or eventually even biased. Whilst model-based approaches have been used across numerous therapeutic areas to assess efficacy and safety, most attempts have been descriptive in nature, without clear focus on the prospective use of models as benchmark or strategy for drug screening and decision-making [8]. The possibility to dissect system from drug-specific features allows the introduction of modularity, which is essential for the assessment of drug-induced effects and risk management. Consequently,
here we propose a general but robust parameterisation which enables integration of historical data into decision making. Moreover, it allows a shift in the efforts usually required for model building into relatively simple parameter estimation procedures. These procedures also provide the basis for validation of screening tools in terms of sensitivity and specificity (i.e., false positive and false negative rates). We also show that decision making requires more than hypothesis testing, as currently assessed during a TQT trial. In fact, assessment of the liability of a compound to prolong QT/QTc-interval requires a well-defined, integrated set of assumptions and scenarios which together describe the observed findings not only in statistical terms (i.e., effect size), but ultimately are predictive of yet-to-be observed events, providing insight into the mechanisms underlying the experimental data.

This concept is introduced in Chapter 3, where a general model parameterisation is proposed in terms of drug-specific and system-specific properties. Using clinical data from Phase I trials in healthy volunteers, we demonstrate the feasibility of a Bayesian approach to characterise the exposure-effect relationship and the corresponding probability of QTc-interval prolongation of three compounds known to prolong the QT/QTc-interval (d,l-sotalol, moxifloxacin, grepafloxacin). The proposed model consists of an individual correction factor for heart rate, an oscillatory component describing the circadian variation and a truncated E$_{\text{max}}$ model to account for drug effect. Probability curves for an increase in QTc-interval $\geq$ 10 msec are derived, which can be easily interpreted for clinical purposes.

From a statistical perspective, in Chapter 3 we also draw attention to the implications of false-positive and false-negative results and emphasise the importance of a parametric approach to overcome high type I error rates, without the burden of traditional statistical methods that impose therefore large study populations. In addition, our approach readily allows for the incorporation of prior knowledge, which is abundant from historical data on system-related parameters describing e.g., QT/RR relationship and circadian variability. This approach also allows for the estimation of the posterior parameter distribution, which fully reflects all acknowledged sources of uncertainty.

The advantages of model parameterisation based on drug- and system-specific properties is further explored in Chapter 4, where we illustrate the suitability of the model to describe the pro-arrhythmic effects of d,l-sotalol,
cisapride and moxifloxacin. Data generated pre-clinically and clinically in dogs and in humans, respectively, is analysed using the same hierarchical model, enabling the assessment of correlations between species. We anticipate that the use of a common model during the course of drug development can facilitate the translation of data from pre-clinical to clinical conditions. From a drug development perspective, these features are highly relevant, as they allow prospective evaluation of compounds.

2.4 FROM PHARMACOLOGY TO EPIDEMIOLOGY: A BRIDGING STRATEGY FOR THE EVALUATION OF CARDIOVASCULAR RISK.

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 observed during Phase IIb, III and IV trials. Although epidemiological evidence suggests a correlation between drug-induced QTc-interval prolongation, the risk of fatal arrhythmias and consequently sudden cardiac death (SCD), little has been done to pro-actively evaluate whether the signal-to-noise ratio observed during a clinical trial changes under real-life conditions. In fact, trials in early- and late-phase clinical development are primarily designed to assess efficacy. The detection of safety signals and adverse events are limited by the duration of the study, by the inclusion/exclusion criteria as well as by the size of the study population. Furthermore, no formal procedures exist to mitigate the impact of such differences or support the management of cardiovascular risk in the target population. Yet, many patient subgroups with higher risk factors are often excluded from the trials, but are ultimately part of the population receiving treatment after regulatory approval. It is, therefore, important to devise a strategy to anticipate the incidence of adverse events in these patients [9]. Based on clinical trial simulation (CTS) concepts, in Chapters 5, 6 and 7 we explore innovative techniques to evaluate safety and cardiovascular risk in the overall population.

In Chapter 5, 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 appear at low frequency or are rare events. Such information can only be derived from epidemiological studies, which are planned and performed only after drug approval.

Instead, an integrated approach is proposed that enables data from clinical trials to be analysed parametrically to predict the overall treatment effect in the general population. Once more d,l-sotalol is used as paradigm compound to illustrate the potential clinical implications of differences in safety signal before and after marketing authorisation. The Rotterdam Study cohort, an observational cohort established in a suburb area of Rotterdam, The Netherlands, is analysed in conjunction with the hierarchical model used previously to characterise QTc-interval in healthy subjects. From a statistical perspective, linear regression methods are combined with nonlinear mixed effects models. In contrast to using normalised daily dose, drug exposure in epidemiological cohorts is characterised by pharmacokinetic modelling. Inferences about individual exposure are based on the expected population parameter estimate and on covariate effects.

Finally, the concept of Not-In-Trial Simulation as a way to predict QTc prolongation during the post-marketing phase is established in Chapter 6, where we attempt to reconcile the differences found between clinical trial population and real patients. Furthermore, we show that despite the widespread acceptance of model-based techniques in clinical pharmacology research and in drug development, the use of simulations remains limited, in particular for the evaluation of safety data. Not-In-Trial Simulations are introduced as an inferential tool. This method offers the basis for improved risk management, dissecting drug-specific contribution to pre- or post-approval safety findings. In fact, we anticipate that simulations can unravel the individual contribution of baseline QTc-interval, drug exposure, concomitant medications and comorbidity conditions to the overall effect in patients. As such, Not-In-Trial Simulations offer an opportunity for proactive risk management.

Given the relevance of the predictive validity of the proposed hierarchical model for QTc-interval prolongation, in Chapter 7 the advantages of Not-In-Trial Simulations will be further assessed using cisapride as a reference compound. Model-predicted QTc-intervals are obtained taking into account differences in patient demographics, baseline QTc values, the exposure-effect relationship observed in a clinical trial population as well as the effects of
concomitant drugs and comorbidities. We endeavour to demonstrate that Not-In-Trial Simulations can provide a framework for the assessment of drug- versus disease-specific properties, facilitating the assessment of causality when describing or analysing safety data. The concept becomes increasingly important with less frequent or rare events, for which larger populations or longer trials are required. On the other hand, there are cases in which it is not the frequency of an adverse event, but its magnitude that cannot be accurately determined during the clinical development phase. In these circumstances the observed risk-benefit ratio may not reflect the actual implications of the therapeutic intervention in the target population. We consider QTc-interval prolongation as an example of such a case.

After having addressed some of the issues associated with the lack of data and demonstrated the translational and predictive value of a hierarchical model to characterise QTc-interval prolongation, the next question to be considered in the evaluation of cardiovascular safety is the relevance of a 10 msec increase and the meaning of such a threshold. Currently, a TQT trial is mandated before drug approval to ensure accurate assessment of the risk of pro-arrhythmias. The underlying assumption is that QTc-interval prolongation >10 msec leads to an increase in the risk of torsade de pointes (TdP). However, the evidence for such a correlation remains debatable. In fact, several large population-based studies exploring the association between the heart-rate corrected QT-interval and mortality in apparently healthy persons did not find a consistent association between the length of the QTc-interval and total or cardiovascular mortality [10]. This very notion is challenged in Chapter 8, where we investigate whether TdP and sudden cardiac death (SCD) are indeed correlated with each other in a cause-and-effect manner. Among other things, it is of interest to evaluate whether increases in QTc values, calculated as changes from baseline or absolute measured QTc values, are associated with SCD. We envisage that any meaningful attempt to further implement QT-interval as a prognostic marker of cardiovascular risk must take into consideration the basic operating characteristics of working variables, i.e., its sensitivity and specificity to detect the signal of interest. To this purpose, SCD data in the general population is obtained from the Rotterdam Study cohort. We apply trend-analysis to describe the correlation between patients showing changes in QTc of <10 msec and ≥10 msec. In addition, a sensitivity analysis is performed to explore the impact of the time span between death and last ECG measurement.


2.5 CLINICAL TRIAL SIMULATION AND STUDY DESIGN OPTIMISATION FOR THE EVALUATION OF CARDIOVASCULAR SAFETY

In this section of the thesis, we consider the advantages of protocol design optimisation as the basis for improved statistical inference and accurate estimation of drug-induced effects (i.e., treatment effect size) in a clinical trial. Different protocol optimisation scenarios are explored under the assumption of treatment with drugs showing varying degrees of QTc-interval prolongation. First, attention is paid to the confounding effects of heart rate, which also affects the QT-interval. Correction for heart rate is required to ensure a single physiological value within the same subject under different conditions. In fact, the relationship between the QT-interval and heart rate can be modelled mathematically. However, despite the accuracy and precision of the parameters describing the so-called physiological relationship, the use of such functions in clinical trial simulations has been limited. In Chapter 9, rather than relying on data resampling using existing pairs of QT-RR values, a novel approach is proposed to describe the correlation between QT- and RR-intervals in healthy volunteers. Based on simulations, a hypothetical population with physiological values of QT and RR values is used in Chapter 10 to explore and compare the sensitivity / specificity of the different methods currently available for the assessment of QTc-interval prolongation. Of particular interest are the differences between the proposed parametric approach, based on a Bayesian hierarchical model and the double-delta method, which is the current gold standard for the analysis of QT-interval prolongation. In addition, the impact of different design factors on type I and type II errors is considered for a variety of study scenarios in which hypothetical drugs with different effect sizes (namely, 10 msec for a positive compound, 5 msec for a borderline compound and 2 msec for a negative compound) are evaluated [11]. Ultimately, the objective is to demonstrate whether the use of concentration-exposure relationships can facilitate the assessment of causality and overcome the statistical issues caused by the double-delta method.

Given the statistical flaws, the ethical burden and financial consequences of the study design recommended by ICH E14, in Chapter 11 we also assess the feasibility of using first-time-in-human (FTIH) studies as the basis for evidence synthesis regarding the propensity for pro-arrhythmic effects.
FTIH studies are a mandatory step in the drug development process and include detailed information on drug exposure at therapeutic and supratherapeutic levels, enabling the characterisation of the concentration-effect curve. It is hypothesised that FTIH trials can be modified and optimised to deliver the same information currently obtained from thorough-QT (TQT) studies by increasing the sampling frequency and by applying a parametric method for the analysis of the data. We also challenge the requirement for continuous assessment of the individual sensitivity to QTc-interval prolongation by the inclusion of a positive control arm. In fact, historical studies with moxifloxacin can be used in a manner as priors during parameter estimation [12]. We anticipate that the possibility to modify traditional FTIH studies will provide invaluable safety information early in development [13]. Given that such modifications will consequently generate more informative data, it will facilitate and improve decision making.

2.6 CONCLUSION AND PERSPECTIVES

The final section of this thesis summarises the main findings and conclusions from the investigations presented throughout the various chapters. In Chapter 12, we continue to advocate the need for data integration and on evidence synthesis based on pharmacokinetic-pharmacodynamic relationships. We provide recommendations for cardiovascular safety evaluation based on the assumption that accurate inferences can be made irrespective of the level of evidence obtained during the development phases. A new paradigm must emerge in which evidence generation does not replace or overshadows the need for evidence synthesis, which is by far more encompassing and comprehensive. We also reiterate the importance of revisiting current guidelines, in particular the ICH E14 document, which shows inconsistency with oncoming evidence from scientific research.

Finally, 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.
REFERENCES


